

Physique des solides

(MP 022)

Table des Matières

1 Généralités sur l'état condensé de la matière ordinaire	1
1.1 Ordres de grandeur	1
1.2 Caractérisation des solides	2
1.2.1 Propriétés mécaniques	3
1.2.2 Propriétés structurales	3
1.2.3 Propriétés électriques	4
1.2.4 Propriétés thermiques	6
1.2.5 Propriétés magnétiques	7
1.3 Classification des solides ordonnés	8
2 Énergie de cohésion des solides ordonnés	11
2.1 Généralités	11
2.2 Cohésion des solides moléculaires	12
2.3 Cohésion des réseaux ioniques	15
2.4 Cohésion des solides covalents	18
2.5 Cohésion des métaux	18
3 Structures cristallines	23
3.1 Réseau de Bravais	23
3.2 Structure physique d'un réseau de Bravais	26
3.3 Quasicristaux	27
4 Réseau réciproque	29
4.1 Définition et propriétés du réseau réciproque	29
4.2 Exemples de réseaux réciproques	32
4.3 Plans réticulaires	33

5	Diffraction	37
5.1	Diffusion des rayons X et des neutrons	37
5.1.1	Rayons X	37
5.1.2	Neutrons	40
5.2	Diffraction par un réseau. Conditions de von Laue et de Bragg	41
5.3	Facteurs de structure (géométrique et atomique)	44
6	Propriétés générales du mouvement d'un électron dans un potentiel périodique	49
6.1	Préliminaires	49
6.2	Théorème de Bloch	51
6.3	Premières conséquences du théorème de Bloch	57
6.3.1	Première zone de Brillouin	57
6.3.2	Équation pour la fonction $u_{\vec{k}}$	58
6.3.3	Impulsion de l'électron dans le cristal	59
6.3.4	Vitesse d'un électron de Bloch	59
6.3.5	Symétrie de la fonction $\varepsilon(\vec{k})$	60
6.3.6	Surface de Fermi	61
7	Applications choisies du théorème de Bloch	63
7.1	L'approximation des électrons presque libres	63
7.2	L'approximation des liaisons fortes	73
7.3	Conducteurs, semi-conducteurs, isolants	78
8	Vibrations d'un solide ordonné	85
8.1	L'approximation harmonique	85
8.2	Modes normaux d'un réseau unidimensionnel	87
8.3	Modes normaux d'un réseau unidimensionnel à deux atomes par maille	93
8.4	Modes normaux d'un réseau tridimensionnel	99
8.4.1	Réseau sans base	99
8.4.2	Réseau avec base	100
8.5	Quantification des vibrations de réseau	101
8.6	Contribution des phonons à la chaleur spécifique	102
8.7	Manifestation des phonons dans les spectres de diffraction	103

Chapitre 1

Généralités sur l'état condensé de la matière ordinaire

1.1 Ordres de grandeur

Il s'agit ici de la matière "ordinaire", royaume de la physique des basses énergies (quelques eV), par opposition à la matière nucléaire par exemple. La densité caractéristique de particules par unité de volume, n , permet d'obtenir l'ordre de grandeur de la distance (moyenne) entre celles-ci :

$$d = n^{-1/3} . \quad (1.1)$$

Si ρ désigne la masse volumique, m est la masse d'un atome (molécule) et M la masse atomique, alors $\rho = nm$ et :

$$d = \left(\frac{m}{\rho}\right)^{1/3} = \left(\frac{M}{\mathcal{N}\rho}\right)^{1/3} . \quad (1.2)$$

La densité n peut s'estimer comme suit :

- Gaz

La masse volumique ρ est de quelques grammes par litre. Par exemple pour l'air, composé en gros de 1/5 d'oxygène O_2 et de 4/5 d'azote N_2 , une mole a pour masse :

$$\frac{1}{5} \times 32 + \frac{4}{5} \times 28 \text{ g} \quad (1.3)$$

et occupe 22,4 litres dans les conditions normales. La masse volumique est donc :

$$\rho_{\text{air}} = \frac{144}{5 \times 22,4} \simeq 1,3 \text{ g/l} . \quad (1.4)$$

Quant à la densité n de molécules, elle vaut :

$$n = \frac{\mathcal{N}}{22,4 \times 10^{-3}} \text{ m}^{-3} \simeq \frac{6 \times 10^{23}}{22,4 \times 10^{-3}} \text{ m}^{-3} , \quad (1.5)$$

soit $n \simeq 3 \times 10^{25} \text{ m}^{-3} = 3 \times 10^{19} \text{ cm}^{-3}$. Ceci, *via* (1.1), donne l'ordre de grandeur de la distance moyenne d entre particules :

$$d_{\text{gaz}} = n^{-1/3} = (30 \times 10^{24})^{-1/3} \text{ m} \simeq (3^3 \times 10^{24})^{-1/3} \text{ m} \sim 30 \text{ \AA} . \quad (1.6)$$

Le gaz est donc typiquement un milieu dilué, où la distance entre deux objets est très grande devant la taille de chacun d'entre eux.

- Liquides

La masse volumique ρ est cette fois de l'ordre du kilogramme par litre, soit 10^3 kg/m^3 . Pour l'eau, par exemple, une mole a pour masse 18 g :

$$\rho_{\text{eau}} \simeq n \frac{18 \times 10^{-3}}{6 \times 10^{23}} \sim 10^3 \text{ kg/m}^3 \iff n \simeq (1/3) \times 10^{29} \text{ m}^{-3}, \quad (1.7)$$

ce qui donne $n \simeq 3 \times 10^{28} \text{ m}^{-3} = 3 \times 10^{22} \text{ cm}^{-3}$ et une distance moyenne typique :

$$d_{\text{liquides}} = n^{-1/3} \simeq (30 \times 10^{-30})^{1/3} \text{ m} \sim 3 \text{ \AA}. \quad (1.8)$$

Maintenant, au contraire, la distance entre deux particules est du même ordre de grandeur que la taille de chacune d'entre elles ; un liquide est donc un milieu dense – mais fluide ! –, dont on peut se faire une image en disant que les particules y “roulent” les unes sur les autres.

- Solides

Pour le fer, par exemple ($M = 56 \text{ g}$), la densité est $7,8 \text{ g/cm}^3$, soit $\rho_{\text{fer}} = 7,8 \times 10^3 \text{ kg/m}^3$. On a donc :

$$\rho_{\text{fer}} = n \frac{56 \times 10^{-3}}{6 \times 10^{23}} \simeq 7,8 \times 10^3 \text{ kg/m}^3 \iff n \simeq (5/56) \times 10^{30} \text{ m}^{-3} \sim 10^{23} \text{ cm}^{-3}, \quad (1.9)$$

d'où :

$$d_{\text{solides}} = (56/5)^{1/3} \times 10^{-10} \text{ m} \sim 2 \text{ \AA}. \quad (1.10)$$

Bien sûr, cette distance est d'ordre atomique, et est un peu plus petite que dans un liquide.

Pour mémoire : il existe des liquides *solides*, ce sont les verres, ou plus généralement les matériaux dits amorphes. Ils tiennent des solides au sens où ils sont relativement rigides¹, mais leur ordre interne est celui d'un liquide : à courte portée, il porte souvent sur l'orientation de molécules, ou de groupements moléculaires (verres orientationnels). On désigne par *transition vitreuse* le passage du liquide au verre, au moment où la viscosité augmente vertigineusement alors que la température varie très peu. Le statut d'une telle transition, comparée à une vraie transition de phase, n'est pas vraiment élucidé.

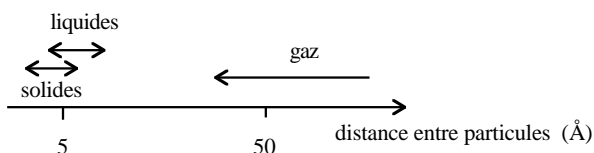


Figure 1.1: Distances typiques entre particules en matière condensée ordinaire.

Les densités estimées ci-dessus peuvent être comparées à celle de la matière nucléaire ; le proton (noyau d'hydrogène) et le neutron ont à peu près la même masse $M \simeq 1,6 \times 10^{-25} \text{ g}$ et le rayon nucléaire est $\sim 1 \text{ F} = 10^{-15} \text{ m}$, d'où une masse volumique² $\sim 1,6 \times 10^{14} \text{ g/cm}^3$! La densité de nucléons est, elle aussi faramineuse : un nucléon par Fermi³ correspond à 10^{38} nucléons par cm^3 ...

Dans toute la suite, s'agissant de la matière condensée ordinaire, on aura donc $d \sim$ quelques Å et on supposera des conditions également ordinaires (température, pression, champs externes, etc.).

1.2 Caractérisation des solides

Ceci étant précisé, les corps à l'état condensé peuvent être caractérisés par diverses propriétés.

¹Par comparaison avec un liquide, la viscosité d'un verre est infinie.

²Un dé à coudre “pèse” cent millions de tonnes !

1.2.1 Propriétés mécaniques

Ces propriétés permettent de rationaliser le comportement sous l'effet de contraintes mécaniques. On peut ainsi définir, par exemple, l'élasticité (déformation réversible), la plasticité (déformation irréversible), la dureté (concernant essentiellement la surface), *etc.* Ces propriétés dépendent fortement de défauts au sein du matériau (les dislocations³ par exemple) et c'est pourquoi il s'agit de disciplines en soi, souvent phénoménologiques, qui ne seront pas abordées dans la suite.

1.2.2 Propriétés structurales

Globalement, il s'agit de répondre à la question : comment la matière est-elle répartie dans l'espace ? Cette question repose inévitablement sur la notion d'ordre :

- Dans un cristal, l'arrangement est spatialement périodique et on parle d'ordre à longue portée (théoriquement infinie pour un cristal parfait). Cependant, en pratique, il existe toujours des défauts (lacunes, impuretés chimiques, dislocations, *etc.*)
- Dans un liquide, il existe un ordre, mais il est à courte portée
- Dans un gaz, le "chaos moléculaire" est la règle, et aucun ordre n'existe.

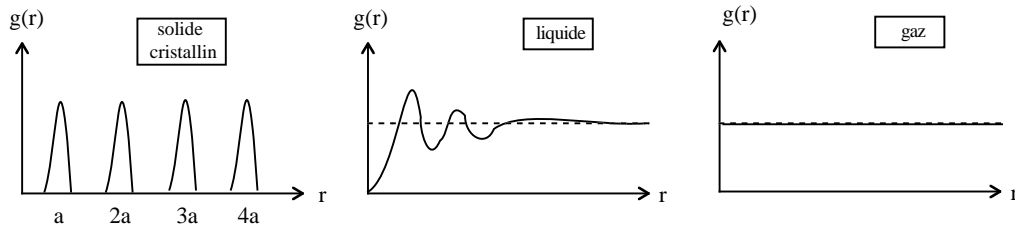


Figure 1.2: Fonction de corrélation $g(r)$.

On peut dire que l'ordre structural du liquide interpole entre le désordre absolu du gaz et l'ordre parfait du cristal. Plus quantitativement, l'ordre se mesure à l'aide de la fonction donnant la (densité de) probabilité de trouver deux particules en deux positions \vec{r}_1 et \vec{r}_2 , $g(\vec{r}_1, \vec{r}_2)$; pour un matériau d'extension infinie, g ne dépend que de $\vec{r} = \vec{r}_2 - \vec{r}_1$. C'est un peigne (édenté) de Dirac pour un cristal (il lui manque une dent, celle à l'origine) et c'est une constante pour un gaz ; pour le liquide, aux petites distances on retrouve une structure réminiscente de Dirac, puis g sature à une valeur constante aux grandes distances ; pour un gaz, où règne le chaos moléculaire, les atomes sont sans corrélation et la fonction de paire est simplement le produit des densités à une particule : elle est donc constante (avec un trou en $r = 0$). C'est bien pourquoi on peut dire que l'ordre du liquide interpole entre ces deux situations extrêmes (voir figure 1.2). On verra dans la suite que la fonction g est directement mesurable par des expériences de diffraction (X, neutrons, *etc.*)

Remarque

Il est naturel de se poser la question : pourquoi un corps, dans des conditions données, est-il plutôt dans un état (phase) que dans un(e) autre ? La réponse est fournie par la théorie des transitions de phase et plus particulièrement celle des phénomènes critiques. Il s'agit encore d'une discipline en soi, qui ne sera pas abordée, sauf sous la forme de brèves incidentes ici et là – notamment à propos des propriétés magnétiques.

³D'une façon générale, une dislocation est une ligne (ou un plan) de défauts qui rend le matériau plus vulnérable à une contrainte extérieure convenablement orientée.

1.2.3 Propriétés électriques

Il s'agit principalement de l'aptitude à conduire le courant électrique, que l'on mesure par la conductivité, σ ; celle-ci varie énormément d'un matériau à l'autre : il existe un facteur 10^9 entre la conductivité du cuivre et celle du germanium pur à l'ambiante, et un facteur 10^{19} entre celle du cuivre et celle du verre ordinaire. Il est rare qu'une même propriété physique varie dans un intervalle aussi gigantesque et ce fait permet de distinguer en gros trois classes de matériaux : les conducteurs, les semi-conducteurs et les isolants⁴. Cette distinction est expliquée par la Mécanique Quantique, et elle seulement : elle ne peut se décrire simplement en termes de propriétés élémentaires dans l'espace réel et impose de se placer dans l'espace des impulsions (espace réciproque), plus précisément d'examiner la dépendance de l'énergie d'un électron en fonction de sa (pseudo-) impulsion (loi de dispersion). En outre, elle résulte d'un jeu subtil entre ces lois de dispersion et la statistique quantique, étant entendu que les porteurs de charge (dans le cas du transport ordinaire) sont des électrons, c'est-à-dire des fermions. La structure géométrique du solide joue un rôle important en imposant une loi de dispersion ou une autre, mais la nature physique (structure électronique) des objets remplissant la structure a également un rôle déterminant. Globalement, la théorie des bandes fournit le cadre théorique permettant de comprendre pourquoi, quand les électrons interagissent peu, un matériau est conducteur ou non.

Ceci étant, il est toutefois facile (et utile) de se faire une idée un peu plus quantitative du transport du courant dans un métal, reléguant la question plus fondamentale de savoir pourquoi un matériau donné est ou n'est pas conducteur. Il suffit d'admettre qu'il existe un gaz d'électrons à peu près libres pour bâtir le modèle dit de Drude ; quoique très rustique, il permet de comprendre quelques faits majeurs et notamment la nécessité de défauts (au sens large) pour expliquer la loi d'Ohm⁵. Les défauts assurent la dissipation d'énergie sans laquelle il n'y aurait pas de loi d'Ohm.

Pour un électron de ce gaz, quand le métal est soumis à un champ électrique $\vec{\mathcal{E}}$, on a envie d'écrire l'équation fondamentale de la dynamique sous la forme :

$$m \frac{d\vec{v}}{dt} = e\vec{\mathcal{E}} . \quad (1.11)$$

Ceci produit un mouvement uniformément accéléré, donc une vitesse croissant linéairement en temps. Comme le courant est proportionnel à la vitesse *via* la densité de porteurs, on trouve un courant qui augmente avec le temps, en contradiction avec l'expérience. Il manque donc quelque chose dans l'équation (1.11), qui vient limiter la croissance de la vitesse et on devine qu'un frottement peut sauver la situation. De fait, si l'on incorpore une force de freinage – la plus simple est celle caractéristique du frottement fluide, proportionnelle à la vitesse –, on a maintenant :

$$m \frac{d\vec{v}}{dt} = e\vec{\mathcal{E}} - \alpha \vec{v} , \quad (1.12)$$

où la constante α est homogène à [masse]/[temps] ; en posant strictement $\alpha = m/\tau$, on a maintenant :

$$m \frac{d\vec{v}}{dt} + \frac{m}{\tau} \vec{v} = e\vec{\mathcal{E}} . \quad (1.13)$$

La solution de cette équation contient un régime transitoire – tributaire de la condition initiale (inconnue !) – de durée τ , suivie d'un régime à vitesse constante ; le transitoire est sans intérêt, en raison de la petitesse de τ (voir ci-dessous). En pratique, seul le régime forcé importe, caractérisé par la vitesse :

$$\vec{v} \simeq \frac{e\tau}{m} \vec{\mathcal{E}} \quad \forall t \gg \tau . \quad (1.14)$$

τ apparaît comme un temps de relaxation, au-delà duquel la condition initiale est oubliée. Physiquement, cette relaxation provient de “collisions” inélastiques⁶ au sens large (collisions avec les ions du réseau, avec des défauts,

⁴Ces énormes différences d'ordres de grandeur traduisent les différences profondes de la nature du transport de l'électricité dans ces divers solides. Pour un métal, le transport est possible grâce à la présence d'électrons presque libres, et la conductivité *décroit* avec la température : plus la température est élevée, plus les ions du métal vibrent avec une grande amplitude et plus les électrons sont freinés. Pour un semi-conducteur intrinsèque – dont la résistivité *décroit* avec la température –, les porteurs de charge sont encore des électrons, mais ils doivent être activés thermiquement pour franchir les (petits) gaps d'énergie entre la bande de valence et la bande de conduction. Enfin, dans le cas du verre, ce sont des ions qui transportent le courant ; la diffusion d'espèces très massives permet à elle seule de se douter que la résistivité est – même toutes choses égales par ailleurs – extrêmement élevée.

⁵D'ailleurs, la formule de la conductivité selon Drude, formule (1.16), est “presque” exacte : on verra par la suite que, dans les cas usuels, la modification essentielle (et importante) consiste à remplacer la masse m (masse nue) de l'électron par sa masse effective.

⁶Il y a bien de l'inélasticité quelque part, puisque (1.13) entraîne que l'énergie mécanique de l'électron n'est pas conservée.

interaction avec les vibrations, *etc.*) ; le champ accélère l'électron, mais chaque collision réduit cette vitesse. Au total, la vitesse sature en moyenne à une valeur finie constante, indépendante de la condition initiale⁷. Maintenant, s'il y a n_e électrons par unité de volume, le courant électrique résultant de la présence du champ est :

$$\vec{j} = n_e e \vec{v} = n_e \frac{e^2 \tau}{m} \vec{\mathcal{E}}, \quad (1.15)$$

ce qui permet, par $\vec{j} = \sigma \vec{\mathcal{E}}$, d'obtenir la conductivité en champ constant⁸ (fréquence nulle), σ_0 :

$$\sigma_0 = n_e \frac{e^2 \tau}{m}. \quad (1.16)$$

Ce résultat se généralise au cas d'un champ variable de pulsation donnée, ω ; en posant :

$$\vec{\mathcal{E}} = \vec{\mathcal{E}}_0 e^{i\omega t}, \quad (1.17)$$

et en reportant dans (1.13), on trouve⁹, pour la composante de Fourier de la vitesse, $V(\omega)$:

$$im\omega \vec{V}(\omega) + \frac{m}{\tau} \vec{V}(\omega) = e\vec{\mathcal{E}}_0 \iff \vec{V}(\omega) = \frac{\tau}{m(1+i\omega\tau)} e\vec{\mathcal{E}}_0. \quad (1.18)$$

La transformée de Fourier du courant, $\vec{J}(\omega) = n_e e \vec{V}(\omega)$ fait apparaître la conductivité en fréquence $\sigma(\omega)$:

$$\sigma(\omega) = \frac{n_e e^2 \tau}{m(1+i\omega\tau)} \equiv \frac{\sigma_0}{1+i\omega\tau}. \quad (1.19)$$

La conductivité tend donc vers σ_0 à fréquence nulle et vers zéro aux très hautes fréquences. Ce dernier résultat se comprend bien physiquement : si la force électrique change de sens trop rapidement, l'électron n'a pas le temps de suivre le mouvement et le courant devient nul.

Le modèle de Drude ne vaut que pour un métal : son ingrédient essentiel est l'existence d'une densité n_e de porteurs libres (ou quasi-libres). Ceci n'est réalisé ni dans un semi-conducteur, ni dans un isolant. Par ailleurs, la relation (1.16) peut être inversée :

$$\tau = \frac{m}{n_e e^2} \sigma_0 \equiv \frac{m}{n_e e^2} \frac{1}{\rho}, \quad (1.20)$$

où $\rho = 1/\sigma_0$ est la résistivité usuelle, afin d'exprimer le temps de relaxation en termes de quantités facilement mesurables. On peut ainsi trouver que τ est, à la température ambiante, de l'ordre de 10^{-14} à 10^{-15} s. Ce résultat permet de vérifier *a posteriori* la cohérence du modèle de Drude – et aussi de rencontrer ses limites quantitatives. En effet, on peut se convaincre que τ est du même ordre de grandeur que l'intervalle de temps entre deux collisions puisqu'il est certain, physiquement, que quelques collisions suffisent pour effacer la mémoire de toute condition initiale, les diffuseurs étant indépendants les uns des autres. Ceci étant, avec τ et une vitesse on peut maintenant former une longueur – qui ne sera rien d'autre que le libre parcours moyen :

$$l = \text{vitesse} \times \tau. \quad (1.21)$$

Comme vitesse, en restant à un niveau rustique puisque le modèle lui-même l'est, on n'a guère le choix et celui de la vitesse thermique s'impose. Sans aller chercher de complication, on prend la répartition de Maxwell, qui conduit à $m\vec{v}_{\text{th}}^2/2 = (3/2)k_B T$; à l'ambiante et avec $m = 9 \times 10^{-31}$ kg, on trouve $v_{\text{th}} \sim 10^5$ m/s. Maintenant (1.21) donne :

$$l = v_{\text{th}} \times \tau \sim 1 \text{ à } 10 \text{ \AA}. \quad (1.22)$$

Cette distance est comparable à la distance entre ions du métal et, interprétée comme un libre parcours moyen, justifie grossièrement l'image de la dissipation d'énergie par collisions successives des porteurs de charge sur les

⁷Cette amnésie est typique des systèmes ordinaires (en particulier non-chaotiques), dont l'état d'équilibre (ou stationnaire), rendu accessible grâce à la relaxation (dissipation), est défini en soi par la seule donnée de conditions externes excluant toute information détaillée de type "métaphysique". Par ailleurs, l'expression au second membre de (1.14) est très exactement la vitesse moyenne (au sens statistique) dans l'hypothèse où la vitesse initiale \vec{v}_0 est distribuée uniformément au hasard.

⁸aussi appelée conductivité statique.

⁹On note que l'usage élémentaire de la transformation de Fourier fait disparaître la nécessité de spécifier une condition initiale.

ions (il s'agit plus précisément d'une interaction entre électrons et phonons¹⁰), les ions "reculant" sous l'effet des collisions et se mettant à vibrer autour de leur position d'équilibre.

L'usage de v_{th} est un peu désinvolte : en réalité, la vitesse typique des électrons, à la température ambiante, est nettement plus élevée. Ceci tient essentiellement au fait que, en vertu du Principe d'exclusion de Pauli, pour un nombre donné d'électrons on est dans l'obligation de peupler des états de forte énergie. Par ailleurs, le temps de relaxation est en fait lui aussi beaucoup plus élevé. Il en résulte que le libre parcours moyen l est plutôt de l'ordre de 10^3 Å, ce qui invalide l'image naïve d'électrons se bornant à rebondir inélastiquement sur le réseau d'ions. En tout état de cause, il est clair que le modèle de Drude, en raison de sa simplicité, ne doit pas être trop pris au sérieux ; visiblement, l'auto-justification ci-dessus est un peu illusoire.

Une fois introduite la classification conducteurs, semi-conducteurs et isolants, il s'avère utile de définir une sous-classification de ces derniers. On distingue ainsi, parmi les isolants, les solides covalents, les solides moléculaires et les solides ioniques – qui seront discutés plus loin.

1.2.4 Propriétés thermiques

Elles se caractérisent principalement par la chaleur spécifique $C_V = (\partial E / \partial T)_V$ et par la conductivité thermique, κ . κ est défini par :

$$\vec{j}_Q = -\kappa \vec{\nabla} T , \quad (1.23)$$

où \vec{j}_Q est le courant de chaleur et T la température. Cette relation, presque naturelle et en tout cas d'apparence anodine, repose sur l'hypothèse forte de l'existence d'un équilibre local : il faut que l'on puisse définir des domaines de taille mésoscopique, contenant un assez grand nombre d'éléments microscopiques pour que chaque domaine ait un nombre énorme de degrés de liberté autorisant une description statistique, que ce soit avec la Mécanique Statistique ou avec sa variante "phénoménologique", la Thermodynamique. C'est seulement à ce prix que l'on peut définir une température *locale* – et son éventuel gradient. Clairement, la relation de simple proportionnalité (1.23) ne vaut que pour les petits gradients de température, et traduit une hypothèse de réponse linéaire – tout comme la loi de Fick conduisant à l'équation de diffusion classique. En définitive, (1.23) repose sur deux hypothèses fortes : l'existence d'un équilibre local et d'une réponse linéaire.

D'une façon générale, la classification des solides sur la base des propriétés thermiques redonne essentiellement la classification entre conducteurs et isolants, fondée, elle, sur un critère de conductivité électrique. Ainsi par exemple, la loi de Wiedemann - Franz pour les métaux affirme la simple proportionnalité entre σ et κ , aux températures pas trop basses :

$$\frac{\kappa}{\sigma} = L T , \quad (1.24)$$

où L est parfois appelé nombre de Lorenz. C'est pourquoi il est à la fois plus aisé et plus utile de discuter les propriétés thermiques séparément pour chaque type de solide.

La loi exprimée par la relation (1.24) peut se comprendre qualitativement comme suit. Le transport de chaleur ou d'électricité dans les métaux est assuré par les électrons "libres". Le courant de chaleur (d'énergie) contient le carré de la vitesse ; au contraire, le courant électrique a en facteur la charge électrique, qui est une constante. Quand on fait le rapport κ/σ , on obtient donc essentiellement le carré de la vitesse, qui est (classiquement) proportionnel à T .

En ce qui concerne la capacité calorifique, on verra que, pour un isolant, et à basse température, la chaleur spécifique varie comme T^3 , c'est la loi dite de Debye :

$$C_V \propto T^3 \quad (\text{isolant}) , \quad (1.25)$$

alors que pour un métal, on a un comportement du genre :

$$C_V = \gamma T + A T^3 \quad (\text{métal}) . \quad (1.26)$$

où γ est une constante, parfois appelée constante de Sommerfeld. Ce terme linéaire en température, absent pour un isolant, est d'origine électronique : il représente la contribution du gaz d'électrons quasi-libres (les électrons

¹⁰donc présente même dans un cristal parfait ne présentant aucun défaut, ni structurel, ni chimique.

qui conduisent le courant), précisément absent dans un isolant. Le terme cubique vient des vibrations du réseau (phonons) ; l'expression (1.26) est assez remarquable en ce sens qu'elle rassemble les contributions des deux types de particules apparaissant en statistique quantique : les fermions (électrons) et les bosons (phonons)¹¹.

1.2.5 Propriétés magnétiques

Les propriétés magnétiques de la matière, condensée ou non, sont extrêmement variées, allant du diamagnétisme universel au ferromagnétisme spécifique (et spectaculaire) de certains matériaux, en passant par le paramagnétisme.

- Le diamagnétisme est universel au sens où il existe dès que l'on applique un champ magnétique à des charges mobiles liées – ici ce sont les électrons. Schématiquement, le champ produit une force radiale supplémentaire au mouvement de rotation représentant le mouvement lié en l'absence de champ appliqué. En considérant le cas où l'orbite est orientée de telle sorte que l'électron tourne dans le sens trigonométrique positif autour du champ, la force s'ajoute à la force responsable du mouvement lié, l'électron doit tourner plus vite pour rester en équilibre, ce qui correspond à une augmentation du moment cinétique, donc à une variation du magnétique associé à la boucle de courant ; mais comme le facteur gyromagnétique est *négligé* – il contient la charge de l'électron –, le moment induit créé par cette rotation plus rapide est en sens contraire du champ – c'est l'image microscopique de la loi de Lenz. Le diamagnétisme est l'apparition d'un moment induit par le champ appliqué, qui se couple précisément au champ qui lui a donné naissance. C'est donc un effet du second ordre et c'est pourquoi, dans la théorie élémentaire d'un atome couplé à un champ \vec{B} , le terme en B^2 est qualifié de diamagnétique. On sait qu'il est très petit par rapport au terme linéaire produisant l'effet Zeeman. Ceci est général : le diamagnétisme est un phénomène universel, parce que, au sein de la matière, il y a des électrons qui tournent autour des noyaux. Toutefois, les moments induits sont très petits ; il en résulte que le diamagnétisme est complètement occulté par le paramagnétisme quand ce dernier existe.
- Le paramagnétisme se produit lorsque les entités constitutives du solide (atomes, ions, molécules) possèdent chacune un moment magnétique non-nul en moyenne, qu'il soit d'origine orbitale, de spin ou des deux – ce qui n'est pas possible si toutes les couches atomiques sont complètes. L'application du champ oriente préférentiellement ces moments et il apparaît une magnétisation moyenne non-nulle (paramagnétisme de Langevin¹²), qui disparaît évidemment en champ nul, et permet de ce fait de définir une susceptibilité magnétique, χ , toujours finie à température finie (loi de Curie, $\chi \propto T^{-1}$). Bien évidemment, l'aimantation n'est pas totale : la tendance à l'alignement sur le champ est en compétition avec l'agitation thermique, se traduisant par des poids de Boltzmann (probabilités).
- Enfin, le ferromagnétisme désigne l'aptitude d'un corps (Fe, Ni, Co) à conserver une aimantation macroscopique (aimantation dite spontanée) en champ (magnétique) nul, à condition d'être à une température inférieure à une température remarquable (critique), appelée température de Curie (environ 1000° K pour le fer). Cette propriété implique l'existence de moments microscopiques qui coopèrent et qui sont tous alignés dans le même sens sur une distance très grande à l'échelle atomique (domaines de Weiss).

Ces deux derniers types d'états (para- et ferromagnétisme) peuvent être rapprochés en introduisant la notion d'ordre à distance. Un ferromagnétique possède un ordre à longue distance, et une aimantation spontanée en champ nul. Au contraire, dans une phase paramagnétique c'est le champ appliqué de l'extérieur qui crée une polarisation (aimantation par unité de volume), et les fonctions de corrélation décroissent exponentiellement (ordre à courte portée), résultat d'une compétition entre la tendance à l'alignement sur le champ appliqué et la tendance vers le désordre provoquée par les fluctuations thermiques ; la polarisation disparaît en champ nul et, à l'instar d'un gaz, il n'y a pas d'ordre à longue distance. Ces arguments, fondés sur la notion d'ordre,

¹¹Les comportements de C_V en température donnés ci-dessus excluent tout effet de type magnétique. Pour un solide magnétique, les *magnons* contribuent à C_V au même titre que les autres degrés de liberté.

¹²ou sa version quantique, paramagnétisme de Brillouin.

permettent de soupçonner qu'il existe, *mutatis mutandis*, des analogies¹³ entre la transition liquide - gaz et la transition ferromagnétique - paramagnétique.

La question de savoir pourquoi on passe – pour certains matériaux – d'un état désordonné (paramagnétique) à un état ordonné (ferromagnétique) quand on abaisse suffisamment la température est l'enjeu de la théorie des transitions de phase. Dans ce contexte, la notion de phase apparaît comme le résultat de compétitions électroniques très subtiles, ne se traduisant pas par une modification structurale au sens premier du terme. Notons enfin que l'état ferromagnétique n'est pas propre à la matière solide au sens usuel : il existe des ferrofluides.

Bien d'autres structures magnétiques sont concevables, et ont été observées (par exemple : antiferromagnétisme – qui peut donner lieu à des phénomènes remarquables¹⁴ de *frustration* –, héliamagnétisme, *etc.*). Enfin, il est utile de rappeler que le magnétisme – quel qu'il soit – n'est explicable que dans un cadre quantique. Le théorème de Miss van Leeuwen stipule que les briques élémentaires du magnétisme – nécessaires pour expliquer ne serait-ce que le paramagnétisme – n'existent pas dans un cadre classique (plus précisément, les moyennes statistiques de ces “ampériens” sont toujours nulles).

1.3 Classification des solides ordonnés

Il s'agit ici par définition des solides possédant un ordre à longue portée, en principe infinie. Un tel solide résulte du rapprochement d'entités bien définies avant la formation du cristal : atomes, molécules, ions. Bien sûr, la constitution du solide en tant qu'édifice stable s'accompagne d'altérations plus ou moins grandes de ces briques élémentaires constitutives ; on admettra “qu'il en reste toujours quelque chose”, ce qui est presque une évidence sur le plan physique. Il y a donc dans le solide la trace des constituants à l'état libre : on parle alors de “cœur” pour désigner ce fantôme du constituant isolé. Les électrons complémentaires sont traditionnellement appelés électrons de valence en ce qui concerne ceux qui sont principalement responsables de la cohésion du solide, ou, s'il en reste, électrons de conduction puisque, presque libres, ils sont aptes à conduire le courant électrique.

Cela étant, comme on l'a vu, une première grande classification se crée spontanément sur la considération des propriétés électriques permettant de séparer nettement conducteurs et isolants, sur la base de leur résistivité électrique. La véritable explication de cette distinction ressort de la Mécanique Quantique et s'exprime par des considérations dans l'espace des impulsions. A défaut d'un critère précis défini sur la considération de grandeurs dans l'espace physique, il est cependant possible d'effectuer une distinction qualitative dans l'espace réel : en général, dans un métal, la distribution électronique est nettement plus diffuse (diluée, délocalisée) que dans les isolants, où les électrons de valence restent davantage situés près des cœurs, ou entre deux cœurs, pour fabriquer la “colle” assurant la stabilité de la liaison chimique.

Il s'avère très utile d'établir une sous-classification pour les isolants, fondée, elle, précisément sur la densité électronique. On distingue ainsi principalement :

- Les solides covalents, où l'on retrouve les caractéristiques *directionnelles* des liaisons chimiques. Le prototype de cette espèce de solide est le diamant, constitué d'atomes de carbone répartis dans l'espace suivant deux réseaux cubiques à faces centrées entremêlés, résultat de la valence 4 du carbone (structure tétraédrique¹⁵). D'autres exemples : le silicium et le germanium, ou même le carbone sous sa forme graphite¹⁶ – qui ont pourtant des propriétés électriques très différentes du diamant ! Alors que le diamant est un isolant remarquable, on sait bien que Ge et Si – même purs – sont des semi-conducteurs.

¹³Ces analogies reposent plus précisément sur la notion de *paramètre d'ordre*, qui n'est pas toujours associé à un ordre spatial (dans le cas liquide - gaz, ce paramètre est la différence des densités entre le solide et le gaz). De façon plus précise, les analogies possibles résultent du fait que, dans un cas comme dans l'autre, le diagramme de phase pertinent contient des lignes de transition séparant les différentes phases, qui sont des lignes de discontinuité pour certaines grandeurs physiques (la densité, ou l'aimantation spontanée en champ nul). Cette *universalité* permet de construire des théories (par exemple celle de Landau) susceptibles de décrire des systèmes physiques à première vue très différents.

¹⁴Par exemple, l'état fondamental d'un système frustré peut être infiniment dégénéré ; il existe alors une entropie résiduelle non nulle à température nulle.

¹⁵C'est alors l'hybridation sp^3 qui est pertinente.

¹⁶expliquée par l'hybridation sp^2 produisant les angles de 120° .

La différence tient bien sûr à la structure de bande et aux valeurs *numériques* des gaps d'énergie (à l'ambiante¹⁷ : 0.67 eV et 1.12 eV pour Ge et Si respectivement, 5.5 eV pour le diamant¹⁸).

- Les solides moléculaires, qui sont fondamentalement une juxtaposition d'atomes ou molécules très peu perturbés par rapport à leur état libre en phase diluée. Il y a peu de densité électronique entre les atomes. Les prototypes sont les solides de gaz rares (Néon, Argon, Krypton, Xénon¹⁹) et les solides de certaines molécules aromatiques (anthracène, par exemple).
- Les solides ioniques, par exemple le sel ordinaire (NaCl). Ils sont formés à partir de couples d'atomes ; l'un des deux est de type métallique (il perd un ou des électrons lors de la formation du solide, c'est ce que fait Na), alors que l'autre est très avide d'électrons (très électronégatif, "antimétallique", c'est le cas de Cl). Les électrons échangés restent très bien localisés près des centres, comme dans un cristal moléculaire, ce qui produit une structure formées de paires d'ions ; mais la réorganisation est importante par rapport à la structure des deux partenaires non-appariés, puisque l'un a perdu des électrons alors que l'autre en a gagné. Au total, les électrons restent bien localisés, l'interaction entre ions n'est pas écrantée et c'est donc l'interaction de Coulomb nue qui est à l'œuvre – bien évidemment complétée par une répulsion de cœur dur.

En tout état de cause, ce sont bien sûr les électrons plus ou moins partagés entre les "briques" du solide qui en constituent la colle, le ciment. Selon le cas, on parlera de liaison métallique, liaison covalente ou liaison ionique. En ce qui concerne les solides moléculaires, il s'agit plutôt d'une pseudo-liaison chimique (fragile) résultant d'une interaction du type van der Waals, interprétable en termes d'échanges de photons (virtuels) entre deux objets (atomes, molécules) globalement neutres. De façon plus élémentaire, il s'agit d'une interaction dipôle induit - dipôle induit : la présence du premier partenaire (et des charges électriques qu'il renferme) induit un moment dipolaire électrique d_2 au sein du deuxième partenaire, via la polarisabilité α :

$$d_2 \sim \alpha \mathcal{E} = \alpha \frac{d_1}{R^3} . \quad (1.27)$$

Comme l'interaction entre deux dipôles est du genre $d_1 d_2 / R^3$, on en déduit une énergie d'interaction attractive variant comme $[d_1(\alpha d_1 / R^3)] / R^3 \propto R^{-6}$. Par $\vec{F} = -\vec{\nabla} V$, il en résulte que les forces de van der Waals décroissent comme $1/R^7$ – aux distances grandes par rapport à l'échelle atomique, mais pas trop grandes : quand les photons (virtuels) – qui sont les vecteurs de l'interaction électromagnétique – mettent trop longtemps pour aller d'un atome à l'autre, il apparaît des corrections donnant finalement une décroissance plus rapide avec la distance (par exemple : $V \propto R^{-7}$).

¹⁷Au voisinage de l'ambiante, le gap varie linéairement en température. Ceci provient de la dilatation du réseau, ce qui induit une expansion du potentiel périodique.

¹⁸Ces valeurs ne sont pas d'un ordre de grandeur différent, mais rentrent dans des exponentielles, d'où les ordres de grandeurs différents pour la conductivité – sans compter avec les valeurs prises par la densité électronique.

¹⁹L'hélium fait bande à part : en vertu de sa petite masse, les effets quantiques ne peuvent jamais être ignorés et toute tentative de description classique est vouée à l'échec.

Chapitre 2

Énergie de cohésion des solides ordonnés

2.1 Généralités

L'énergie de cohésion d'un solide est l'énergie qu'il faut fournir pour le séparer en ses constituants à l'état libre. Une telle définition est clairement ambiguë, car tout dépend des fragments que l'on définit. D'une façon générale, on choisit les éléments libres au vu de considérations physiques mais tout choix reste forcément un peu conventionnel ; les distinctions qui en résultent sont de peu d'intérêt. Par exemple, s'agissant de l'azote solide, on pourrait décider de décomposer le solide en molécules N_2 plutôt qu'en atomes N , mais peu importe au fond, puisque l'on connaît l'énergie de dissociation de la molécule N_2 .

D'une façon générale, un solide se forme par *condensation* d'objets élémentaires (typiquement des atomes ou des petites molécules) initialement en phase vapeur. Lors de la formation du solide, ces entités peuvent perdre une partie de leur intégrité, ce qui donne lieu précisément aux différents types de solides (moléculaires, ioniques, métalliques, *etc.*), comme brièvement discuté au chapitre précédent. On continuera ici à désigner par *cœurs* ce qui reste des objets libres après ce réarrangement ; en général, ceux-ci se définissent d'eux-mêmes sans aucune ambiguïté.

Afin de rester au niveau le plus simple, la discussion ci-dessous de l'énergie de cohésion est faite dans l'hypothèse où les cœurs sont fixes, et sont situés aux nœuds d'un réseau géométrique : on n'envisagera donc que des solides ordonnés. Il s'agit bien sûr d'une approximation, à la fois qualitative et quantitative : d'après la Mécanique Quantique, une particule ne peut être localisée strictement et est soumise aux inégalités de Heisenberg du genre $\Delta x \Delta p_x \sim \hbar$. Il en résulte que pour une particule de masse M , à une incertitude en position Δx correspond une énergie $\Delta E \sim \Delta p^2/2M \sim \hbar^2/(M\Delta x^2)$ – appelée énergie de point zéro quand il s'agit de vibrations – et qui, en toute rigueur, devrait être prise en compte dans le calcul de l'énergie de cohésion. Pour fixer les idées, avec $\Delta x \sim 0.5 \text{ \AA}$, et pour un atome de carbone, on trouve $\Delta E \sim 0.05 \text{ eV}$, ce qui est loin d'être négligeable devant l'énergie par atome que l'on trouvera dans l'hypothèse où le réseau est rigide. Clairement, traiter les cœurs comme des objets ponctuels classiques sera, de ce point de vue, d'autant meilleure que leur masse est élevée.

Il est possible de construire une théorie élémentaire de la cohésion pour les solides moléculaires et ioniques, n'utilisant que les idées classiques. En revanche, rien de tel n'existe pour les solides covalents – sous-jacente, il s'agit en fait de la théorie de la liaison chimique. Quant aux métaux, la situation est difficile à décrire convenablement : on ne peut pas ignorer les électrons presque libres – qu'il faut traiter quantiquement – et leurs interactions, même si elles sont en général assez fortement écrantées.

2.2 Cohésion des solides moléculaires

Le prototype du solide moléculaire est le cristal de gaz rare. Excluant l'hélium – si léger que les effets quantiques de point zéro ne peuvent être ignorés – on se place dans l'approximation où les atomes constitutifs sont traités comme des objets ponctuels classiques. Par définition d'un cristal moléculaire, ses atomes sont très peu modifiés par rapport à leur état libre lors de la formation du cristal ; c'est le cas quand il s'agit d'atomes très robustes comme les gaz rares avec leurs couches électroniques complètes.

L'interaction entre deux objets neutres, robustes et bien localisés se compose de deux contributions dont l'origine physique est évidente :

1. une répulsion forte à courte distance, chaque atome ayant par nature un domaine propre réservé résistant à toute invasion extérieure, en particulier à l'investissement d'un autre atome.
2. une interaction attractive à longue portée, connue sous le nom d'interaction de van der Waals, donnant une énergie potentielle décroissant comme R^{-6} :

$$V(R) \simeq -\frac{C^{\text{ste}}}{R^6} \quad (R \gg a_0) . \quad (2.1)$$

Au total, l'interaction entre deux atomes neutres a l'allure représentée sur la figure (2.1).

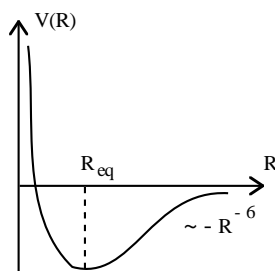


Figure 2.1: Allure de l'énergie potentielle d'interaction entre deux objets neutres.

Autant la forme à longue distance est parfaitement déterminée, et en un sens universelle, autant la répulsion de cœur varie d'un élément à l'autre. La modélisation précise de cette dernière n'a pas beaucoup d'importance (ce qui importe, c'est d'inclure de fait une répulsion), puisqu'elle intervient sur une petite échelle de distance, très peu "explorée" pour le problème ici analysé¹. Le choix d'une dépendance en R ou d'une autre est largement dicté par le souci de la simplicité analytique. Un choix très courant consiste à poser :

$$V(R) \simeq +\frac{A}{R^{12}} \quad (R \lesssim a_0) , \quad (2.2)$$

et on obtient alors le "potentiel 6-12" de Lennard - Jones :

$$V(R) = 4v_0 \left[\left(\frac{d}{R}\right)^{12} - \left(\frac{d}{R}\right)^6 \right] \quad (v_0 > 0) . \quad (2.3)$$

Le 12 n'est finalement choisi que par commodité ($2 \times 6 = 12$!), aucune raison théorique ne le justifie, la seule contrainte est qu'il soit fortement répulsif² – et il l'est. La distance d'équilibre pour une paire unique est $R_{\text{eq}} = 2^{1/6} d \simeq 1.12d$, ce qui donne le sens physique du paramètre d dans (2.3) ; quant à l'énergie de dissociation d'une paire, elle vaut $V(R_{\text{eq}})$; avec le calibrage (2.3), on a tout simplement $V(R_{\text{eq}}) = -v_0$.

¹Tôt ou tard, des exponentielles de Boltzmann du genre $e^{-\beta V(R)}$ interviennent dans des intégrales ; à très courte distance et aux températures usuelles, ces exponentielles tendent vers zéro à toute vitesse et viennent donc "tuer" toute contribution à l'intégrale ; de ce fait, la forme précise de $V(R)$ est en pratique sans grande importance.

²En tout cas, il doit diverger plus vite à l'origine que R^{-6} . D'une façon générale, les calculs quantiques avec les fonctions d'onde donnent plutôt une "divergence" exponentielle du genre $e^{2\alpha(R_{\text{eq}}-R)}$ ($R < R_{\text{eq}}$), que l'on trouve dans le potentiel de Morse $V(x) = -V_0[e^{-2\alpha(R-R_{\text{eq}})} - 2e^{-\alpha(R-R_{\text{eq}})}]$.

L'existence d'une interaction entre les atomes entraîne que le gaz qu'ils constituent ne saurait être parfait. On peut donc espérer trouver les deux paramètres phénoménologiques³ d et v_0 en mesurant l'écart à la loi $PV = Nk_B T$. De fait, loin du point critique de la transition liquide - gaz, il existe un développement de l'équation d'état des fluides en puissances de la densité ρ de particules ($\rho = N/V$), appelé développement du Viriel. Celui-ci permet d'exprimer toutes les grandeurs thermodynamiques sous la forme d'un développement du même type ; en particulier, la pression P apparaît maintenant sous la forme :

$$P = N \frac{k_B T}{V} [1 - B(T)\rho - C(T)\rho^2 - \dots] . \quad (2.4)$$

Dans cette expression, le premier terme redonne la loi des gaz parfaits. Les fonctions $B(T)$, $C(T)$, \dots , sont par définition les coefficients du Viriel⁴, et s'expriment à l'aide d'intégrales impliquant justement l'interaction entre deux atomes ; par exemple :

$$B = 2\pi \int_0^{+\infty} r^2 dr [e^{-\beta V(r)} - 1] . \quad (2.5)$$

La détermination de la quantité⁵ B (via la mesure des isothermes par exemple) permet de caractériser $V(R)$ et donne donc accès à v_0 et d – étant à nouveau entendu que des calculs de chimie quantique permettent aussi d'obtenir des valeurs numériques plus ou moins précises. B dépend de la température, mais pas de la densité ρ : ceci permet de comprendre que les valeurs obtenues par l'étude du *gaz* conservent toute leur utilité pour l'étude de la phase condensée. Quelques valeurs numériques :

	Ne	Ar	Kr	Xe	
v_0 (eV)	0.0031	0.0104	0.0140	0.0200	.
R_{eq} (Å)	2.74	3.40	3.65	3.98	

(2.6)

Comme le choix de la partie répulsive n'est pas critique, il arrive aussi souvent que l'on prenne un potentiel de sphère dure pour la représenter. Alors :

$$V(R) = \begin{cases} +\infty & \text{si } R < d \\ -\frac{C^{\text{ste}}}{R^6} & \text{si } R > d \end{cases} , \quad (2.7)$$

où d est une longueur atomique, d'ordre a_0 (voir figure 2.2).

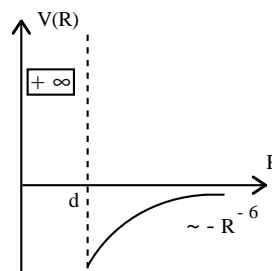


Figure 2.2: Allure de l'énergie potentielle d'interaction entre deux objets neutres avec un cœur dur d'extension d .

Dans tous les cas, les énergies de liaison par paire sont faibles – comme en attestent les valeurs de v_0 , comparées aux valeurs usuelles de la physique atomique ou de la chimie, ce qui n'est pas surprenant compte tenu de la nature physique d'un cristal moléculaire. En tout cas, c'est bien parce que les composants sont très peu altérés lors de la condensation que l'on peut utiliser les paramètres déduits de l'analyse en phase vapeur.

³Rappelons qu'en toute rigueur, seul le coefficient du terme en R^{-12} est d'essence phénoménologique. On peut en principe calculer le terme de van der Waals si l'on connaît les fonctions d'onde des atomes (ou des molécules).

⁴Il est clair qu'un tel développement n'a d'intérêt et d'utilité que pour des densités pas trop élevées. Par ailleurs, son existence est plausible physiquement, à condition d'être loin d'un point critique.

⁵Noter que B est une fonction de la température.

L'interaction d'une paire étant précisée, il convient maintenant de considérer effectivement un *crystal*, construit en disposant aux nœuds d'un réseau géométrique les atomes en interaction par $V(\vec{R})$. Si \vec{R}_i repère un point de ce réseau, l'interaction totale est :

$$E = \frac{1}{2} \sum_{\vec{R}_i, \vec{R}_j, i \neq j} V(\vec{R}_i - \vec{R}_j) , \quad (2.8)$$

où le facteur 1/2 évite de compter deux fois l'interaction au sein d'une même paire. Dans la limite d'un cristal macroscopique contenant N ($\sim 10^{23}$) atomes, il y a une quasi-invariance par translation (discrète), l'énergie E est une grandeur extensive⁶ et il vient simplement, aux termes de surface près :

$$E \rightarrow \frac{N}{2} \sum_{\vec{R} \neq \vec{0}} V(\vec{R}) \quad (N \gg 1) . \quad (2.9)$$

Dans cette limite, on peut donc définir une énergie de cohésion par particule, ε :

$$\varepsilon \equiv \frac{E}{N} = \frac{1}{2} \sum_{\vec{R}} V(\vec{R}) . \quad (2.10)$$

On peut aussi dire que, fixant une particule (à l'origine par exemple), son interaction avec toutes les autres est $\sum_{\vec{R} \neq \vec{0}} V(\vec{R})$, ce qui fournit $N \sum_{\vec{R} \neq \vec{0}} V(\vec{R})$ pour les N particules ; mais alors, on a compté deux fois l'interaction d'une même paire (i, j) , d'où le facteur 1/2.

Si l'on choisit le potentiel (2.3), il faut donc calculer :

$$\varepsilon = 4v_0 \frac{1}{2} \sum_{\vec{R}} \left[\left(\frac{d}{R} \right)^{12} - \left(\frac{d}{R} \right)^6 \right] . \quad (2.11)$$

où la somme court sur tous les nœuds du réseau, chacun d'entre eux étant caractérisé par un vecteur \vec{R} . Par exemple, pour un réseau cubique de côté a , on doit calculer des sommes⁷ du genre :

$$\left(\frac{d}{a} \right)^n \sum_{m_1, m_2, m_3} \frac{1}{\|m_1 \vec{i} + m_2 \vec{j} + m_3 \vec{k}\|^n} \quad (2.12)$$

avec $n = 6, 12$ et où les m_i sont des entiers naturels ; le calcul de telles sommes (triples !) n'est pas si simple et des méthodes ingénieuses ont été mises au point pour accélérer la convergence.

Dans le cas d'un réseau quelconque, il est commode de poser $\vec{R} = l\vec{\rho}(\vec{R})$, où l est la distance entre plus proches voisins (par exemple, pour un cubique faces centrées (CFC) de maille conventionnelle de côté a , $l = a\frac{\sqrt{2}}{2}$, pour un cubique centré (CC), $l = a\frac{\sqrt{3}}{2}$). Avec ces notations, il faut calculer les constantes A_n :

$$\left(\frac{d}{l} \right)^n \sum_{\vec{R} \in \mathcal{B}} \frac{1}{\|\vec{\rho}(\vec{R})\|^n} \equiv \left(\frac{d}{l} \right)^n A_n \quad (2.13)$$

Pour un réseau CFC (le cas des gaz rares), les sommes de réseau donnent :

$$A_6 \simeq 14.45 , \quad A_{12} \simeq 12.13 , \quad (2.14)$$

Il en résulte que l'énergie de cohésion par particule est :

$$\varepsilon = 2v_0 \left[A_{12} \left(\frac{d}{l} \right)^{12} - A_6 \left(\frac{d}{l} \right)^6 \right] . \quad (2.15)$$

⁶L'argument ne vaut que dans la mesure où les forces ne sont pas à portée infinie : de ce fait, les effets de surface/volume s'évanouissent à la limite thermodynamique. En effet, si ξ désigne la portée des interactions, le nombre d'atomes de surface en interaction mutuelle est $\sim (N/L^3) \xi L^2$, pour un échantillon cubique de côté L , soit $N(\xi/L)$. Dans la limite thermodynamique, pour une grandeur extensive comme l'énergie, les corrections de surface s'annulent comme L^{-1} , soit comme $N^{-1/3}$ lorsque la densité est maintenue constante.

⁷dites *sommes de réseau*.

Par minimisation, on trouve la distance d'équilibre l_{eq} , reliée au côté de la maille cubique conventionnelle du cristal CFC constitué ($a_{\text{eq}} = \sqrt{2}l_{\text{eq}}$) :

$$l_{\text{eq}} = \left(\frac{2A_{12}}{A_6} \right)^{1/6} d \simeq 1.09 d , \quad a_{\text{eq}} \simeq 1.54 d . \quad (2.16)$$

La distance l_{eq} inférieure est à la valeur d'équilibre d'une seule paire ($\simeq 1.12 d$), traduisant l'augmentation de densité d'un "dimère" quand on passe en phase solide – indépendamment de la structure cristalline (compacte puisqu'il s'agit d'un CFC) : une grande quantité d'atomes qui tous s'attirent ont évidemment tendance à former un assemblage relativement compact – seule la répulsion à courte distance prévenant le *collapse*. Les valeurs numériques sont en assez bon accord avec l'expérience :

	Ne	Ar	Kr	Xe	
l_{eq} (expérimental, Å)	3.13	3.75	3.99	4.33	. (2.17)
$l_{\text{eq}} = 1.09 d$ (Å)	2.99	3.71	3.98	4.34	

On constate d'ailleurs que l'accord est d'autant meilleur que l'atome est plus lourd : l'oubli des termes de point-zéro est bien de moins en moins préjudiciable quand la masse des atomes augmente. D'une façon générale, les vibrations irréductibles d'après Heisenberg jouent le rôle d'une répulsion effective : quand les atomes vibrent, ils occupent plus de place et leur distance d'équilibre se trouve, de fait, augmentée.

Une fois obtenue la distance d'équilibre, on trouve l'énergie de cohésion par particule en reportant l'expression (2.16) de l_{eq} dans (2.15) :

$$\varepsilon_{\text{eq}} = 2v_0 \left[A_{12} \left(\frac{A_6}{2A_{12}} \right)^2 - A_6 \frac{A_6}{2A_{12}} \right] = -v_0 \frac{A_6^2}{2A_{12}} . \quad (2.18)$$

Pour le CFC, on a ainsi :

$$\varepsilon_{\text{eq}} \simeq -8.60 v_0 . \quad (2.19)$$

L'accord est convenable, comme le montre la table ci-dessous ; à nouveau, il est d'autant meilleur que l'atome est plus massif. Les énergies expérimentales sont plus élevées (en valeurs algébriques, l'énergie de liaison est plus faible, la liaison est donc plus fragile) puisque l'on a négligé une contribution *positive* à l'énergie, à savoir l'énergie de vibration⁸. On note la faiblesse de la liaison comparée à une liaison chimique ordinaire, plutôt de l'ordre de l'eV.

	Ne	Ar	Kr	Xe	
ε (expérimental, eV)	-0.02	-0.08	-0.11	-0.17	. (2.20)
$\varepsilon_{\text{eq}} = -8.60 v_0$ (eV)	-0.027	-0.089	-0.120	-0.172	

On note qu'il y a presque un facteur 10 entre ε_{eq} et v_0 : rapportée à un constituant, l'énergie de cohésion du solide CFC est beaucoup plus grande que l'énergie d'une molécule X_2 .

2.3 Cohésion des réseaux ioniques

On considère à nouveau le cas le plus simple : un ensemble d'ions fixés aux nœuds d'un réseau et en prenant comme fragments les ions isolés. Il n'y a toujours pas d'énergie cinétique (les ions sont fixés) et l'interaction dominante à une distance $\gtrsim a_0$ est maintenant celle de Coulomb, qui décroît très lentement comme R^{-1} . Il existe toujours une répulsion à courte distance, de sorte qu'il faut écrire (dans la limite macroscopique) :

$$E = \frac{N}{2} \sum_{\vec{R}} [V_{\text{cœurs}}(\vec{R}) + V_{\text{Coulomb}}(\vec{R})] \equiv E_{\text{cœurs}} + E_{\text{Coulomb}} . \quad (2.21)$$

En raison de la décroissance lente de l'interaction de Coulomb, le calcul de la somme impliquant V_{Coulomb} n'est pas aussi simple qu'il pourrait paraître à première vue. La difficulté est visible immédiatement pour un réseau à une dimension, où alternent des ions positifs et des ions négatifs, portant chacun une charge élémentaire et

⁸L'énergie de vibration contient deux termes (cinétique et potentiel), tous deux positifs.

situés à la distance a les uns des autres. L'énergie potentielle de l'ion situé à l'origine liée à son interaction avec tous les autres est :

$$2 \frac{e'^2}{a} \left[-1 + \frac{1}{2} - \frac{1}{3} + \dots \right]. \quad (2.22)$$

Le facteur 2 global prend en compte les deux moitiés du réseau, à gauche et à droite de l'origine. La série harmonique alternée est convergente :

$$\sum_{n=1}^{+\infty} \frac{(-1)^n}{n} = -\ln 2, \quad (2.23)$$

mais converge avec une lenteur désespérante. D'ailleurs, on pourrait décider de compter autrement, en prenant d'abord l'interaction entre l'ion à l'origine avec tous ses partenaires de même signe ; ainsi viendrait la première série :

$$2 \frac{e'^2}{a} \sum_{n=1}^{+\infty} \frac{1}{2n}. \quad (2.24)$$

L'interaction entre le même ion central et ceux de charge opposée est :

$$-2 \frac{e'^2}{a} \sum_{n=1}^{+\infty} \frac{1}{2n+1}. \quad (2.25)$$

Les deux séries (2.24) et (2.25) sont ... divergentes !

Le même type de difficulté se produit dans le cas tridimensionnel ; la série est mal conditionnée et il faut en fait construire le cristal de façon astucieuse, afin d'obtenir rapidement la bonne réponse. Par exemple, le cristal de sel, NaCl, est un cubique à faces centrées : deux réseaux CFC de même côté (paramètre) a , mais décalés de $a/2$ le long d'un côté du cube. La distance entre l'ion à l'origine (Na^+ , par exemple) et ses homologues situés en \vec{R} est de la forme $\rho(\vec{R})a$, où ρ est la fonction vectorielle donnant, en unités a , la distance entre deux ions Na^+ . L'énergie électrostatique prend alors la forme :

$$E_{\text{Coulomb}} = \frac{1}{2} \frac{Ne'^2}{a} \left[- \sum_{\forall \vec{R}} \frac{1}{\rho(\vec{R} + \vec{d})} + \sum_{\forall \vec{R} \neq \vec{0}} \frac{1}{\rho(\vec{R})} \right], \quad (2.26)$$

où \vec{d} est la translation d'un demi-côté du cube. L'énergie par paire d'ions⁹ est donc :

$$\varepsilon_{\text{Coulomb}} = \frac{e'^2}{a} \left[- \frac{1}{\rho(\vec{d})} - \sum_{\forall \vec{R} \neq \vec{0}} \left(\frac{1}{\rho(\vec{R} + \vec{d})} - \frac{1}{\rho(\vec{R})} \right) \right]. \quad (2.27)$$

Compte tenu de la décroissance lente de l'interaction de Coulomb, la série apparaissant dans (2.27) est *conditionnellement* convergente : suivant les regroupements de termes effectués, on peut lui donner à peu près n'importe quelle valeur !

La difficulté peut aussi se sentir qualitativement en invoquant le théorème de Gauss, à propos d'une distribution de charge à symétrie sphérique, de rayon R . On sait que le champ ressenti par une charge-test située à la distance $r > R$ est le même que si toute la charge était concentrée au centre. Autrement dit, quand on ajoute des charges en surface, l'effet global est le même, indépendant de la distance R de ces charges aux charges situées vers le centre de la distribution. Ainsi, une distribution de charge, aussi étendue qu'elle soit, est sensible à ce qui se passe en surface.

Pour contourner la difficulté, on peut construire mentalement le cristal en le complétant à chaque fois par des fragments de charge totale nulle. Clairement, l'adjonction d'un tel fragment donnera des contributions positives et négatives à l'énergie d'interaction et, la distance aidant, les termes supplémentaires seront d'autant plus petits que la taille du système est grande. Dans ces conditions, la série devient (plus ou moins) rapidement convergente.

Dans le cas d'un cristal à symétrie cubique, il est possible de fournir un argument plus quantitatif permettant de comprendre ce qui se passe, et par là de construire des méthodes introduisant des séries rapidement

⁹S'agissant des cristaux ioniques, l'usage est de compter l'énergie par paire d'ions, plutôt que par ion.

convergentes. Le développement multipolaire s'obtient en faisant un développement de Taylor, ce qui peut s'écrire en utilisant la relation formelle (employée en Mécanique quantique à plusieurs reprises) :

$$f(\vec{r} + \vec{a}) = e^{\vec{a} \cdot \vec{\nabla}} f(\vec{r}) = [1 + \vec{a} \cdot \vec{\nabla} + \frac{1}{2!}(\vec{a} \cdot \vec{\nabla})^2 + \dots] f(\vec{r}) ; \quad (2.28)$$

ainsi, en développant autour du point \vec{r}' ($\vec{a} = -\vec{r}'$), il vient :

$$\frac{1}{\|\vec{r} - \vec{r}'\|} = \frac{1}{r} - (\vec{r}' \cdot \vec{\nabla}) \frac{1}{r} + \frac{1}{2} (\vec{r}' \cdot \vec{\nabla})^2 \frac{1}{r} + \dots = \frac{1}{r} + \frac{\vec{r} \cdot \vec{r}'}{r^3} + \frac{3(\vec{r} \cdot \vec{r}')^2 - \vec{r}^2 \vec{r}'^2}{r^5} + \frac{1}{r} \mathcal{O}\left(\frac{r'}{r}\right)^3 . \quad (2.29)$$

Avec :

$$\vec{r}' \cdot \vec{\nabla} \frac{1}{r} = \vec{r}' \cdot \frac{\vec{r}}{r^3} , \quad (\vec{r}' \cdot \vec{\nabla})^2 = \frac{3(\vec{r} \cdot \vec{r}')^2 - \vec{r}^2 \vec{r}'^2}{r^5} , \quad (2.30)$$

l'interaction entre le cristal déjà constitué de densité $\rho(\vec{r})$, et celle du fragment ajouté de densité $\rho_{\text{fr}}(\vec{r}')$ est :

$$\begin{aligned} \frac{1}{2} \int d^3 r \int_{\mathbb{R}^3} d^3 r' \frac{\rho(\vec{r}) \rho_{\text{fr}}(\vec{r}')}{4\pi\epsilon_0 \|\vec{r} - \vec{r}'\|} d^3 r' = \\ \frac{1}{2} \int d^3 r \frac{\rho(\vec{r})}{4\pi\epsilon_0} \left[\int d^3 r' \frac{1}{r} \rho_{\text{fr}}(\vec{r}') + \int d^3 r' \frac{\vec{r} \cdot \vec{r}'}{r^3} \rho_{\text{fr}}(\vec{r}') + \int d^3 r' \frac{3(\vec{r} \cdot \vec{r}')^2 - \vec{r}^2 \vec{r}'^2}{r^5} \rho_{\text{fr}}(\vec{r}') + \dots \right] . \end{aligned} \quad (2.31)$$

Quand le fragment ajouté est de charge totale nulle, le premier terme du crochet est nul. Le second l'est tout autant si $\rho_{\text{fr}}(\vec{r}')$ est pair dans l'inversion d'espace, c'est-à-dire si la densité du morceau ajouté est invariante dans le changement $\vec{r}' \rightarrow -\vec{r}'$. Enfin, si le fragment supplémentaire a la symétrie cubique du cristal, le terme en r^{-5} est aussi nul (rapporté au vecteur unitaire \vec{r}/r , il est en r^{-3}), puisque la symétrie cubique assure que $3\cos^2\theta - 1$ est nul. Le terme suivant (non-écrit) – en r^{-4} suivant la même convention – est encore nul par parité. Finalement le premier terme non nul est en r^{-5} : clairement, la série devient rapidement convergente¹⁰.

D'autres méthodes de sommation existent et donnent d'une façon ou d'une autre l'énergie de Coulomb ; pour un réseau cubique de côté a , elle est nécessairement de la forme :

$$E_{\text{Coulomb}} \propto -N \frac{e'^2}{a} . \quad (2.32)$$

En comptant par paire d'ions (l'habitude pour les cristaux ioniques), on définit précisément une constante A , dite *constante de Madelung*, par la relation :

$$E_{\text{Coulomb}} = -A \frac{N}{2} \frac{e'^2}{l} ; \quad (2.33)$$

A est un simple facteur numérique, l étant toujours la distance entre plus proches voisins. En définitive, l'enjeu des différentes méthodes de sommation se réduit au calcul de cette constante. Par paire d'ions, (2.33) donne :

$$\epsilon_{\text{Coulomb}} = -A \frac{e'^2}{l} . \quad (2.34)$$

La constante de Madelung ne dépend que de la structure cristalline et peut être calculée numériquement¹¹ ; par exemple :

Structure cristalline	Constante de Madelung, A	
CsCl	1.7627	(2.35)
NaCl	1.7476	
Pechblende	1.6381	

¹⁰À la réflexion, l'argumentation peut paraître un peu spéieuse, car la Nature se moque de savoir comment le Physicien s'y prend pour faire ses calculs. Et, en définitive, comment peut-on être sûr que le calcul, fait d'une façon ou d'une autre, donne le bon résultat ? En réalité, il faut bien comprendre ceci : quoiqu'il en soit, les cristaux sont toujours de taille *finie*, contenant N ions. Alors, quel que soit N , aussi gigantesque que l'on veut, le résultat est unique, indépendant de la procédure de sommation utilisée. Finalement, s'agissant de calculer la somme d'un grand nombre de termes, c'est le souci d'une bonne efficacité numérique qui dicte le choix d'une méthode ou d'une autre. Physiquement, la réponse est satisfaisante si, une fois la somme calculée par la méthode la plus rapide, l'énergie devient – à l'approximation souhaitée – une grandeur extensive.

¹¹On la connaît avec un nombre de décimales très surdimensionné pour les applications usuelles (voir André Hautot, "New applications of Poisson's summation formula", J. of Phys, A **8**, 853(1975)).

Ceci est une borne supérieure de l'énergie de liaison, puisque la répulsion de cœur (déstabilisante) n'est pas prise en compte. En adoptant une loi-puissance pour cette dernière, l'énergie de cohésion dans le modèle à ions fixes est :

$$E = \frac{N}{2} \varepsilon \quad \varepsilon = -A \frac{e'^2}{l} + \frac{C}{l^m} \quad (N \gg 1) . \quad (2.36)$$

La distance d'équilibre s'obtient par minimisation :

$$A \frac{e'^2}{l^2} - m \frac{C}{l^{m+1}} = 0 \quad \iff \quad l_{\text{eq}} = \left(\frac{mC}{Ae'^2} \right)^{1/(m-1)} . \quad (2.37)$$

On en tire l'énergie par paire d'ions à l'équilibre en utilisant (2.37) au passage :

$$\varepsilon_{\text{eq}} = -A \frac{e'^2}{l_{\text{eq}}} + \frac{C}{l_{\text{eq}}^m} = -A \frac{e'^2}{l_{\text{eq}}} + \frac{A}{m} \frac{e'^2}{l_{\text{eq}}} = -A \frac{m-1}{m} \frac{e'^2}{l_{\text{eq}}} . \quad (2.38)$$

Pour m assez grand, ce dernier résultat est très peu différent de $-A(e'^2/l_{\text{eq}})$; l_{eq} dépend faiblement de m : finalement, pour les cristaux ioniques, le cœur fortement répulsif joue peu de rôle. En raison de son intensité et de sa longue portée, l'interaction de Coulomb est, dans ces systèmes, de loin l'interaction dominante. Noter aussi que ε_{eq} est ici de l'ordre de quelques eV, une valeur qui, sans surprise, est très supérieure à son homologue pour les cristaux moléculaires

2.4 Cohésion des solides covalents

Avec les solides covalents (exemple : la variété diamant du carbone), on rentre dans le domaine des solides dont les constituants, une fois rassemblés, sont grandement perturbés par rapport à leur état libre isolé (il en va de même pour les métaux). La liaison est maintenant du type liaison chimique, ce qui constitue une théorie en soi, débordant largement le cadre de ce cours. Il n'existe donc pas de théorie simple bâtie sur des objets libres ; il reste toutefois possible de donner quelques idées qualitatives.

Il est facile de comprendre pourquoi le carbone, par exemple, ne relève pas du tout du même traitement qu'un gaz rare. Dans un gaz rare, toutes les couches atomiques sont saturées et l'addition d'un électron demanderait une énergie importante destinée à peupler un état initialement vide : le principe de Pauli joue à fond son rôle. En revanche, le carbone, avec ses 6 électrons, a la configuration $(1s)^2 (2s)^2 (2p)^2$: la sous-couche (2p) est incomplète (elle peut recevoir 6 électrons), il y a donc beaucoup de place disponible, la case (2p) est de ce fait très malléable. La liaison chimique s'établit grâce au partage des électrons entre plusieurs atomes de carbone, cette mise en commun étant rendue possible par la place disponible chez l'un pour les électrons de l'autre. Ceci produit une délocalisation des électrons partagés et le calcul montre que cette délocalisation produit un abaissement de l'énergie – très schématiquement, quand Δx augmente, Δp_x diminue – et la formation d'une liaison se trouve donc favorisée¹². Une caractéristique importante de la liaison chimique est son aspect directionnel, qui se traduit par une forte densité électronique le long de l'axe joignant les deux membres de la paire d'atomes engagés dans la liaison. Physiquement, c'est cette directivité qui est en grande partie responsable de la "rigidité" des molécules.

2.5 Cohésion des métaux

En première approximation, un métal peut se représenter comme un réseau d'ions tous positifs noyés dans une mer d'électrons presque libres, mais néanmoins soumis au potentiel périodique du réseau. Ce sont ces électrons, libérés lors de la formation du cristal (dont les fragments sont neutres au départ), qui constituent la colle assurant la cohésion du cristal – autrement, tous les ions positifs se repoussant, l'édifice ne pourrait être stable

¹²C'est ce que l'on appelle en Chimie la stabilisation par *résonance* (exemple : les formules résonnantes du benzène).

et exploserait. Dans la théorie des métaux, il est d'usage d'introduire le rayon de la sphère "occupée" par un électron, noté r_s , défini comme¹³ :

$$V = N \frac{4\pi r_s^3}{3} , \quad (2.40)$$

où N est le nombre total d'électrons et V le volume de l'échantillon ; pour les métaux ordinaires, r_s vaut quelques a_0 .

L'énergie de cohésion se définit par référence à l'énergie des constituants séparés. Dans le cas des métaux, des électrons sont libérés lors de la formation du solide et se retrouvent presque libres (dans le cas d'un cristal ionique, chaque électron libéré par l'atome d'une espèce est piégé par un atome de l'autre). L'énergie du cristal constitué doit donc incorporer l'énergie d'interaction entre les ions (comme avant), mais également toutes les énergies relatives aux électrons (énergie potentielle et énergie cinétique).

L'énergie potentielle est celle de l'interaction de Coulomb ; elle donne lieu à plusieurs sortes de contributions :

- les interactions entre ions, répulsives
- les interactions entre électrons, également répulsives
- les interactions entre les ions et les électrons, attractives.

Dans une approche élémentaire, on oublie les vibrations possibles des ions positifs ; ceux-ci se bornent alors à jouer le rôle de charges ponctuelles fixes immergées dans un fond continu négatif compensateur (les électrons), le système étant globalement neutre. Le calcul peut se faire suivant les mêmes techniques que pour les réseaux ioniques ; pour un alcalin, on trouve, par atome, dans le cas d'un réseau cubique centré :

$$\varepsilon_{\text{Coulomb, class}} = - \frac{24.35}{r_s/a_0} \text{ eV} . \quad (2.41)$$

Le résultat (2.41) est purement classique et incorpore toutes les interactions coulombiennes *classiques* de la distribution de charge (électrons et ions) ainsi définie (réseau d'ions fixes, de structure déterminée et distribution électronique uniforme et continue).

En fait, le calcul du terme de répulsion entre électrons ne se réduit pas au terme classique inclus dans (2.41) qui, par nature, ignore le Principe de Pauli. Dans l'approximation minimale prenant ce dernier en compte pour des fermions (une fonction d'onde en forme de déterminant de Slater), on sait de l'étude de l'atome d'hélium, qu'apparaissent en plus des termes dits d'échange donnant au total une contribution toujours négative pour deux électrons de même spin (trou de Fermi). Ceci réduit fortement la répulsion de Coulomb (qui est déstabilisante) par rapport à la situation "spinless"¹⁴ ; au total, la prise en compte de ces termes abaisse la moyenne de l'énergie potentielle totale de façon importante. Pour la contribution d'échange, on trouve très précisément¹⁵ :

$$\varepsilon_{\text{échange}} = - \frac{3}{2\pi} \left(\frac{9\pi}{4} \right)^{1/3} \frac{e'^2}{2a_0} \simeq - \frac{12.4}{r_s/a_0} \text{ eV} , \quad (2.42)$$

de sorte que la prise en compte de l'indiscernabilité des électrons donne au total l'énergie de Coulomb :

$$\varepsilon_{\text{Coulomb}} \simeq - \frac{36.8}{r_s/a_0} \text{ eV} . \quad (2.43)$$

¹³La définition de r_s fluctue d'un auteur à l'autre ; l'autre définition usuelle donne un r_s adimensionné (on compte le rayon de la sphère en unités a_0) :

$$V = N \frac{4\pi r_s^3}{3} a_0^3 . \quad (2.39)$$

¹⁴En effet, pour la moitié des paires d'électrons, la répulsion est fortement diminuée par le trou de Fermi.

¹⁵Ce calcul n'est pas difficile, mais n'est pas pour autant élémentaire. Il passe par le calcul de l'énergie à deux corps dans une approximation de type déterminant de Slater ou, ce qui ici revient au même comme on peut s'en convaincre, un calcul de perturbation au premier ordre.

Ceci étant, à cause de la faible inertie des électrons et de leur mouvement presque libre au sein du métal, il faut prendre en compte leur énergie cinétique, qui entre bien dans le bilan d'énergie par rapport aux constituants (atomes neutres) séparés. En raison de la rusticité du modèle en cours d'analyse, on va calculer cette énergie cinétique comme si les électrons étaient complètement libres – ce qui suppose que l'énergie potentielle de chaque électron dans le champ des ions positifs est une petite correction¹⁶. L'énergie de l'un d'entre eux est alors :

$$\varepsilon(\vec{k}) = \frac{\hbar^2 k^2}{2m} , \quad (2.44)$$

où \vec{k} est le vecteur d'onde. Dans le cas d'un alcalin par exemple, il y a un électron "libre" par maille, c'est-à-dire que s'il y a N ions dans le cristal il y a aussi N électrons, qu'il faut répartir dans les états à une particule conformément au Principe de Pauli. Les électrons étant considérés comme libres, ils sont représentés par des ondes planes $\propto e^{i\vec{k}\cdot\vec{r}}$; en adoptant des conditions cycliques de Born - von Karman dans une boîte cubique de côté L , les valeurs de k sont quantifiées, chaque composante étant un multiple entier de $\frac{2\pi}{L}$: l'élément de volume dans l'espace réciproque est donc un petit cube "infinitésimal" de côté $\frac{2\pi}{L}$.

Le principe de Pauli exige de répartir les électrons dans ces états mettant au plus 2 électrons par état, compte tenu du spin. En procédant ainsi, on remplit ceux-ci jusqu'à un certain vecteur d'onde k_F , appelé vecteur d'onde de Fermi. Dans la sphère de rayon k_F , il y a :

$$\frac{\frac{4\pi}{3} k_F^3}{\left(\frac{2\pi}{L}\right)^3} \quad (2.45)$$

petits cubes et chacun peut accepter deux électrons. Pour un alcalin, le nombre d'électrons "libres" est égal au nombre d'atomes, soit N ; au total, on a :

$$2 \frac{(4\pi/3) k_F^3}{(2\pi/L)^3} = N \iff k_F = \frac{2\pi}{L} \left(\frac{3N}{8\pi}\right)^{1/3} \propto \left(\frac{N}{V}\right)^{1/3} . \quad (2.46)$$

Maintenant, l'énergie cinétique de ces N électrons est :

$$E_{\text{cin}} = \sum_{\text{spin}} \sum_{k \leq k_F} \frac{\hbar^2 k^2}{2m} \rightarrow 2 \left(\frac{L}{2\pi}\right)^3 \int_0^{k_F} 4\pi k^2 dk \frac{\hbar^2 k^2}{2m} ; \quad (2.47)$$

une intégration élémentaire fournit :

$$E_{\text{cin}} = N \frac{3}{5} \frac{\hbar^2 k_F^2}{2m} \equiv N \frac{3}{5} \varepsilon_F , \quad (2.48)$$

la dernière égalité définissant l'énergie de Fermi ε_F pour le gaz d'électrons libres. En introduisant la longueur r_s , on trouve finalement l'énergie cinétique par électron sous la forme :

$$\varepsilon_{\text{cin}} = \frac{3}{5} \frac{e'^2}{2a_0} \left(\frac{9\pi}{4}\right)^{2/3} \left(\frac{a_0}{r_s}\right)^2 \iff \varepsilon_{\text{cin}} \simeq \frac{30.1}{(r_s/a_0)^2} \text{ eV} . \quad (2.49)$$

Notons que comme k_F augmente avec la densité N/V , il en va de même de E_{cin} .

En additionnant maintenant toutes les contributions à l'énergie (2.43) et (2.42), on obtient :

$$\varepsilon = \frac{30.1}{(r_s/a_0)^2} - \frac{36.8}{r_s/a_0} \text{ eV} . \quad (2.50)$$

La minimisation par rapport à r_s donne :

$$\left(\frac{r_s}{a_0}\right)_{\text{eq}} \simeq 1.6 . \quad (2.51)$$

¹⁶Cette idée est le fondement de l'approximation dite des *électrons presque libres*, qui sera exposée ultérieurement. Notons toutefois que même si l'effet du réseau est quantitativement petit, il provoque néanmoins une modification qualitative spectaculaire de la loi de dispersion de l'électron : la parabole libre devient une succession d'arcs déformés et disjoints, d'où l'apparition de *gaps* d'énergie.

L'énergie de cohésion à l'équilibre s'en déduit¹⁷ :

$$|\varepsilon(r_{s, \text{eq}})| \simeq 11.2 \text{ eV} . \quad (2.52)$$

Ces résultats sont médiocres. Notamment, les valeurs expérimentales de $\frac{r_s}{a_0}$ varient en fait entre 2 et 6 ; par ailleurs, on trouve évidemment la même valeur d'équilibre pour tous les alcalins¹⁸, indépendamment de leur masse atomique, puisque les ions ont été pris ponctuels. Les raisons de ces insuccès sont de plusieurs ordres.

D'abord, comme les ions ne sont pas ponctuels, les électrons de conduction ne peuvent pas en fait s'approcher aussi près des ions qu'on l'a supposé : il existe une sorte de volume exclu. Il en résulte d'une part que l'interaction électrostatique électron - ion n'est pas aussi grande en valeur absolue et que d'autre part, là où elle prend des valeurs notables, la densité électronique est en réalité plus élevée, ce qui produit une énergie cinétique plus grande. En conséquence, le terme $\varepsilon_{\text{Coulomb}}$ est moins négatif et l'énergie ε_{cin} est plus élevée. Ces deux tendances se conjuguent pour donner une valeur d'équilibre plus grande que ne le prévoit la formule simpliste (2.50). En outre, en prenant de fait en compte la variation du rayon ionique avec la masse, un calcul excluant les cœurs peut de fait prédire que les différents alcalins ont bien des distances d'équilibre distinctes.

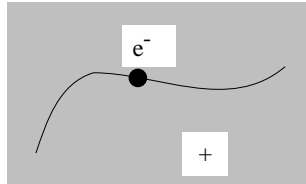


Figure 2.3: Représentation schématique du *jellium*.

Pour les métaux, on dispose aussi d'un autre modèle élémentaire donnant de meilleurs résultats pour r_s – en améliorant le résultat numérique pour les termes de Coulomb – tout en continuant à donner, par construction, une valeur indépendante de l'alcalin considéré. Il s'agit du *jellium*, qui reprend l'essentiel du modèle précédent mais remplace les ions ponctuels par un fond continu positif, d'où son nom. Le calcul de l'énergie cinétique des électrons est évidemment inchangé, et celle-ci est donc toujours donnée par (2.49). En ce qui concerne l'énergie de Coulomb, elle contient toujours des termes "classiques" et des termes d'échange venant du Principe de Pauli. Le calcul donne¹⁹ :

$$\varepsilon_{\text{Coulomb}} = -\frac{3}{2\pi} \left(\frac{9\pi}{4}\right)^{1/3} \frac{e^2}{2a_0} \simeq -\frac{12.4}{r_s/a_0} \text{ eV} . \quad (2.53)$$

Au total, l'énergie par électron est :

$$\varepsilon(r_s) = \frac{30.1}{(r_s/a_0)^2} - \frac{12.4}{r_s/a_0} \text{ eV} . \quad (2.54)$$

Comme le terme attractif est plus petit que précédemment en valeur absolue, la valeur d'équilibre de r_s s'en trouve augmentée ; elle vaut maintenant :

$$\left(\frac{r_s}{a_0}\right)_{\text{eq}} \simeq 4.8 , \quad (2.55)$$

d'où l'énergie de cohésion à l'équilibre :

$$|\varepsilon(r_{s, \text{eq}})| \simeq 1.3 \text{ eV} . \quad (2.56)$$

¹⁷ $\varepsilon(r_{s, \text{eq}})$ est évidemment négatif.

¹⁸Rappelons que le calcul a été fait pour N atomes (hypothèse nécessaire pour arriver à (2.41)), et que l'on a considéré un nombre égal d'électrons pour obtenir (2.48). Le calcul ne vaut donc bien que pour les éléments de la première colonne, soit Li, Na, K, Rb, Cs.

¹⁹On note, par comparaison avec (2.42), que la contribution totale des interactions coulombiennes est maintenant égale au terme d'échange ; en effet, le calcul détaillé montre que, dans le cas du *jellium*, les termes directs (positifs) compensent très exactement la contribution (qui se trouve être négative) liée à l'existence du fond continu positif (*self*-répulsion de ce fond et attraction entre les électrons et le fond).

La valeur de $r_{s,eq}$ est nettement plus satisfaisante, mais ne doit pas faire illusion et faire oublier le caractère un peu artificiel du *jellium*.

Clairement, c'est du côté des hautes densités électroniques que le *jellium* trouve sa pleine validité ; en effet, lorsque r_s est petit, c'est l'énergie cinétique ε_{cin} qui est dominante et la prise en compte élémentaire (*i.e.* par perturbations) de l'interaction coulombienne concerne alors des termes énergétiques relativement petits. À l'autre extrême (faibles densités, grand r_s), le calcul précédent tend à montrer (malgré ses limites) que l'énergie potentielle est dominante, l'énergie cinétique est petite (les électrons sont peu mobiles). En 1938, par de tout autres moyens, Wigner a prévu l'existence d'un cristal bidimensionnel triangulaire d'électrons. Cette prévision théorique a reçu une confirmation expérimentale dans les années '80, par la mesure de la modification des ondes capillaires de l'hélium liquide, à la surface duquel un gaz d'électrons avait été déposé. Les altérations observées sont compatibles avec le réseau triangulaire proposé par Wigner.

D'une façon générale, la théorie de perturbation ordinaire pour traiter l'interaction de Coulomb entre électrons n'est pas satisfaisante : le calcul à des "ordres supérieurs" fait apparaître des termes singuliers du genre $r_s \ln r_s$.

Chapitre 3

Structures cristallines

Le cours est centré sur l'étude de la matière ordinaire à l'état cristallin (ordre à longue portée, en fait infinie pour un cristal parfait). De ce fait, il est indispensable de donner les bases de la cristallographie ; c'est l'objet de ce chapitre.

3.1 Réseau de Bravais

Soit, dans \mathbb{R}^3 , trois vecteurs linéairement indépendants \vec{a}_1 , \vec{a}_2 , et \vec{a}_3 . On appelle réseau de Bravais l'ensemble des points dont le rayon-vecteur \vec{R} est de la forme :

$$\boxed{\vec{R} = n_1\vec{a}_1 + n_2\vec{a}_2 + n_3\vec{a}_3} \quad (n_i \in \mathbb{Z}) \quad (3.1)$$

À l'ensemble des vecteurs \vec{R} , on peut associer un ensemble de translations $\mathcal{T}(\vec{R})$, appelé translations du réseau ; toute translation du réseau relie deux points du réseau. On peut définir de la même façon un réseau uni- ou bi-dimensionnel ($d = 1$ et 2 respectivement) ; généralement, un réseau en dimension d peut être noté \mathbb{Z}^d . Un réseau de Bravais est infini dans toutes ses directions et offre exactement la même vision de quelque point qu'on le regarde (de chaque point, on voit exactement le même paysage que de tous les autres) ; chaque point du réseau est aussi appelé nœud. Les vecteurs \vec{a}_i sont appelés vecteurs primitifs, puisqu'ils permettent d'engendrer le réseau par la relation de définition (3.1) ; leur choix n'est pas unique en général, voir l'exemple donné sur la figure 3.1, dans le cas d'un réseau \mathbb{Z}^2 .

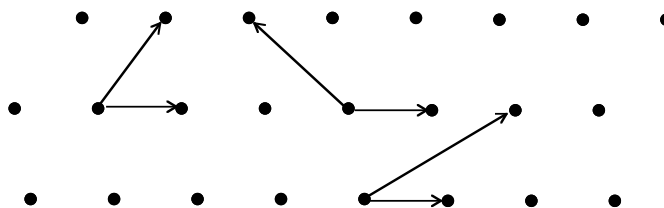


Figure 3.1: Différents choix possibles pour les vecteurs primitifs \vec{a}_i d'un réseau bidimensionnel.

Le mot réseau désigne plus généralement un ensemble infini de points régulièrement répartis dans l'espace suivant une règle précise et non-ambiguë. Tous les réseaux ne sont pas des réseaux de Bravais, l'exemple le plus connu étant le réseau en nid d'abeille (figure 3.2). En effet, si on considère les deux vecteurs \vec{u} et \vec{v} , on peut bien engendrer tout le réseau, mais aussi *en plus* des points (par exemple les centres des hexagones) qui n'appartiennent pas au réseau initialement défini ; il est facile de se convaincre géométriquement qu'il n'existe pas deux vecteurs permettant de construire tous les points du réseau nid d'abeille et eux seulement : le réseau en nid d'abeille n'est pas un réseau de Bravais. Le réseau de Bravais engendré à partir de \vec{u} et \vec{v} (donc avec un nœud au centre de chaque hexagone) s'appelle réseau hexagonal.

L'analyse systématique montre qu'il existe en tout et pour tout 14 réseaux de Bravais distincts (cubique, quadratique, orthorhombique, monoclinique, triclinique), chacun pouvant exister sous différentes formes (par exemple, pour le cubique : simple, centré et à faces centrées).

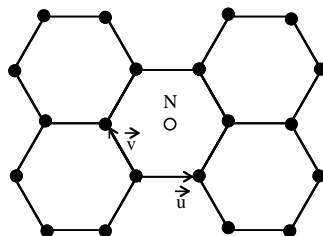


Figure 3.2: Le réseau en nid d'abeille n'est pas un réseau de Bravais : le point N obtenu par la somme $\vec{u} + \vec{v}$ ne lui appartient pas.

Évidemment, outre la symétrie de translation, un réseau peut posséder d'autres éléments de symétrie (rotation, réflexion dans un plan, inversion d'espace, etc.), qui se distinguent des translations en ce sens qu'elles possèdent des points fixes (invariants). L'ensemble de ces opérations constitue la symétrie dite ponctuelle¹. Un point mérite d'être mentionné : la symétrie d'ordre 5 (une rotation de $\frac{2\pi}{5}$ par exemple) est incompatible avec la symétrie de translation d'un réseau de Bravais. On a toutefois découvert récemment (1984) des matériaux remarquables en ce sens qu'ils exhibent une symétrie pentagonale, tout en donnant des clichés diffraction à l'instar d'un réseau de Bravais. Ces systèmes, appelés quasi-cristaux, semblent donc à première vue invalider le dogme de la cristallographie interdisant la symétrie d'ordre 5. On reviendra sur ce point dans la suite (section 3.3).

Les réseaux cubiques sont des réseaux de Bravais. Le plus simple d'entre eux est le cubique simple (C) – un nœud en chacun des huit sommets –, de côté a , engendré par trois vecteurs \vec{a}_i de même longueur a et orthogonaux entre eux ; si $(\vec{i}, \vec{j}, \vec{k})$ désigne une base orthonormalisée de \mathbb{R}^3 , on peut prendre :

$$\vec{a}_1 = a \vec{i} \quad \vec{a}_2 = a \vec{j} \quad \vec{a}_3 = a \vec{k} . \tag{3.2}$$

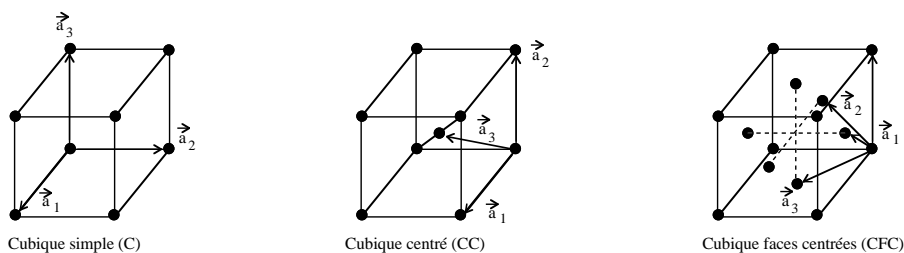


Figure 3.3: Réseaux cubique simple (C), cubique centré (CC) et cubique faces centrées (CFC), avec un choix de vecteurs primitifs.

Un autre réseau cubique est le cubique centré (CC) : il y a un nœud aux huit sommets d'un cube et un autre au centre du cube. Un choix de vecteurs primitifs est :

$$\vec{a}_1 = a \vec{i} \quad \vec{a}_2 = a \vec{j} \quad \vec{a}_3 = \frac{a}{2} (\vec{i} + \vec{j} + \vec{k}) . \tag{3.3}$$

Un autre choix plus symétrique est :

$$\vec{a}'_1 = \frac{a}{2} (-\vec{i} + \vec{j} + \vec{k}) \quad \vec{a}'_2 = \frac{a}{2} (\vec{i} - \vec{j} + \vec{k}) \quad \vec{a}'_3 = \frac{a}{2} (\vec{i} + \vec{j} - \vec{k}) . \tag{3.4}$$

¹le qualificatif veut rappeler, précisément, qu'il existe des points fixes.

Le dernier réseau cubique important est le cubique faces centrées (CFC), qui, outre les huit nœuds aux sommets d'un cube, a un nœud au centre de chacune des 6 faces de ce dernier ; on peut l'engendrer avec les vecteurs primitifs suivants :

$$\vec{a}_1 = a \vec{i} \quad \vec{a}_2 = \frac{a}{2} (\vec{i} + \vec{j}) \quad \vec{a}_3 = \frac{a}{2} (\vec{j} + \vec{k}) . \quad (3.5)$$

À nouveau, il existe des choix plus symétriques, par exemple :

$$\vec{a}_1 = \frac{a}{2} (\vec{i} + \vec{j}) \quad \vec{a}_2 = \frac{a}{2} (\vec{j} + \vec{k}) \quad \vec{a}_3 = \frac{a}{2} (\vec{i} + \vec{k}) . \quad (3.6)$$

Ces trois réseaux jouent un rôle éminent, car, dans les conditions ordinaires, de nombreux matériaux cristallisent dans le système cubique. Toutefois, le cubique simple est très rare, un exemple connu étant la phase dite α du Polonium.

Etant donné un point d'un réseau, le nombre de ses plus proches voisins, souvent noté z , est appelé *coordination*. Pour un cubique simple, $z = 6$ ($z = 2d$). De surcroît, on voit facilement que pour le cubique centré, un atome a 8 premiers voisins (raisonner avec l'atome au centre du cube avec les huit atomes situés aux sommets de celui-ci) ; pour le cubique faces centrées, et raisonnant avec la base de vecteurs primitifs (3.6), on voit que l'atome en $(0, 0, 0)$ a effectivement 12 premiers voisins, repérés par $\frac{a}{2} (\pm \vec{i} \pm \vec{j})$, $\frac{a}{2} (\pm \vec{j} \pm \vec{k})$ et $\frac{a}{2} (\pm \vec{k} \pm \vec{i})$ (trois fois quatre nœuds, les signes étant indépendants). En définitive :

$$\boxed{z_{\text{CC}} = 8 \quad z_{\text{CFC}} = 12} \quad (3.7)$$

Dans ces réseaux, pour a donné, la matière est donc nettement plus dense que dans le cubique simple.

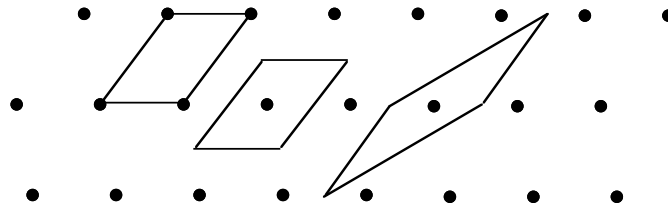


Figure 3.4: Différents choix possibles pour la maille primitive d'un réseau bidimensionnel.

On définit également la maille primitive d'un réseau de Bravais. C'est le volume (ou la surface si $d = 2$) qui, soumis à *toutes* les translations du réseau (construites sur les vecteurs primitifs) remplit complètement l'espace sans aucun recouvrement : c'est la brique élémentaire (incompressible !) pavant l'espace sans laisser le moindre trou. Un réseau étant donné, il y a un choix infini de mailles primitives (voir figure 3.4). Une maille primitive contient exactement un point du réseau et un seul ; tout point situé en surface est partagé entre les mailles adjacentes et compte donc pour une fraction de point – un demi-point s'il est sur la face commune à deux mailles. Soit un ensemble fini de N points construits sur trois vecteurs primitifs, dont le volume est V . La densité de points est donc $n = \frac{N}{V}$; si v désigne le volume d'une maille primitive, on a (aux effets de surface près) :

$$V \simeq N v = (nV) v \quad (N \gg 1) . \quad (3.8)$$

Dans la limite du cristal infini, les effets de surface s'annulent strictement et il vient, après simplification par V :

$$1 = n v \quad \Longleftrightarrow \quad v = \frac{1}{n} . \quad (3.9)$$

Ceci est vrai quelle que soit la maille primitive choisie : toutes les mailles primitives ont donc le même volume v , égal au module du produit mixte construit sur trois vecteurs primitifs :

$$v = |(\vec{a}_1, \vec{a}_2, \vec{a}_3)| . \quad (3.10)$$

On voit en outre que, deux mailles primitives étant choisies, l'une reproduit l'autre exactement quand on la découpe judicieusement en morceaux, auxquels on applique les bonnes translations du réseau.

Une fois définis les vecteurs primitifs, le choix le plus naturel pour la maille primitive est le volume dont tous les points ont pour rayon-vecteur \vec{r} tel que :

$$\vec{r} = x_1 \vec{a}_1 + x_2 \vec{a}_2 + x_3 \vec{a}_3 \quad (0 \leq x_i < 1) . \quad (3.11)$$

Il est d'usage courant, pour la commodité des représentations géométriques, de manipuler des mailles qui ne sont pas primitives au sens où on peut reproduire tout le réseau en leur appliquant seulement un *sous-ensemble* des translations. De telles mailles sont appelées mailles élémentaires conventionnelles. Par exemple, pour le cubique centré, une telle maille est constituée par le cube de côté a , munie de son atome central ; cette maille conventionnelle est deux fois plus grande en volume que toute maille primitive du CC puisqu'elle contient deux points du réseau².

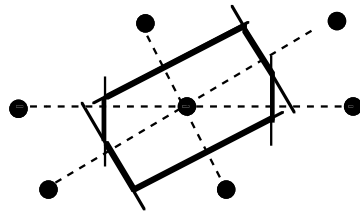


Figure 3.5: La maille primitive de Wigner-Seitz contient tous les points de l'espace plus proches d'un nœud donné que de tous les autres nœuds.

Par ailleurs, il est toujours possible de choisir comme maille primitive une maille ayant par elle-même la symétrie (ponctuelle) du réseau. Un exemple est la maille dite³ de Wigner - Seitz, constituée des points de \mathbb{R}^3 qui sont plus proches d'un nœud donné que de tous les autres nœuds. De façon évidente, il s'agit d'un polyèdre (ou d'un polygone pour $d = 2$) dont les faces (ou les côtés) sont dans les plans médiateurs (ou sur les médiatrices) des segments joignant deux voisins du réseau de Bravais (voir figure 3.5). Il s'agit bien d'une maille primitive, car elle restitue convenablement tout l'espace sans recouvrements quand on lui applique les translations du réseau. Par ailleurs, elle présente l'avantage d'être indépendante du choix des vecteurs primitifs (elle est construite directement sur les points du réseau) et c'est d'ailleurs pour cette raison qu'elle possède la même symétrie que ce réseau.

3.2 Structure physique d'un réseau de Bravais

Un réseau de Bravais est donc une construction purement géométrique dont les principales caractéristiques viennent d'être énoncées et constituée de points mathématiques. Pour définir un système physique construit de la sorte, il faut préciser la nature des objets situés aux nœuds de ce réseau. On appelle base l'unité physique (atome, molécule, ion, *etc.*) attachée à un nœud et dont la duplication à l'infini par translation reproduit le cristal physique, supposé parfait d'un point de vue structural. Dans la figure 3.6, on donne l'exemple d'un matériau où une molécule diatomique hétéronucléaire occupe chaque nœud.

Un autre exemple intéressant (voir fig 3.7) montre comment un réseau en nid d'abeille – qui n'est pas un réseau de Bravais – peut être construit à partir d'un réseau triangulaire muni d'une base à deux points.

L'ensemble réseau et base définit complètement la structure cristalline.

²Le même argument vaut pour le CFC. Si on prend une maille cubique de côté a , elle contient 6 points (ceux au centre de chaque face) partagés entre deux cubes et 8 points (ceux des sommets) partagés chacun entre 8 cubes. Le nombre de points est donc $6 \times \frac{1}{2} + 8 \times \frac{1}{8} = 4$ et le volume de cette maille conventionnelle est $4v_{\text{CFC}}$.

³Dans la littérature mathématique des réseaux de points, la maille de Wigner-Seitz s'appelle maille (ou cellule) de Voronoï. Par ailleurs, cette maille a été introduite en premier par Brillouin pour le réseau réciproque.

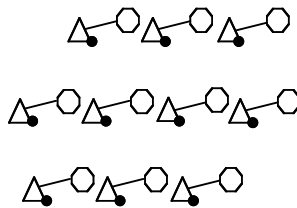


Figure 3.6: Réseau bidimensionnel avec pour base (motif) une molécule diatomique hétéronucléaire.

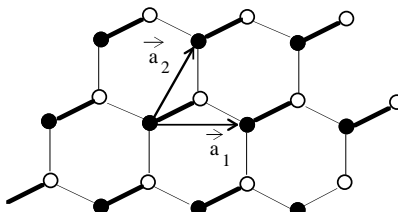


Figure 3.7: Réseau en nid d'abeille construit sur un réseau triangulaire avec deux atomes par nœud.

À l'inverse, pour mettre en avant la symétrie et/ou pour pouvoir raisonner géométriquement plus facilement, certains réseaux de Bravais authentiques peuvent être définis comme des réseaux plus "simples", mais munis d'une base non ponctuelle. Par exemple, le cubique centré peut être défini comme un cubique simple avec sa maille primitive cubique de côté a construite sur une base $(\vec{i}, \vec{j}, \vec{k})$ et muni de la base à deux points de coordonnées :

$$(0, 0, 0) \quad \left(\frac{a}{2}, \frac{a}{2}, \frac{a}{2}\right) . \quad (3.12)$$

De la même façon, le cubique faces centrées peut être construit sur un cubique simple avec une base à quatre points de coordonnées :

$$(0, 0, 0) \quad \left(\frac{a}{2}, \frac{a}{2}, 0\right) \quad \left(\frac{a}{2}, 0, \frac{a}{2}\right) \quad \left(0, \frac{a}{2}, \frac{a}{2}\right) . \quad (3.13)$$

3.3 Quasicristaux

Chapitre 4

Réseau réciproque

4.1 Définition et propriétés du réseau réciproque

Le réseau réciproque d'un réseau de Bravais apparaît spontanément dès que l'on fait de l'analyse de Fourier. En effet, un réseau de points étant donné, l'objet fondamental de l'analyse de Fourier est l'onde plane $e^{i\vec{k}\cdot\vec{r}}$ où \vec{k} est le vecteur d'onde. Par ailleurs, on connaît l'importance du fait que $e^{i\vec{k}\cdot\vec{r}} = 1$: c'est la condition pour avoir des interférences constructives. On en vient ainsi tout naturellement, un réseau de Bravais $\{\vec{R}\}$ étant donné, à considérer l'ensemble des points caractérisés par un vecteur d'onde, noté \vec{K} pour le distinguer, satisfaisant la relation :

$$e^{i\vec{K}\cdot\vec{R}} = 1 . \quad (4.1)$$

L'ensemble des vecteurs \vec{K} est plongé dans \mathbb{R}^3 , mais ce n'est pas l'espace réel – ses vecteurs sont homogènes à l'inverse d'une longueur (comme un vecteur d'onde) – et s'appelle espace réciproque. Il est clair que l'ensemble des points satisfaisant (4.1) ne forme pas un *continuum* mais a la nature discrète d'un réseau, appelé réseau réciproque. En définitive, le réseau réciproque de \mathcal{B} , noté $\tilde{\mathcal{B}}$, est l'ensemble des points extrémités des vecteurs \vec{K} satisfaisant la relation de définition :

$$\boxed{\vec{K} \in \tilde{\mathcal{B}} \iff e^{i\vec{K}\cdot\vec{R}} = 1 \quad \forall \vec{R} \in \mathcal{B}} \quad (4.2)$$

où \mathcal{B} désigne un réseau de Bravais. De la relation de définition (4.2), il résulte immédiatement :

$$\forall \vec{r} \in \mathbb{R}^3 : \quad e^{i\vec{K}\cdot(\vec{R}+\vec{r})} = e^{i\vec{K}\cdot\vec{r}} . \quad (4.3)$$

Une fois introduit le réseau réciproque d'un réseau de Bravais, ce dernier est parfois appelé réseau direct, pour la symétrie du langage.

Il est facile de voir que le réseau réciproque est lui-même un réseau de Bravais. En effet, soit \vec{K}_1, \vec{K}_2 et \vec{K}_3 trois vecteurs linéairement indépendants satisfaisant (4.2). Formons maintenant le vecteur \vec{K} :

$$\vec{K} = m_1 \vec{K}_1 + m_2 \vec{K}_2 + m_3 \vec{K}_3 \quad (m_j \in \mathbb{Z}) . \quad (4.4)$$

Alors, $\forall \vec{R} \in \mathcal{B}$ et puisque les \vec{K}_j satisfont (4.2) et que les m_j sont entiers :

$$e^{i\vec{K}\cdot\vec{R}} = e^{i\sum_{j=1}^3 m_j \vec{K}_j \cdot \vec{R}} = \prod_{j=1}^3 e^{i m_j \vec{K}_j \cdot \vec{R}} = \prod_{j=1}^3 1^{m_j} = \prod_{j=1}^3 1 = 1 . \quad (4.5)$$

Le vecteur \vec{K} satisfait (4.2) lui aussi, c'est donc un vecteur de $\tilde{\mathcal{B}}$; sa construction avec des coordonnées entières selon (4.4) montre que $\tilde{\mathcal{B}}$ est un réseau de Bravais. Il s'agit sûrement de la preuve la plus simple et la plus directe de cette propriété fondamentale du réseau réciproque.

Une autre démonstration passe par un choix de vecteurs primitifs pour le réseau direct \mathcal{B} , supposé à trois dimensions. Ceux-ci étant donnés, on considère les trois vecteurs \vec{b}_i définis comme suit :

$$\boxed{\vec{b}_1 = 2\pi \frac{\vec{a}_2 \times \vec{a}_3}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)} \quad \vec{b}_2 = 2\pi \frac{\vec{a}_3 \times \vec{a}_1}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)} \quad \vec{b}_3 = 2\pi \frac{\vec{a}_1 \times \vec{a}_2}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)}} \quad (4.6)$$

On va montrer que les \vec{b}_i sont des vecteurs primitifs de $\tilde{\mathcal{B}}$. On note d'abord que $\vec{a}_1 \cdot \vec{b}_1 = 2\pi$ et que $\vec{a}_1 \cdot \vec{b}_2 = 0, \dots$, soit plus généralement :

$$\boxed{\vec{a}_i \cdot \vec{b}_j = 2\pi \delta_{ij}} \quad (4.7)$$

Chaque \vec{b}_j est orthogonal aux deux \vec{a}_i qui rentrent dans sa définition (4.6), ce qui est évident par la définition du produit vectoriel. Soit maintenant d'une part un vecteur \vec{k} défini comme une combinaison linéaire des \vec{b}_j :

$$\vec{k} = \sum_{j=1}^3 k_j \vec{b}_j \quad (k_j \in \mathbb{R}) , \quad (4.8)$$

d'autre part un vecteur $\vec{R} \in \mathcal{B}$ d'un réseau de Bravais (direct) :

$$\vec{R} = \sum_{i=1}^3 n_i \vec{a}_i \quad (n_i \in \mathbb{Z}) . \quad (4.9)$$

Alors, compte tenu de (4.7) :

$$\vec{k} \cdot \vec{R} = \sum_{i=1}^3 \sum_{j=1}^3 n_i k_j \vec{a}_i \cdot \vec{b}_j = 2\pi \sum_{i=1}^3 n_i k_i \quad \forall \vec{R} \in \mathcal{B} . \quad (4.10)$$

Prenons maintenant $k_j = m_j \in \mathbb{Z}$; alors la somme $\sum_{j=1}^3 n_i k_j$ est elle-même un nombre entier. Il en résulte que tout vecteur \vec{k} de la forme $\sum_{j=1}^3 m_j \vec{b}_j$ est tel que le produit scalaire $\vec{k} \cdot \vec{R}$ est égal à $2\pi \times$ entier, $\forall \vec{R} \in \mathcal{B}$ ¹ ; autrement dit, ce vecteur est un vecteur du réseau réciproque $\tilde{\mathcal{B}}$, que l'on peut noter \vec{K} :

$$\vec{K} = \sum_{j=1}^3 m_j \vec{b}_j \quad (m_j \in \mathbb{Z}) \iff \vec{K} \in \tilde{\mathcal{B}} . \quad (4.11)$$

Ceci établit que les \vec{b}_j sont bien des vecteurs primitifs de $\tilde{\mathcal{B}}$.

Le produit mixte $\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)$ est, au signe près, le volume de la maille primitive de \mathcal{B} :

$$|\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)| = v . \quad (4.12)$$

Calculons le volume de la maille primitive de $\tilde{\mathcal{B}}$; on a :

$$\vec{b}_2 \times \vec{b}_3 = \frac{(2\pi)^2}{[\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)]^2} [(\vec{a}_3 \times \vec{a}_1) \times (\vec{a}_1 \times \vec{a}_2)] . \quad (4.13)$$

En vertu de la relation :

$$(\vec{U} \times \vec{V}) \times \vec{W} = (\vec{W} \cdot \vec{U}) \cdot \vec{V} - (\vec{V} \cdot \vec{W}) \cdot \vec{U} , \quad (4.14)$$

on a :

$$(\vec{a}_3 \times \vec{a}_1) \times (\vec{a}_1 \times \vec{a}_2) = [(\vec{a}_1 \times \vec{a}_2) \cdot \vec{a}_3] \cdot \vec{a}_1 - [\vec{a}_1 \cdot (\vec{a}_1 \times \vec{a}_2)] \cdot \vec{a}_3 . \quad (4.15)$$

Le dernier crochet est nul, d'où :

$$\vec{b}_2 \times \vec{b}_3 = \frac{(2\pi)^2}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)} \vec{a}_1 . \quad (4.16)$$

¹Le quantificateur $\forall \vec{R} \in \mathcal{B}$ est essentiel. Pour un choix donné des k_i (par exemple des rationnels), on pourrait trouver des n_i particuliers tels que $\sum_{j=1}^3 n_i k_j$ prend une valeur entière.

Il en résulte :

$$\vec{b}_1 \cdot (\vec{b}_2 \times \vec{b}_3) = \frac{(2\pi)^2}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)} \vec{b}_1 \cdot \vec{a}_1 = \frac{(2\pi)^3}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)} , \quad (4.17)$$

la dernière égalité résultant de (4.7). Le volume de $\tilde{\mathcal{B}}$, \tilde{v} , est au signe près le produit mixte $\vec{b}_1 \cdot (\vec{b}_2 \times \vec{b}_3)$:

$$\tilde{v} = |\vec{b}_1 \cdot (\vec{b}_2 \times \vec{b}_3)| . \quad (4.18)$$

v et \tilde{v} sont donc reliés par :

$$\boxed{v \tilde{v} = (2\pi)^3} \quad (4.19)$$

Le produit est bien un nombre pur puisque $[v] = [\text{longueur}^3]$ et que $[\tilde{v}] = [\text{longueur}]^{-3}$.

Le réseau réciproque étant un réseau de Bravais, il possède donc lui aussi un réseau réciproque – qui n'est autre que le réseau direct de départ :

$$\tilde{\tilde{\mathcal{B}}} = \mathcal{B} . \quad (4.20)$$

Ce résultat peut se démontrer de plusieurs façons. Une démonstration directe utilise la définition (4.6) et considère les trois vecteurs \vec{c}_k :

$$\vec{c}_1 = 2\pi \frac{\vec{b}_2 \times \vec{b}_3}{\vec{b}_1 \cdot (\vec{b}_2 \times \vec{b}_3)} \quad \vec{c}_2 = 2\pi \frac{\vec{b}_3 \times \vec{b}_1}{\vec{b}_1 \cdot (\vec{b}_2 \times \vec{b}_3)} \quad \vec{c}_3 = 2\pi \frac{\vec{b}_1 \times \vec{b}_2}{\vec{b}_1 \cdot (\vec{b}_2 \times \vec{b}_3)} . \quad (4.21)$$

D'après (4.16) et (4.17), on a par exemple :

$$\vec{c}_1 = 2\pi \frac{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)}{(2\pi)^3} \frac{(2\pi)^2}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)} \vec{a}_1 = \vec{a}_1 , \quad (4.22)$$

et de même $\vec{c}_2 = \vec{a}_2$, $\vec{c}_3 = \vec{a}_3$. L'identité $\vec{c}_i = \vec{a}_i$ est d'ailleurs évidente géométriquement.

On peut aussi raisonner plus formellement comme suit. Par définition, le réseau réciproque de $\tilde{\mathcal{B}}$ est l'ensemble de tous les vecteurs \vec{T} tels que :

$$e^{i\vec{K} \cdot \vec{T}} = 1 \quad \forall \vec{K} \in \tilde{\mathcal{B}} . \quad (4.23)$$

Par ailleurs, les vecteurs \vec{K} sont tels que :

$$e^{i\vec{K} \cdot \vec{R}} = 1 \quad \forall \vec{R} \in \mathcal{B} . \quad (4.24)$$

Il en résulte que tous les vecteurs \vec{R} appartiennent au réciproque du réciproque :

$$\{\vec{R}\} \subseteq \{\vec{T}\} . \quad (4.25)$$

En outre, soit $\vec{t} = \sum_{i=1}^3 x_i \vec{a}_i$ un vecteur n'appartenant pas à \mathcal{B} (l'un des réels x_i au moins n'est pas entier relatif, par exemple x_1). Prenant $\vec{K} = \vec{b}_1$ et $\vec{T} = \vec{t}$, il vient :

$$e^{i\vec{K} \cdot \vec{T}} = e^{i \sum_i x_i 2\pi \delta_{ij}} = e^{i x_1 2\pi} \neq 1 \quad (x_1 \notin \mathbb{Z}) . \quad (4.26)$$

Ainsi, un vecteur n'appartenant pas à \mathcal{B} n'appartient pas non plus à $\tilde{\tilde{\mathcal{B}}}$. Compte tenu de (4.23) et 4.26, on a donc :

$$\tilde{\tilde{\mathcal{B}}} = \mathcal{B} . \quad (4.27)$$

Toutes les définitions ont été données dans le cas d'un réseau \mathcal{B} à trois dimensions. Elles s'appliquent également en dimension inférieure (en pratique $d = 1$ ou 2). Par exemple, soit à construire le réseau réciproque d'un réseau bidimensionnel engendré par les deux vecteurs primitifs \vec{a}_1 et \vec{a}_2 , faisant entre eux l'angle θ . Rien n'interdit de définir formellement (et provisoirement) une troisième direction, de vecteur primitif \vec{a}_3 ; de surcroît, la prendre perpendiculaire au plan défini par \vec{a}_1 et \vec{a}_2 ne nuit aucunement à la généralité. L'astuce consiste à faire tendre en temps utile, la norme de \vec{a}_3 vers l'infini. Autrement dit, on introduit un réseau monoclinique dont la direction droite est en bout de course dilatée à l'infini.

Compte tenu de la définition (4.6) et du choix de \vec{a}_3 (perpendiculaire aux plans réticulaires), le vecteur \vec{b}_1 est dans le plan (\vec{a}_1, \vec{a}_2) et sa norme est :

$$\|\vec{b}_1\| = 2\pi \frac{a_2 a_3}{a_1 a_2 a_3 \sin \theta} = \frac{2\pi}{a_1 \sin \theta} , \quad (4.28)$$

et, de même :

$$\|\vec{b}_2\| = 2\pi \frac{a_3 a_1}{a_1 a_2 a_3 \sin \theta} = \frac{2\pi}{a_2 \sin \theta} \quad \|\vec{b}_3\| = 2\pi \frac{a_1 a_2 \sin \theta}{a_1 a_2 a_3 \sin \theta} = \frac{2\pi}{a_3} . \quad (4.29)$$

On prend maintenant la limite $a_3 \rightarrow +\infty$; la norme de \vec{b}_3 s'annule et on obtient² le réseau réciproque, de dimension 2 et situé géométriquement dans le même plan que le réseau direct ; \vec{b}_1 et \vec{b}_2 sont orthogonaux respectivement à \vec{a}_2 et \vec{a}_1 . Enfin, pour un réseau unidimensionnel engendré par \vec{a} , le réseau réciproque engendré par le seul et unique vecteur \vec{b} , colinéaire à \vec{a} (il reste orthogonal à deux dimensions fictives que l'on peut toujours choisir orthogonales à \vec{a}) et de norme $b = 2\pi/a$.

◆ *Remarque*

Affirmer que le réseau réciproque d'un réseau de Bravais bidimensionnel est de dimension $d = 2$ est conforme à l'usage ; toutefois, si l'on s'en tient à la définition première (4.1), il est plus logique de dire que, en pareil cas, \vec{B} est constitué de droites infinies perpendiculaires au réseau direct et dont les intersections avec \mathbb{R}^2 constituent le réseau réciproque bidimensionnel défini par le processus de limite ci-dessus. En effet, soit un vecteur \vec{K} ainsi défini dans le plan ($\vec{K} \in \vec{B}_{d=2} \equiv \mathbb{Z}^2$) et satisfaisant donc :

$$e^{i\vec{K} \cdot \vec{R}} = 1 \quad \forall \vec{R} \in \mathcal{B}_{d=2} \equiv \mathbb{Z}^2 . \quad (4.30)$$

Alors, si on ajoute à \vec{K} (dans le plan) n'importe quel vecteur perpendiculaire \vec{K}_\perp à ce plan :

$$\vec{K}' = \vec{K} + \vec{K}_\perp , \quad (4.31)$$

l'égalité (4.30) sera encore satisfaite. Le même argument permet de dire que le réseau réciproque d'un réseau de Bravais unidimensionnel $\mathcal{B}_{d=1} \equiv \mathbb{Z}$ de pas a est un ensemble de plans parallèles perpendiculaires à ce réseau. Le réseau direct coupe cette famille de plans en des points équidistants d'abscisses $m(2\pi)/a$, $m \in \mathbb{Z}$.

Ces distinctions peuvent jouer un rôle important quand la loi de conservation de l'impulsion est pertinente. Par exemple, si des photons (définis dans \mathbb{R}^3 !) d'impulsion $\hbar\vec{k}$ sont couplés à de la matière bidimensionnelle, la loi de conservation de l'impulsion ne contraint que la *projection* de \vec{k} : tous les photons ne différant que par un vecteur quelconque perpendiculaire au plan de la matière satisfont la conservation de l'impulsion. Autrement dit, une excitation matérielle d'impulsion $\hbar\vec{q}$ se trouve couplée à un *continuum* de photons de vecteur d'onde $\vec{k} = \vec{q} + \vec{q}_\perp$, avec $\vec{q} \cdot \vec{q}_\perp = 0$.

◆

4.2 Exemples de réseaux réciproques

Donnons maintenant quelques exemples de réseaux réciproques. Il est évident que le réseau cubique simple de paramètre a a pour réseau réciproque un réseau également cubique simple, de côté $2\pi/a$. Pour le cubique centré, choisissons les vecteurs primitifs suivants :

$$\vec{a}_1 = a\vec{i} \quad \vec{a}_2 = a\vec{j} \quad \vec{a}_3 = \frac{a}{2}(\vec{i} + \vec{j} + \vec{k}) . \quad (4.32)$$

On a $(\vec{a}_1 \times \vec{a}_2) \cdot \vec{a}_3 = a^2(a/2) = a^3/2$. Maintenant :

$$\vec{b}_1 = \frac{2\pi}{a^3/2} \frac{a^2}{2} \vec{j} \times (\vec{i} + \vec{j} + \vec{k}) = \frac{2\pi}{a} (\vec{i} - \vec{k}) , \quad (4.33)$$

²En réalité, les choses sont plus subtiles. Il est sûr que si l'on prend la limite $a_3 \rightarrow +\infty$, tout point d'abscisse entière *finie* se retrouve dans le plan. Mais *quid* des points à l'infini ?

$$\vec{b}_2 = \frac{2\pi}{a^3/2} \frac{a^2}{2} (\vec{i} + \vec{j} + \vec{k}) \times \vec{i} = \frac{2\pi}{a} (\vec{j} - \vec{k}) , \quad (4.34)$$

$$\vec{b}_3 = \frac{2\pi}{a^3/2} a^2 \vec{i} \times \vec{j} = \frac{4\pi}{a} \vec{k} . \quad (4.35)$$

De façon équivalente, on peut aussi introduire :

$$\vec{b}'_1 = \vec{b}_1 + \vec{b}_3 = \frac{2\pi}{a} (\vec{k} + \vec{i}) , \quad \vec{b}'_2 = \vec{b}_2 + \vec{b}_3 = \frac{2\pi}{a} (\vec{j} + \vec{k}) , \quad \vec{b}'_3 = \vec{b}_1 + \vec{b}_2 + \vec{b}_3 = \frac{2\pi}{a} (\vec{i} + \vec{j}) , \quad (4.36)$$

et on reconnaît en $(\vec{b}'_1, \vec{b}'_2, \vec{b}'_3)$ un jeu de vecteurs primitifs du CFC de maille cubique de côté $4\pi/a$. La maille primitive du CC direct a pour volume $v = a^3/2$, celle du réseau réciproque a pour volume $\tilde{v} = (2\pi)^3/(a^3/2) = 16\pi^3/a^3$. D'un autre côté, le volume de la maille primitive d'un CFC de côté $4\pi/a$ est $(1/4)(4\pi/a)^3 = 16\pi^3/a^3 = \tilde{v}$, comme il se doit.

Comme le réseau réciproque du réseau réciproque est le réseau direct, le réseau réciproque d'un CFC est un réseau CC. Si le volume v du CFC est $\frac{a^3}{4}$, celui de son réseau réciproque est $\frac{(2\pi)^3}{a^3/4} = \frac{32\pi^3}{a^3} \equiv \frac{1}{2} \left(\frac{4\pi}{a}\right)^3$. Le réciproque d'un CFC de côté a est donc le CC de côté $\frac{4\pi}{a}$.

Comme dernier exemple, prenons un réseau hexagonal, engendré par deux vecteurs \vec{a}_1 et \vec{a}_2 de même norme et faisant entre eux un angle de $\frac{\pi}{3}$, et par un vecteur \vec{c} perpendiculaire au plan défini par \vec{a}_1 et \vec{a}_2 et de longueur arbitraire c . Pour faire le calcul, on définit un repère orthonormé $(\vec{i}, \vec{j}, \vec{k})$ sur lequel :

$$\vec{a}_1 = a\vec{i} \quad \vec{a}_2 = \frac{a}{2} (\vec{i} + \sqrt{3}\vec{j}) \quad \vec{a}_3 = c\vec{k} . \quad (4.37)$$

On trouve alors :

$$\vec{a}_2 \times \vec{a}_3 = \frac{ac}{2} (\sqrt{3}\vec{i} - \vec{j}) \quad \vec{a}_3 \times \vec{a}_1 = ac\vec{j} \quad \vec{a}_1 \times \vec{a}_2 = a^2 \frac{\sqrt{3}}{2} \vec{k} , \quad (4.38)$$

et :

$$(\vec{a}_1 \times \vec{a}_2) \cdot \vec{a}_3 = \frac{\sqrt{3}}{2} a^2 c . \quad (4.39)$$

Il en résulte :

$$\vec{b}_1 = \frac{2\pi}{a\sqrt{3}} (\sqrt{3}\vec{i} - \vec{j}) \quad \vec{b}_2 = \frac{4\pi}{a\sqrt{3}} \vec{j} \quad \vec{b}_3 = \frac{2\pi}{c} \vec{k} . \quad (4.40)$$

Ces résultats montrent que le réseau réciproque de l'hexagonal (a, c) est un réseau hexagonal $(4\pi/a\sqrt{3}, 2\pi/c)$, simplement tourné de $\pi/6$ par rapport au réseau direct. Si on prend le réciproque de ce dernier, il faut encore tourner de $\pi/6$ et on retrouve le réseau de départ. Enfin, en faisant tendre c vers l'infini pour construire le réseau réciproque du réseau bidimensionnel, la longueur du vecteur \vec{b}_3 tend vers zéro : les points de $\tilde{\mathcal{B}}$ le long de \vec{k} deviennent infiniment proches les uns des autres et engendrent un pseudo-droite ayant la puissance de \mathbb{Z} (voir la note 2).

Comme le réseau réciproque est un réseau de Bravais, tout ce qui a été dit concernant le choix des vecteurs primitifs, la maille primitive, la maille conventionnelle, *etc.* est applicable à $\tilde{\mathcal{B}}$. En particulier, on peut définir une maille du genre Wigner - Seitz pour $\tilde{\mathcal{B}}$. Pour des raisons qui seront plus claires dans la suite, la maille de Wigner - Seitz du réseau réciproque est appelée 1^{ère} zone de Brillouin (en abrégé : BZ1). Ainsi, ce que l'on appelle 1^{ère} zone de Brillouin d'un CFC est géométriquement identique à la maille de Wigner - Seitz d'un CC.

4.3 Plans réticulaires

On appelle plan réticulaire un plan contenant au moins trois nœuds d'un réseau de Bravais, c'est-à-dire, partant de là, une infinité. Il existe une relation très étroite entre les plans réticulaires et les vecteurs du réseau réciproque ; cette relation joue un rôle important pour les phénomènes de diffraction et peut être appréhendée sur une stricte base géométrique.

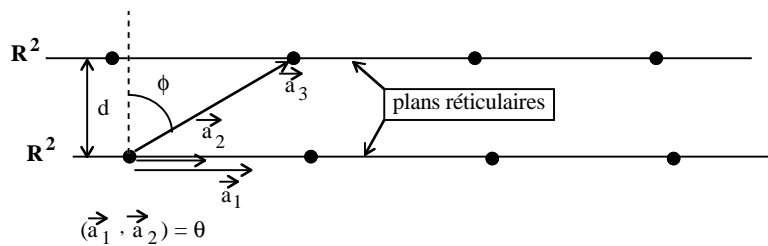


Figure 4.1: La distance d entre deux plans adjacents d'une même famille contenant \vec{a}_1 et \vec{a}_2 est simplement reliée au module de \vec{b}_3 (voir (4.42)).

Trois nœuds définissent un plan ; tous les plans parallèles à un même plan forment une famille de plans, équidistants les uns des autres d'une distance d dépendant de la famille considérée. Il y a manifestement une infinité de façons de résoudre un réseau en un ensemble de plans réticulaires.

Un plan réticulaire est un sous-réseau de dimension 2, où on peut définir deux vecteurs primitifs notés \vec{a}_1 et \vec{a}_2 . Par définition, le vecteur \vec{b}_3 est orthogonal à \vec{a}_1 et \vec{a}_2 , c'est donc une normale à ce plan. Prenons pour \vec{a}_3 le vecteur indiqué sur la figure 4.1. Si θ est l'angle entre \vec{a}_1 et \vec{a}_2 et ϕ l'angle entre \vec{a}_3 et la normale au plan, on a :

$$\|\vec{b}_3\| = 2\pi \frac{|a_1 a_2 \sin \theta|}{|\vec{a}_3 \cdot (\vec{a}_1 \times \vec{a}_2)|} = 2\pi \frac{a_1 a_2 |\sin \theta|}{a_1 a_2 a_3 |\sin \theta \cos \phi|} . \quad (4.41)$$

Comme $a_3 |\cos \phi| = d$, distance entre deux plans adjacents, il vient finalement :

$$\|\vec{b}_3\| = \frac{2\pi}{d} \quad (4.42)$$

Ainsi, une famille de plans étant précisée, il existe des vecteurs du réseau réciproque orthogonaux à cette famille et le plus petit d'entre eux³ a une norme s'exprimant simplement, par (4.42), en fonction de la distance entre deux plans consécutifs. Clairement, l'argument précédent ne privilégie pas un certain vecteur primitif de $\vec{\mathcal{B}}$ (celui noté \vec{b}_3 ci-dessus) et vaut pour l'ensemble du réseau réciproque : tout en conservant les \vec{a}_i précédents (donc les mêmes \vec{b}_j), mais en choisissant une autre famille de plans, on montrerait de la même façon qu'il existe un vecteur \vec{K} orthogonal à cette famille et ayant *a priori* trois composantes non-nulles sur ces \vec{b}_j .

La relation étroite entre une famille de plans réticulaires et les vecteurs du réseau réciproque permet de définir commodément l'orientation de la famille considérée. Comme il est usuel de définir un plan par sa normale, on peut se donner un certain vecteur du réseau réciproque pour qualifier une famille de plans. On en vient ainsi à ce que l'on appelle les indices de Miller d'un plan : ce sont les composantes du plus petit vecteur de $\vec{\mathcal{B}}$ normal à ce plan, dont la longueur est de ce fait reliée à la distance entre deux plans réticulaires voisins selon (4.42) ; on les note traditionnellement (h, k, l) et la distance correspondante est notée d_{hkl} . Ainsi, un plan dont les indices de Miller sont (h, k, l) est un plan normal au vecteur $\vec{K} = h\vec{b}_1 + k\vec{b}_2 + l\vec{b}_3$. Les indices de Miller sont donc des entiers et comme il s'agit du plus petit vecteur réciproque, les trois nombres h, k et l sont premiers entre eux (par exemple, ils ne sauraient être tous pairs). Bien sûr, le triplet (h, k, l) dépend du choix des vecteurs primitifs \vec{b}_j .

Les indices de Miller ont une interprétation géométrique simple dans le réseau direct, laquelle sert d'ailleurs parfois de définition. Etant donné (h, k, l) , l'équation du plan perpendiculaire au vecteur réciproque correspondant, \vec{K} , est $\vec{K} \cdot \vec{r} = C$, avec $\vec{K} = h\vec{b}_1 + k\vec{b}_2 + l\vec{b}_3$ et C une constante. Ce plan coupe les axes portant les \vec{a}_i en trois points d'abscisses x_1, x_2 et x_3 . Ces points, appartenant au plan, satisfont son équation et on a donc :

$$\vec{K} \cdot (x_i \vec{a}_i) = C \quad (i = 1, 2, 3) . \quad (4.43)$$

³ \vec{b}_3 est bien le plus petit vecteur réciproque normal aux plans considérés.

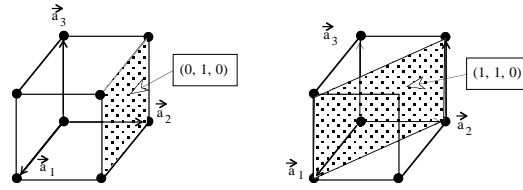


Figure 4.2: Exemples d'indices de Miller pour un réseau cubique simple

Comme $\vec{K} \cdot \vec{a}_i = (h\vec{b}_1 + k\vec{b}_2 + l\vec{b}_3) \cdot \vec{a}_i$ et compte tenu de (4.7), on trouve :

$$2\pi x_1 h = 2\pi x_2 k = 2\pi x_3 l = C \quad , \quad (4.44)$$

soit :

$$x_1 = \frac{C}{2\pi h} \quad x_2 = \frac{C}{2\pi k} \quad x_3 = \frac{C}{2\pi l} \quad . \quad (4.45)$$

Les coordonnées des points d'intersection des plans d'une famille sont donc inversement proportionnelles aux indices de Miller (voir figure 4.2).

Chapitre 5

Diffraction

5.1 Diffusion des rayons X et des neutrons

La diffusion des rayons X et des neutrons¹ est un outil de tout premier plan pour étudier la structure de la matière, qu'elle soit en phase diluée ou condensée. Le principe en est simple : on envoie sur le système à étudier une "sonde" externe parfaitement caractérisée (énergie ou longueur d'onde, polarisation²). L'interaction sonde - système est complexe en général ; pour simplifier, seules les interactions *élastiques* seront considérées dans la suite³. Dans ce cadre, la "lumière" issue du milieu a la même fréquence (énergie) que l'onde incidente ; son vecteur d'onde \vec{k}_f a donc le même module que celui du faisceau incident, \vec{k}_i :

$$\boxed{\text{diffusion élastique} \iff \|\vec{k}_f\| = \|\vec{k}_i\|} \quad (5.1)$$

En tout cas, l'interaction sonde - cible est supposée faible, puisque l'on veut une photographie de la matière telle qu'elle est, c'est-à-dire perturbée aussi peu que possible.

5.1.1 Rayons X

Dans le cas des rayons X, l'énergie des photons est de quelques dizaines de keV, très supérieure aux énergies intra-atomiques ou intra-moléculaires, qui sont de l'ordre de quelques eV au plus. Le rayonnement électromagnétique incident, essentiellement par son champ électrique (de module \mathcal{E}_0), met en vibration forcée, à la même fréquence, les électrons du milieu (qui lui apparaissent comme presque libres) : classiquement, c'est ce que l'on appelle généralement la diffusion Thomson⁴. Pour une onde incidente non polarisée, la diffusion Thomson est caractérisée à haute fréquence par une section efficace indépendante de la fréquence, $\sigma_T = (8\pi/3) r_e^2$, où r_e est appelé historiquement le "rayon classique" de l'électron ($r_e = e^2/(4\pi\epsilon_0 mc^2) \approx 3 \times 10^{-15} \text{ m}$)⁵ ; r_e est la longueur la plus simple que l'on peut former avec la vitesse de la lumière, la masse et la charge de l'électron. La vibration forcée donne aux électrons une accélération non-nulle et, comme toute particule chargée accélérée rayonne, ceux-ci émettent un rayonnement électromagnétique, évidemment de même fréquence puisqu'il

¹et des électrons pour les surfaces.

²Dans le cas des neutrons, la polarisation est liée au spin ($\pm\hbar/2$) des neutrons incidents ; un faisceau non polarisé est un faisceau dans lequel l'orientation du spin est aléatoire.

³Bien sûr, les expériences de diffusion inélastique sont aussi d'usage courant et apportent d'autres informations.

⁴Il existe également une diffusion inélastique, appelée diffusion Compton, qui peut donner de précieux renseignements sur la distribution des vitesses au sein du système.

⁵La section efficace Thomson est donc microscopique à l'échelle atomique – c'est une surface dont le rayon est de l'ordre des dimensions nucléaires. Au contraire, la section efficace de diffusion *résonnante* est, elle, gigantesque par rapport à la taille atomique puisqu'elle vaut en gros λ_0^2 , où λ_0 est la longueur d'onde de résonance, qui est dans le domaine optique ou UV pour les transitions électroniques ordinaires des atomes et des molécules.

s'agit d'une vibration forcée. La puissance rayonnée vers l'extérieur par une charge accélérée, \mathcal{P} , est donnée par la formule classique de Larmor :

$$\mathcal{P} = \frac{2}{3} \frac{e'^2 \dot{v}^2}{c^3} \quad e'^2 \stackrel{\text{déf}}{=} \frac{e^2}{4\pi\epsilon_0} \quad (5.2)$$

En ne retenant que les effets majoritaires, l'accélération d'un électron, \vec{v} , se déduit immédiatement du champ électrique de l'onde qui le met en mouvement⁶. Pour un électron accéléré bien localisé dans un domaine très petit par rapport à la longueur d'onde incidente, le champ électrique qu'il rayonne à une grande distance R a pour amplitude :

$$\mathcal{E} = \frac{r_e}{R} \mathcal{E}_0 \quad , \quad (5.3)$$

\mathcal{E}_0 étant l'amplitude du champ externe. En réalité, pour les atomes et les molécules, les électrons sont répartis sur des distances de l'ordre de l'Å ; le caractère diffus de la densité électronique produit un premier type de déphasage spatial et donne lieu à des effets d'interférences dont il faut tenir compte⁷. Ceci introduit un facteur multiplicatif, usuellement noté f et appelé facteur de structure (intra-atomique) ; c'est essentiellement la transformée de Fourier de la densité électronique, calculée pour le transfert de moment $\vec{q} = \vec{k}_f - \vec{k}_i$; $f(\vec{q} = 0)$ est égal à Z , nombre total d'électrons appartenant au centre diffuseur.

En raisonnant d'emblée avec les composantes de Fourier, il convient de trouver les déphasages relativement à un point arbitraire. Par rapport à l'origine $\vec{r} = 0$, le déphasage de la lumière émise par un seul centre diffuseur de facteur de structure f_1 situé au point \vec{r}_1 est $f_1 e^{i(\vec{k}_f - \vec{k}_i) \cdot \vec{r}_1}$; l'amplitude (complexe) du champ électrique émis par vibration forcée est, à la distance R , donnée par⁸ :

$$\mathcal{E}_1 = \mathcal{E}_0 f_1 \frac{r_e}{R} e^{i(\vec{k}_f - \vec{k}_i) \cdot \vec{r}_1} \quad . \quad (5.4)$$

Pour un ensemble de N diffuseurs situés aux points $\vec{r}_n, n = 1, 2, \dots, N$, l'amplitude diffusée est la somme :

$$\mathcal{E} = \mathcal{E}_0 \frac{r_e}{R} \sum_{n=1}^N f_n e^{i(\vec{k}_f - \vec{k}_i) \cdot \vec{r}_n} \quad . \quad (5.5)$$

Ceci suppose que le champ est polarisé rectilignement, mettant tous les oscillateurs en vibration forcée le long de la *même* direction. Le carré du champ électrique rayonné vaut donc :

$$\mathcal{E}^2 = \left(\mathcal{E}_0 \frac{r_e}{R} \right)^2 \sum_{n,m=1}^N f_n f_m^* e^{i(\vec{k}_f - \vec{k}_i) \cdot (\vec{r}_n - \vec{r}_m)} \quad . \quad (5.6)$$

Enfin, si la source n'est pas polarisée, il convient de moyenner sur les angles, ce qui revient à moyenner sur toutes les directions possibles les petites "antennes" électroniques. Comme toujours, on suppose que l'absence de polarisation de l'onde incidente se traduit par une équiprobabilité des directions de vibration de son champ électrique. \vec{k}_f et \vec{k}_i étant fixés une fois pour toutes et si θ désigne l'angle de diffusion, $\theta = (\vec{k}_i, \vec{k}_f)$, on trouve que le carré du module du champ moyen est :

$$\bar{\mathcal{E}}^2 = \frac{1}{2} (1 + \cos^2 \theta) \left(\mathcal{E}_0 \frac{r_e}{R} \right)^2 \sum_{n,m=1}^N f_n f_m^* e^{i(\vec{k}_f - \vec{k}_i) \cdot (\vec{r}_n - \vec{r}_m)} \quad . \quad (5.7)$$

En introduisant finalement l'angle $\alpha_{nm} = (\vec{k}_f - \vec{k}_i, \vec{r}_n - \vec{r}_m)$, en désignant par λ la longueur d'onde incidente et en posant $r_{nm} = \|\vec{r}_n - \vec{r}_m\|$, l'intensité diffusée I peut se mettre sous la forme ($\|\vec{q}\| = 2\|\vec{k}_i\| \sin(\theta/2)$) :

$$I = I_0 \sum_{n,m=1}^N f_n f_m^* e^{iK r_{nm} \cos \alpha_{nm}} \quad , \quad K = \frac{4\pi}{\lambda} \sin \frac{\theta}{2} \quad . \quad (5.8)$$

Le cas de diffuseurs immobiles situés aux nœuds d'un réseau sera traité dans la section suivante. Dans un milieu désordonné (un fluide par exemple, ou un verre), l'intensité mesurée correspond en fait à la moyenne

⁶Pour une onde polarisée, l'électron est mis en vibration le long de la direction fixée par la polarisation électrique.

⁷Dans l'optique ou le visible, ce déphasage est négligeable pour les transitions de type dipolaire électrique.

⁸On omet tous les facteurs temporels $e^{i\omega t}$ qui, pour les processus élastiques, seront communs à toutes les ondes diffusées. Le calcul détaillé du déphasage, faisant apparaître le transfert de moment, est rappelé plus loin (section 5.2, figure 5.1).

des orientations de $\vec{r}_n - \vec{r}_m$. L'exponentielle portant les déphasages spatiaux dans (5.8) doit donc être moyennée sur l'angle solide $d\Omega = 2\pi \sin \alpha_{nm} d\alpha_{nm}$; en admettant un désordre complet, pour toute distance r_{nm} donnée, toutes les orientations possibles sont équiprobables ; dans ces conditions, cette moyenne s'écrit :

$$\int \frac{d\Omega}{4\pi} e^{iKr_{nm} \cos \alpha_{nm}} , \quad (5.9)$$

et se calcule aisément :

$$\frac{1}{4\pi} \int_0^\pi 2\pi \sin \alpha d\alpha e^{iKr_{nm} \cos \alpha} = \frac{\sin Kr_{nm}}{Kr_{nm}} . \quad (5.10)$$

L'intensité moyenne est ainsi donnée par :

$$\bar{I} = I_0 \sum_{n,m=1}^N f_n f_m^* \frac{\sin Kr_{nm}}{Kr_{nm}} . \quad (5.11)$$

Ceci constitue la formule de Debye, à laquelle on peut donner une forme plus agréable. On commence par isoler les termes carrés $n = m$ ($\sum_{n,m=1}^N = \sum_{n=1}^N + \sum_{n \neq m=1}^N$) ; dans l'hypothèse d'un échantillon de taille macroscopique (où l'on peut oublier les effets de surface), tout se passe comme s'il existait une quasi-invariance par translation : il en résulte que la somme double est extensive en N . Enfin, en prenant des diffuseurs monoatomiques tous identiques ($f_n = f \forall n$), (5.11) prend finalement la forme :

$$\bar{I} = N I_0 |f|^2 \left(1 + \sum_n \frac{\sin Kr_n}{Kr_n} \right) . \quad (5.12)$$

La somme discrète peut toujours s'écrire sous forme intégrale. En effet, s'il s'agit de se borner aux propriétés à grande échelle, on peut définir une densité ρ : c'est, pour un volume mésoscopique δV , le nombre de particules dans ce volume "élémentaire" divisé par δV et alors une description pseudo-continue peut être utilisée ; si l'aspect granulaire de la matière doit être conservé, il suffit d'introduire des fonctions de Dirac pour représenter les densités atomiques.

Notant $G(r)$ le nombre de particules par unité de volume situées à la distance r de l'origine⁹, l'intensité moyenne (5.12) s'écrit comme suit dans la description continue¹⁰ :

$$\bar{I} = N I_0 |f|^2 \left(1 + \int_0^{\mathcal{R}} 4\pi r^2 dr G(r) \frac{\sin Kr}{Kr} \right) , \quad (5.13)$$

où \mathcal{R} représente le rayon de l'échantillon, supposé sphérique. Pour extraire la composante représentant la diffusion triviale vers l'avant ($\vec{q} = \vec{0}$), il suffit d'ajouter et de retrancher la densité moyenne ρ :

$$\bar{I} = N I_0 |f|^2 \left[1 + \int_0^{\mathcal{R}} 4\pi r^2 dr [G(r) - \rho] \frac{\sin Kr}{Kr} + \rho \int_0^{\mathcal{R}} 4\pi r^2 dr \frac{\sin Kr}{Kr} \right] . \quad (5.14)$$

Pour un échantillon macroscopique (où l'on peut oublier les effets de surface), la dernière intégrale peut être remplacée par $(2\pi)^3 \delta(\vec{K})$; ce terme représente la diffusion vers l'avant, sans intérêt car ne contenant aucune information relative au système étudié. Au contraire, cette dernière est contenue dans les deux premiers termes du crochet ; on définit ainsi précisément, en posant $G(r) = \rho g(r)$:

$$S(K) = 1 + \rho \int_0^{+\infty} 4\pi r^2 dr [g(r) - 1] \frac{\sin Kr}{Kr} \quad (5.15)$$

qui est réellement la signature du système (on voit apparaître la fonction $h(r) = g(r) - 1$, qui tend vers zéro aux grands r , une condition nécessaire pour pouvoir lui associer une transformée de Fourier). En jouant avec la symétrie sphérique, on voit que $S(\vec{K})$ (qui, en fait, ne dépend ici que de $\|\vec{K}\|$) est aussi :

$$S(\vec{K}) = 1 + \rho \int d^3r [g(r) - 1] e^{i\vec{K} \cdot \vec{r}} , \quad (5.16)$$

⁹La fonction G est supposée à symétrie sphérique. C'est en fait le nombre de particules situé à la distance r de l'une d'entre elles, c'est donc la fonction de corrélation de paire.

¹⁰ $\int_{0 \leq r \leq \mathcal{R}} d^3r e^{i\vec{K} \cdot \vec{r}} = \int_0^{\mathcal{R}} 4\pi r^2 dr (\sin Kr)/(Kr)$

ce qui permet d'effectuer à vue l'inversion de Fourier pour obtenir :

$$g(r) = 1 + \frac{1}{8\pi^3\rho} \int d^3K [S(\vec{K}) - 1] e^{-i\vec{K}\cdot\vec{r}} \quad (5.17)$$

et en déduire finalement la relation d'intérêt pratique :

$$g(r) - 1 \equiv h(r) = \frac{1}{2\pi^2\rho r} \int_0^{+\infty} dK [S(K) - 1] K \sin Kr . \quad (5.18)$$

En effet, en mesurant $S(K)$ avec une très bonne résolution en K (et en explorant un grand intervalle en K), on peut reconstruire $g(r)$ par la transformation intégrale (5.18). Cette relation joue donc un rôle majeur pour toutes les expériences de diffusion X.

5.1.2 Neutrons

La diffusion des neutrons¹¹ relève de la même description. La longueur d'onde à prendre en compte est la longueur d'onde associée par la relation de de Broglie $\lambda = \frac{h}{Mv} = \frac{h}{\sqrt{2ME}}$; pour les neutrons ceci s'exprime commodément comme suit :

$$\lambda_{\text{\AA}} \simeq \frac{0.286}{\sqrt{E_{\text{eV}}}} \quad (5.19)$$

Pour un neutron thermique, $E \approx (k_B T)_{T=300\text{K}} \approx 25$ meV, de sorte que $\lambda \approx 1.8$ \AA. C'est dire que ce sont les neutrons thermiques qu'il convient d'utiliser pour sonder la matière ordinaire, qu'elle soit à l'état fluide ou solide. Bien sûr, un faisceau neutronique est en général non monochromatique et la distribution du module de la vitesse (une distribution de Maxwell) reflète l'état thermique du modérateur ; si celui-ci est à une température de l'ordre de 100 K, le maximum de la maxwellienne survient précisément pour E de l'ordre de 25 meV.

Les neutrons incidents et les constituants du fluide peuvent interagir principalement de deux façons, étant entendu que le neutron est insensible à l'interaction électromagnétique¹². La première, relevant de l'interaction forte, est à très courte portée (la dimension nucléaire) et peut être considérée comme étant de contact¹³ pour la diffusion sur des atomes ou des molécules. Le neutron, grâce à son moment magnétique associé à son spin ($S = \frac{1}{2}$ en unités \hbar), peut en outre se coupler aux électrons non appariés, auquel cas la portée de l'interaction effective est d'ordre atomique (l'échelle de longueur des fonctions d'onde de ces électrons).

En ce qui concerne le couplage de contact résultant de l'interaction forte, on l'écrit usuellement :

$$V_c(\vec{r}) = \frac{2\pi\hbar^2}{M} b \delta(\vec{r} - \vec{R}) \quad (5.20)$$

C'est ce que l'on appelle le pseudo-potentiel de Fermi. \vec{r} et \vec{R} sont respectivement les positions du neutron et du noyau diffuseur ; b est la longueur de diffusion¹⁴, qui varie d'un noyau à l'autre (et donc en particulier d'un isotope à l'autre) et dépend également du spin nucléaire. De la forme (5.20), il résulte que pour un diffuseur unique les sections efficaces différentielle et totale sont (voir (5.54)) :

$$\frac{d\sigma}{d\Omega} = b^2 , \quad \sigma = 4\pi b^2 . \quad (5.21)$$

b rentre dans ces expressions de la même façon que le rayon d'une sphère dure pour un problème classique de diffusion¹⁵.

¹¹Les techniques de diffusion neutronique sont nettement plus récentes que celles utilisant les rayons X ; elles ont commencé à se développer dans les années 50, lorsque des sources de neutrons suffisamment intenses sont devenues opérationnelles. Le formalisme théorique de base est principalement dû à Fermi et Schwinger.

¹²en mettant de côté des effets extrêmement fins impliquant la structure de charge du neutron.

¹³Une interaction est dite *de contact* si elle est de portée nulle (les objets n'interagissent que s'ils sont au contact l'un de l'autre) ; on la représente alors par une fonction de Dirac.

¹⁴Dans \mathbb{R}^3 , la fonction $\delta(\vec{r})$ est homogène à L^{-3} ; il en résulte que b a bien la dimension d'une longueur.

¹⁵au facteur 4 près lié à la spécificité quantique (voir par exemple [8], complément CVIII).

En ce qui concerne le couplage magnétique, notamment avec les spins des électrons célibataires, l'interaction est principalement de la forme dipolaire magnétique :

$$V_{\text{magn}} = -\gamma\mu_N \vec{\sigma} \cdot \vec{B} \equiv -\frac{2\gamma}{\hbar} \mu_N \vec{S} \cdot \vec{B} , \quad (5.22)$$

où $\gamma = -1.91$ est le facteur gyromagnétique du neutron, $\mu_N = |e|\hbar/(2M) > 0$ est le magnéton nucléaire et \vec{B} le champ magnétique produit par les électrons non-appariés – dont l'expression est loin d'être simple [5] : elle contient non seulement le champ dipolaire dû aux spins électroniques (couplage spin-spin) mais également le champ magnétique créé par le mouvement orbital des électrons (couplage spin-orbite). $\vec{\sigma}$ est l'opérateur vectoriel construit avec les trois matrices de Pauli.

D'une façon ou d'une autre, la diffusion élastique des neutrons, avec ses spécificités, donne accès, tout comme la diffusion X, à la fonction de corrélation $g(r)$, par l'intermédiaire de $S(K)$ (par exemple, par la relation (5.17)). Il est à noter que la diffusion *inélastique* (diffusion Brillouin), avec variation du nombre de phonons, donne des renseignements sur la loi de dispersion de ces derniers, principalement sur la vitesse du son au sein du milieu¹⁶.

5.2 Diffraction par un réseau. Conditions de von Laue et de Bragg

En raison de l'ordre géométrique à longue distance présent dans un cristal, il existe un très grand nombre N de diffuseurs qui vont pouvoir *coopérer* et produire des figures de diffusion extrêmement sélectives en direction, la longueur d'onde du rayonnement étant fixée. On parle alors traditionnellement de *diffraction* pour désigner cette sélectivité étant entendu que, au niveau de chaque diffuseur, il s'agit toujours de la simple diffusion élastique de la lumière au sens rappelé ci-dessus. C'est l'énormité de N allié à l'existence d'un ordre à portée infinie qui change qualitativement les phénomènes observés quand on passe d'un milieu désordonné (ou présentant un ordre à courte distance) à un cristal. En parlant schématiquement, on passe d'une réponse *floue* dans le premier cas à une réponse par tout ou rien dans le second.

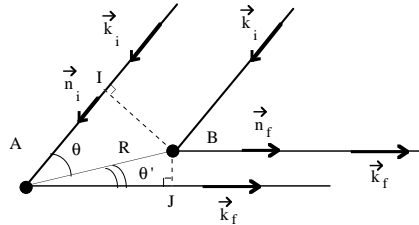


Figure 5.1: Déphasage pour deux diffuseurs séparés de la distance R .

Le lien entre sélectivité et effets coopératifs de diffuseurs ordonnés est facile à saisir. Soit \vec{k}_i le vecteur d'onde incident, \vec{k}_f le vecteur d'onde diffusé élastiquement. L'onde diffusée par l'atome situé au nœud \vec{R} d'un réseau de Bravais est caractérisé par un déphasage $\vec{q} \cdot \vec{R}$ où $\vec{q} = \vec{k}_f - \vec{k}_i$ est le transfert de moment. En effet (voir figure 5.1), la différence de chemin pour les ondes diffusées par A et B est $IA + AJ$, soit :

$$R \cos \theta + R \cos \theta' \equiv \vec{R} \cdot (-\vec{n}_i) + \vec{R} \cdot \vec{n}_f = \vec{R} \cdot (\vec{n}_f - \vec{n}_i) \quad (\vec{R} = \vec{AB}) . \quad (5.23)$$

Pour deux diffuseurs séparés par \vec{R} le déphasage est donc :

$$\frac{2\pi}{\lambda} \vec{R} \cdot (\vec{n}_f - \vec{n}_i) \equiv \vec{R} \cdot (\vec{k}_f - \vec{k}_i) \equiv \vec{R} \cdot \vec{q} \quad \left(\lambda = \frac{2\pi}{\|\vec{k}_i\|} = \frac{2\pi}{\|\vec{k}_f\|} \right) . \quad (5.24)$$

¹⁶Si $\omega(\vec{k})$ est la loi de dispersion des phonons acoustiques, la vitesse du son, usuellement notée c , est simplement reliée à $(\vec{\nabla}\omega)_{\vec{k}=0}$ (pour un milieu isotrope, $c = \|(\vec{\nabla}\omega)_{\vec{k}=0}\|$).

Il en résulte que l'amplitude de l'onde diffusée par un réseau de Bravais \mathcal{B} est de la forme¹⁷ :

$$\Re A_0 \sum_{\vec{R} \in \mathcal{B}} e^{-i(\omega t - \vec{q} \cdot \vec{r})} = \Re(A_0 e^{-i\omega t} S) , \quad \boxed{S \stackrel{\text{d\'ef}}{=} \sum_{\vec{R}} e^{i\vec{q} \cdot \vec{R}}} \quad (5.25)$$

La sommation sur \vec{R} est¹⁸ :

$$S = \sum_{\vec{R}} e^{i\vec{q} \cdot \vec{R}} = \sum_{\{n_i\} \in \mathbb{Z}^3} \prod_{i=1}^3 e^{in_i \vec{q} \cdot \vec{a}_i} = \prod_{i=1}^3 \sum_{n_i \in \mathbb{Z}} e^{in_i \vec{q} \cdot \vec{a}_i} = \prod_{i=1}^3 \left[2\pi \sum_{p_i \in \mathbb{Z}} \delta(\vec{q} \cdot \vec{a}_i - p_i 2\pi) \right] . \quad (5.27)$$

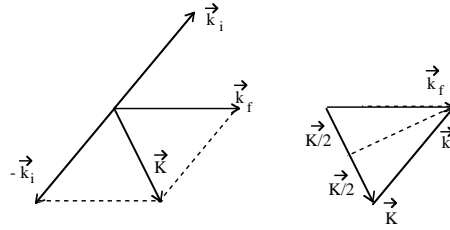


Figure 5.2: Illustration géométrique de la condition de von Laue.

L'apparition des fonctions de Dirac révèle la sélectivité de la diffusion coopérative d'une infinité dénombrable d'objets ordonnés dans l'espace. $\vec{q} = \vec{k}_f - \vec{k}_i$ étant un vecteur de l'espace réciproque¹⁹, il est naturel de le représenter (décomposer) sur les vecteurs primitifs \vec{b}_j du réseau réciproque :

$$\vec{q} = \sum_{j=1}^3 \mu_j \vec{b}_j , \quad (5.28)$$

où les μ_j sont des réels quelconques. Alors :

$$\vec{q} \cdot \vec{a}_i = \sum_{j=1}^3 \mu_j \vec{b}_j \cdot \vec{a}_i = 2\pi \mu_i , \quad (5.29)$$

de sorte que la somme S définie en (5.27) est :

$$S = \prod_{i=1}^3 \left[2\pi \sum_{p_i \in \mathbb{Z}} \delta(2\pi \mu_i - p_i 2\pi) \right] = \prod_{i=1}^3 \left[\sum_{p_i=-\infty}^{+\infty} \delta(\mu_i - p_i) \right] . \quad (5.30)$$

Ceci, compte tenu de la définition (5.28) des μ_j , montre que la somme S est nulle sauf si le transfert de moment \vec{q} est précisément un vecteur du réseau réciproque :

$$\boxed{\text{diffusion par le réseau} \iff \exists \vec{K} \in \tilde{\mathcal{B}} : \vec{K} = \vec{k}_f - \vec{k}_i} \quad (5.31)$$

La condition (5.31), dite *condition de von Laue*, a une traduction géométrique simple ; compte tenu du fait que

¹⁷ \Re désigne la partie réelle.

¹⁸On utilise trois fois le résultat :

$$\sum_{n=-\infty}^{+\infty} e^{inx} = 2\pi \sum_{p=-\infty}^{+\infty} \delta(x - p 2\pi) = \sum_{p=-\infty}^{+\infty} \delta\left(\frac{x}{2\pi} - p\right) \stackrel{\text{d\'ef}}{=} \text{III}(x) , \quad (5.26)$$

où $\text{III}(x)$ est le peigne de Dirac.

¹⁹Il s'agit bien de l'espace réciproque, pas du réseau réciproque $\tilde{\mathcal{B}}$, où celui-ci est plongé. Les composantes de \vec{q} sont pour l'instant des nombres quelconques, pas entiers.

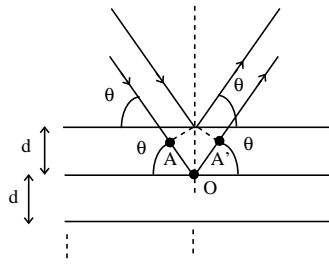


Figure 5.3: Illustration géométrique de la condition de Bragg.

les deux vecteurs d'onde \vec{k}_i et \vec{k}_f ont même module, $-\vec{k}_i$ et \vec{k}_f forment un losange dont \vec{K} est une diagonale ; ou encore : le vecteur formant la base d'un triangle isocèle construit sur \vec{k}_i et \vec{k}_f est un vecteur \vec{K} du réseau réciproque (voir figure 5.2).

Notons que la condition de von Laue ne fait aucune hypothèse particulière (autre qu'élastique) sur la nature de la diffusion de la lumière et résulte uniquement de l'existence d'un ordre à longue portée, matérialisé par le réseau de Bravais. Elle est équivalente à la condition dite de Bragg, qui repose au contraire sur une hypothèse supplémentaire, celle de la réflexion *spéculaire* sur les plans réticulaires (voir figure 5.3), chaque plan se comportant exactement comme un miroir. La différence de marche entre deux plans consécutifs est $AO + OA' = 2d \sin \theta$. L'interférence entre ces deux plans est donc constructive si :

$$\boxed{2d \sin \theta = \text{entier} \times \text{longueur d'onde}} \quad (5.32)$$

Pour un ensemble semi-infini de plans, la sélectivité en direction est absolue et il n'y aura réflexion que pour les angles θ satisfaisant :

$$\boxed{2d \sin \theta = n \lambda \quad (n \in \mathbb{N})} \quad (5.33)$$

Cette relation est appelée *condition de diffraction de Bragg*²⁰.

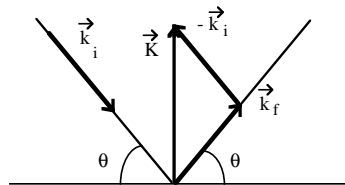


Figure 5.4: Equivalence entre les conditions de von Laue et de Bragg.

L'équivalence entre les deux conditions de von Laue et de Bragg se voit comme suit. En partant du schéma de Bragg (hypothèse d'une réflexion spéculaire sur un plan), on peut construire le vecteur $\vec{k}_f - \vec{k}_i$, qui est perpendiculaire au plan choisi, et a pour longueur (voir figure (5.4)) :

$$2 k_i \sin \theta = 2k_i \frac{n\lambda}{2d} = n \frac{2\pi}{d} . \quad (5.34)$$

Mais $(2\pi/d)$ est précisément la longueur du plus petit vecteur réciproque perpendiculaire à la famille de plans considérée, et, par ailleurs, $\vec{k}_f - \vec{k}_i \in \vec{\mathcal{B}}$ est la condition de von Laue – ce qui achève de montrer l'équivalence.

Que l'on raisonne d'une façon ou d'une autre, il existe donc une condition stricte pour que la diffraction se produise, que l'on peut traduire géométriquement par la construction dite d'Ewald. Étant donné le vecteur \vec{k}_i , on fixe son extrémité sur un nœud quelconque, O, du réseau réciproque et on trace une sphère centrée sur l'extrémité de ce vecteur, de rayon k_i ; par construction, tous les vecteurs \vec{k}_f allant du centre de la sphère à

²⁰Elle montre en particulier qu'il ne peut y avoir réflexion de Bragg que si $\lambda/(2d) < 1$.

n'importe quel point de sa surface satisfait la condition (5.31). Il en résulte que tous les points de $\tilde{\mathcal{B}}$ se trouvant sur la surface de la sphère donnent lieu à la diffraction (voir figure 5.5). En général, il n'y a aucune coïncidence, ce qui traduit le fait que pour \vec{k}_i quelconque (en module et en orientation par rapport au cristal), il n'y a pas de diffraction de Bragg. \vec{k}_i étant fixé, il convient donc d'orienter convenablement le cristal pour observer un pic de Bragg.

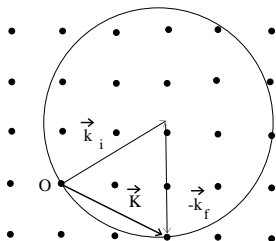


Figure 5.5: Construction d'Ewald.

Toutefois, plusieurs méthodes permettent d'obtenir des pics de diffraction sans prendre de précautions particulières :

- La méthode du cristal tournant. Maintenant, la source monochromatique étant fixe, on fait tourner le cristal lentement après l'avoir entouré d'un film sensible aux rayons X. Ainsi, la sphère d'Ewald tourne, balaie le réseau réciproque et rencontre donc un certain nombre de nœuds.
- La méthode des poudres (dite de Debye - Scherrer), qui utilise non un monocristal, mais une poudre polycristalline dont les grains, quoique petits (en fait, mésoscopiques), contiennent encore un nombre gigantesque de points du réseau direct, donnant une hypersélectivité, même avec les plus grandes résolutions spectrales.
- on utilise une source de rayons X polychromatique, comprise entre deux longueurs d'onde λ_1 et λ_2 . Alors, en traçant les deux sphères d'Ewald correspondantes, on voit que tous les nœuds du réseau situés *entre* les deux sphères donneront une diffusion sélective (voir figure 5.6).

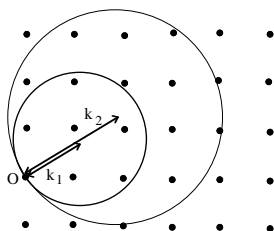


Figure 5.6: Utilisation d'une source polychromatique : tous les points de $\tilde{\mathcal{B}}$ situés entre les deux sphères donnent lieu à une diffraction de Bragg.

5.3 Facteurs de structure (géométrique et atomique)

Jusqu'à présent, seule la diffusion coopérative d'un ensemble ordonné d'objets ponctuels identiques a été considérée. Autrement dit, on a supposé que chaque maille primitive contient un objet à symétrie sphérique assez petit pour que sa dimension linéaire soit non-pertinente, autorisant à l'assimiler à un point. Par ailleurs, rien n'a été dit sur l'objet qui se tient en chaque nœud, mais il est clair que l'amplitude de l'onde diffusée est proportionnelle à un certain facteur f mesurant l'aptitude de cet objet à diffuser plus ou moins le rayonnement incident.

En pratique, les solides cristallisés sont construits le plus souvent sur des réseaux avec base, formée de n objets (toujours ponctuels pour l'instant) assemblés suivant une géométrie connue (par exemple, un cubique simple avec une molécule diatomique à chaque nœud, et orientée d'une façon connue). Dès lors, l'interférence constructive des mailles primitives équivalentes par translation est modulée en amplitude par des déphasages additionnels de l'onde sur le volume contenant les objets non-structurés constituant la base. Ainsi, pour une diffraction associée au vecteur du réseau réciproque \vec{K} , il apparaît en outre un déphasage $\vec{K} \cdot (\vec{d}_i - \vec{d}_j)$ pour tout couple de diffuseurs situés aux points définis par les vecteurs \vec{d}_i et \vec{d}_j , représentant leur position \vec{d} relativement à une maille primitive. La lumière diffusée par une maille primitive a pour amplitude la somme des amplitudes diffusées par chaque élément du motif (base) ; supposant toujours pour l'instant que tous les diffuseurs sont physiquement identiques, chaque terme $e^{i\vec{q} \cdot \vec{R}}$ de la somme S apparaissant en (5.27) est alors remplacé par :

$$\sum_{j=1}^n e^{i\vec{q} \cdot (\vec{R} + \vec{d}_j)} . \quad (5.35)$$

Il en résulte que S devient :

$$S \equiv \sum_{\vec{R}} e^{i\vec{q} \cdot \vec{R}} \rightsquigarrow S' = \sum_{\vec{R}} \sum_{j=1}^n e^{i\vec{q} \cdot (\vec{R} + \vec{d}_j)} . \quad (5.36)$$

Par définition, toutes les mailles primitives se valent et contiennent donc la même somme \sum_j . Celle-ci se factorise, d'où :

$$S' = \sum_{\vec{R}} \sum_{j=1}^n e^{i\vec{q} \cdot (\vec{R} + \vec{d}_j)} = \sum_{j=1}^n e^{i\vec{q} \cdot \vec{d}_j} \sum_{\vec{R}} e^{i\vec{q} \cdot \vec{R}} \equiv S_{\vec{q}} S , \quad (5.37)$$

de sorte que l'effet du motif se traduit par l'introduction d'un facteur global d'amplitude $S_{\vec{K}}$:

$$\boxed{S_{\vec{K}} = \sum_{j=1}^n e^{i\vec{K} \cdot \vec{d}_j}} \quad (5.38)$$

La quantité $S_{\vec{K}}$ est connue sous le nom de *facteur de diffusion géométrique*, puisqu'elle ne dépend que de la géométrie de la base associée à une maille primitive ; en général $S_{\vec{K}} \in \mathbb{C}$.

A titre d'exemple, soit un réseau cubique centré, dont le volume de la maille primitive est $v = a^3/2$. On peut évidemment raisonner avec son réseau réciproque $\tilde{\mathcal{B}}$, qui est CFC. Il est toutefois presque plus commode de considérer le réseau \mathcal{B} comme un cubique simple muni de la base diatomique : un atome à l'origine du cube, un autre au centre du cube, de coordonnées $(\frac{a}{2}, \frac{a}{2}, \frac{a}{2})$. Ceci étant fait, on voit que le facteur de forme est :

$$S_{\vec{K}} = 1 + e^{i\frac{a}{2}\vec{K} \cdot (\vec{i} + \vec{j} + \vec{k})} . \quad (5.39)$$

Dans cette optique, le réseau réciproque est un cubique simple de côté $2\pi/a$, et on a :

$$\forall \vec{K} \in \tilde{\mathcal{B}}_C \quad \vec{K} = \frac{2\pi}{a} (n_1 \vec{i} + n_2 \vec{j} + n_3 \vec{k}) , \quad (5.40)$$

d'où :

$$\vec{K} \cdot (\vec{i} + \vec{j} + \vec{k}) = \frac{2\pi}{a} (n_1 + n_2 + n_3) . \quad (5.41)$$

Finalement, d'après (5.39) :

$$S_{\vec{K}} = 1 + e^{i\pi (n_1 + n_2 + n_3)} = \begin{cases} 0 & \text{si } n_1 + n_2 + n_3 \text{ est impair} \\ 2 & \text{si } n_1 + n_2 + n_3 \text{ est pair} \end{cases} . \quad (5.42)$$

Ceci constitue une sorte de règle de sélection – d'origine purement géométrique – aussi appelée règle d'extinction. Ici, $S_{\vec{K}}$ est réel.

Bien évidemment, le spectre de diffraction ne dépend pas de la convention choisie pour le raisonnement : on obtient les mêmes résultats physiques en raisonnant directement avec le CC sans base. Dans cette optique,

le réseau réciproque est un CFC dont le côté²¹ de la maille conventionnelle est $\frac{4\pi}{a}$, et tout vecteur réciproque de $\tilde{\mathcal{B}}_{\text{CFC}}$ est de la forme :

$$\vec{K} = \frac{4\pi}{a} \left(m_1 \frac{\vec{i} + \vec{j}}{2} + m_2 \frac{\vec{j} + \vec{k}}{2} + m_3 \frac{\vec{k} + \vec{i}}{2} \right) = \frac{2\pi}{a} \left[(m_1 + m_3)\vec{i} + (m_1 + m_2)\vec{j} + (m_2 + m_3)\vec{k} \right]. \quad (5.43)$$

La comparaison des deux expressions de \vec{K} , (5.40) et (5.43), montre que l'on a :

$$n_1 = m_1 + m_3, \quad n_2 = m_1 + m_2, \quad n_3 = m_2 + m_3, \quad (5.44)$$

de sorte que :

$$n_1 + n_2 + n_3 = 2m_1 + 2m_2 + 2m_3. \quad (5.45)$$

Cette somme est bien un nombre pair, comme le prescrit la règle de sélection additionnelle (5.42) obtenue quand on raisonne avec le réseau cubique simple muni de la base diatomique restituant le CC. Ceci montre bien que le choix d'une maille conventionnelle (par opposition à une maille primitive) impose la prise en compte du facteur géométrique associé à la base ainsi introduite *de facto*.

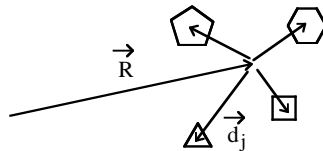


Figure 5.7: Notations utilisées dans (5.47).

Enfin, le plus souvent, les différents objets constituant la base ne sont pas tous identiques physiquement. Chaque onde réémise par un diffuseur contient un facteur d'amplitude spécifique à ce diffuseur, noté f et appelé facteur de forme atomique, qui dépend notamment du transfert de moment $\vec{q} = \vec{K}$. Si tous les objets de la base sont les mêmes, il s'agit d'un simple facteur commun à toutes les amplitudes partielles, et on peut l'absorber dans d'autres constantes multiplicatives. En revanche, si les constituants de la base sont distincts physiquement, le facteur de forme géométrique doit en tenir compte (voir figure 5.7) ; dans un tel cas, on définit un facteur de structure englobant les deux aspects²² :

$$S_{\vec{K}} = \sum_{j=1}^n f_j(\vec{K}) e^{i\vec{K} \cdot \vec{d}_j} \quad (5.47)$$

La dépendance de f_j vis-à-vis de \vec{K} provient du déphasage calculé sur la densité $\rho(\vec{r})$ des particules qui se couplent à la sonde à l'intérieur d'un diffuseur. Il s'agit donc essentiellement d'une transformée de Fourier²³ :

$$f_j(\vec{K}) = \int_{\mathbb{R}^3} d^3r e^{i\vec{K} \cdot \vec{r}} \rho(\vec{r}). \quad (5.48)$$

$f_j(\vec{K})$ est appelé *facteur de forme atomique* ; sa dépendance angulaire vis-à-vis de la direction de \vec{K} donne une idée précise de la forme du diffuseur (anisotropie), sa dépendance par rapport au module de \vec{K} donne

²¹La maille élémentaire du cubique centré a pour volume $v = \frac{a^3}{2}$, celui de la maille primitive du réseau réciproque CFC est $\tilde{v} = \frac{(2\pi)^3}{v} = \frac{16\pi^3}{a^3}$. Comme la maille conventionnelle du CFC possède 4 atomes, son volume est $\frac{1}{4}\tilde{a}^3$ où \tilde{a} est le côté de la maille conventionnelle du CFC ; de $\frac{1}{4}\tilde{a}^3 = \frac{16\pi^3}{a^3}$, on déduit $\tilde{a} = \frac{4\pi}{a}$: la maille conventionnelle du réseau réciproque CFC est donc bien un cube de côté $\frac{4\pi}{a}$.

²²Il s'agit de la définition usuelle. Ceci étant, il serait alors plus logique, dans le cas où les atomes du motif sont identiques, de définir $S_{\vec{K}}$ (voir éq. (5.38)) comme :

$$S_{\vec{K}} = f(\vec{K}) \sum_{j=1}^n e^{i\vec{K} \cdot \vec{d}_j} \quad (5.46)$$

où $f(\vec{K})$ est le facteur commun à tous les éléments du motif. Les définitions utilisées ici sont strictement celles de Ashcroft et Mermin [6].

²³Avec cette définition, le facteur de forme atomique f_j est sans dimension.

des informations sur sa dimension (son extension spatiale). Avec les rayons X, $f_j(\vec{K})$ fournit la photo de la distribution électronique au sein de l'atome.

Pour des objets de dimensions atomiques et pour des rayons X, la quantité $\vec{K} \cdot \vec{r}$ est d'ordre 1 et la dépendance en \vec{K} ne doit être *a priori* négligée. Au contraire, pour la diffusion des neutrons sur la matière ordinaire, on pourra – en ce qui concerne l'interaction de Fermi – négliger cette dépendance de f_j : l'interaction est en effet alors essentiellement de contact, examinée à l'échelle de longueur de la matière condensée usuelle et seul $f_j(\vec{K} = 0)$ est pertinent.

◆ *Remarque*

La diffraction de Bragg résulte de la diffusion (élastique) coopérative de diffuseurs répartis dans l'espace suivant un ordre parfait, un fait dont rend compte le calcul du déphasage spatial pour deux ondes planes $e^{i\vec{k}_i \cdot \vec{r}}$ et $e^{i\vec{k}_f \cdot \vec{r}}$. Ce calcul est manifestement d'inspiration classique.

Il est possible de donner une version quantique du même phénomène. Essentiellement, l'ingrédient que la description classique laisse sous le tapis est la *probabilité* pour le photon (ou le neutron) d'être diffusé de \vec{k}_i en \vec{k}_f . Cette probabilité peut se calculer par la théorie des perturbations et plus particulièrement par la règle d'or de Fermi ; ceci étant obtenu, on en déduit la section efficace de diffusion (finalement, c'est la quantité mesurée), qui fait ressortir la condition de von Laue. La présentation ci-dessous donne les idées principales pour la diffusion des neutrons.

D'après la règle d'or, la probabilité (par unité de temps) de la diffusion $\vec{k}_i \rightarrow \vec{k}_f$ est :

$$\mathcal{P}_{\vec{k}_i \rightarrow \vec{k}_f} = \frac{2\pi}{\hbar} \left| \int_{L^3} d^3r \psi_{\vec{k}_f}^*(\vec{r}) V(\vec{r}) \psi_{\vec{k}_i}(\vec{r}) \right|^2 \rho_{\vec{k}_f}^-(E_f) , \quad (5.49)$$

où $\rho_{\vec{k}}^-(E)$ est la densité d'états en énergie telle que :

$$\rho_{\vec{k}}^-(E) dE = \left(\frac{L}{2\pi} \right)^3 d^3k , \quad d^3k = k^2 dk d\Omega . \quad (5.50)$$

Pour des neutrons (masse M , vecteur d'onde \vec{k}) et avec des conditions cycliques dans un cube de côté L , (5.50) donne :

$$\rho_{\vec{k}}^-(E) dE = \left(\frac{L}{2\pi} \right)^3 \frac{Mk}{\hbar^2} d\Omega . \quad (5.51)$$

Dans (5.49), les fonctions d'onde initiale et finale d'un neutron seront ²⁴ les ondes planes $L^{-3/2} e^{i\vec{k}_s \cdot \vec{r}}$ avec $s = i, f$; $V(\vec{r})$ est l'interaction entre un neutron et un seul diffuseur : c'est par exemple le pseudo-potentiel de Fermi donné en (5.20).

Une fois connue la probabilité de transition par unité de temps, on peut écrire la section efficace élémentaire $d\sigma$:

$$d\sigma = \frac{\mathcal{P}_{\vec{k}_i \rightarrow \vec{k}_f}}{\text{courant incident}} . \quad (5.52)$$

Pour un neutron de vitesse $v = \hbar k_i / M$ dans le volume L^3 , le courant est $L^{-3} (\hbar k_i / M)$. En combinant (5.49), (5.51) et (5.52), on trouve :

$$\frac{d\sigma}{d\Omega} = \left(\frac{M}{2\pi\hbar^2} \right)^2 \left| \int_{L^3} d^3r e^{i\vec{q} \cdot \vec{r}} V(\vec{r}) \right|^2 , \quad \vec{q} = \vec{k}_i - \vec{k}_f . \quad (5.53)$$

Si maintenant on fait le choix selon Fermi, (5.20), pour un noyau situé au site \vec{R} , on trouve le résultat très simple :

$$\frac{d\sigma}{d\Omega} = \left| b e^{i\vec{q} \cdot \vec{R}} \right|^2 = b^2 . \quad (5.54)$$

La longueur de diffusion de Fermi donne directement la section efficace, et joue exactement le même rôle que le rayon d'une sphère dure dans une collision classique. La section efficace totale est $\sigma = 4\pi b^2$.

²⁴En toute rigueur, l'état final est un état de diffusion du potentiel $V(\vec{r})$.

Pour des diffuseurs tous identiques ordonnés suivant le réseau \mathcal{B} , $V(\vec{r})$ est :

$$V(\vec{r}) = \frac{2\pi\hbar^2}{M} b \sum_{\vec{R} \in \mathcal{B}} \delta(\vec{r} - \vec{R}) . \quad (5.55)$$

La section efficace est alors :

$$\frac{d\sigma}{d\Omega} = \left| b \sum_{\vec{R} \in \mathcal{B}} e^{i\vec{q} \cdot \vec{R}} \right|^2 \equiv b^2 |S|^2 . \quad (5.56)$$

Pour un réseau avec une base formée de n noyaux de longueurs de diffusion b_j , on a :

$$\frac{d\sigma}{d\Omega} = \left| \sum_{\vec{R} \in \mathcal{B}} \sum_{j=1}^n b_j e^{i\vec{q} \cdot (\vec{R} + \vec{d}_j)} \right|^2 = \left| \sum_{\vec{R} \in \mathcal{B}} e^{i\vec{q} \cdot \vec{R}} \right|^2 \left| \sum_{j=1}^n b_j e^{i\vec{q} \cdot \vec{d}_j} \right|^2 . \quad (5.57)$$

Le facteur de diffusion géométrique dû au motif est maintenant :

$$\mathcal{S}_{\vec{q}} = \sum_{j=1}^n b_j e^{i\vec{q} \cdot \vec{d}_j} , \quad (5.58)$$

et a ici la dimension d'une longueur ; les b_j jouent dans (5.58) le même rôle que les f_j dans (5.47) et n'apportent pas de dépendance en \vec{q} puisque le potentiel de Fermi est de contact : pour cette interaction, ce qui tient lieu de facteur de forme "atomique" est indépendant du transfert de moment. \blacklozenge

Chapitre 6

Propriétés générales du mouvement d'un électron dans un potentiel périodique

6.1 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, il existe dans le solide constitué deux types de particules. D'une part les électrons, de faible inertie (masse m), qui peuvent être liés à un centre (noyau ou cœur) – c'est le cas des cristaux moléculaires, ioniques et covalents –, ou au contraire presque libres – dans le cas des métaux (toutes les situations intermédiaires étant *a priori* possibles) ; d'autre part les cœurs dont la masse est d'ordre une masse nucléaire M . Cette énorme différence de masses est la clé de voûte de la toute première approximation, dite approximation de Born et Oppenheimer, qui repose sur un développement plus ou moins perturbatif par rapport au petit paramètre $m/M \ll 1$ et permet de justifier la possibilité d'une étude en deux étapes ; schématiquement, les électrons se déplacent très rapidement (qu'ils soient liés ou quasi-libres) et, de ce fait, voient des noyaux quasi-immobiles ; à l'inverse, les noyaux ne voient que le mouvement *moyenné* des électrons. En définitive, le système contient deux types de degrés de liberté : des degrés rapides et des degrés lents.

Ainsi, dans un premier temps, on détermine les états quantiques des électrons dans le champ des noyaux *fixes* (énergies et fonctions propres) ; ensuite, on détermine le mouvement des noyaux, l'énergie propre électronique trouvée précédemment jouant le rôle d'une énergie potentielle. Historiquement, cette approximation, ou l'une de ses variantes, a été introduite lors des premières applications de la Mécanique Quantique aux molécules – un solide n'est jamais qu'une gigantesque molécule ; sa justification détaillée, quelque peu laborieuse, met en évidence ses limites et les situations où un traitement plus subtil s'impose (lorsque, notamment, deux énergies électroniques sont dégénérées, ou quasi-dégénérées). Dans toute la suite, on suppose remplies les conditions de validité de l'approximation de Born et Oppenheimer.

La séparation entre noyaux (ou cœurs) et électrons étant radicalisée, 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).

A l'issue de ce dégrossissage, et en se plaçant dans la première étape de l'approximation de Born et Oppenheimer, on reste *a priori* avec des ions et les électrons de valence, ces derniers étant soumis au potentiel plus ou moins fort des premiers³. En plus du potentiel périodique du réseau, ces 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 de quelques modèles très 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 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.

Landau a été le premier à proposer une théorie *phénoménologique* formalisant ces phénomènes, et introduisant fondamentalement la notion de *quasi-particule*, particule *nue* enveloppée d'un nuage d'autres particules (Mattuck [9] emploie l'image d'un cow-boy traversant le désert au galop sur son cheval) ; par la suite, Nozières et Luttinger ont donné une justification théorique (ardue) de la théorie de Landau, tout en délimitant son champ d'application, permettant ainsi de préciser les situations physiques où une autre approche est nécessaire.

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 :

$$H = \frac{\vec{p}^2}{2m} + V(\vec{r}) , \quad V(\vec{r} + \vec{R}) = V(\vec{r}) \quad \forall \vec{R} \in \mathcal{B} \tag{6.1}$$

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 $K_\alpha, K_\beta, etc.$

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 8.

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.

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 peu mobiles, au sens où leur énergie cinétique est très petite devant leur 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.

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 chapitre 7), 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 $\vec{R} \in \mathcal{B}$.

6.2 Théorème de Bloch

La symétrie de translation caractéristique de V entraîne que le Hamiltonien (6.1) 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 classique est le suivant⁷ :

Pour un Hamiltonien périodique du type (6.1), 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 :

$$\boxed{\forall \psi(\vec{r}), H\psi(\vec{r}) = \varepsilon \psi(\vec{r}), \quad \exists u(\vec{r}), \psi(\vec{r}) = e^{i\vec{k}\cdot\vec{r}} u(\vec{r}), \quad u(\vec{r}) = u(\vec{r} + \vec{R}) \quad \forall \vec{R} \in \mathcal{B}} \quad (6.2)$$

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 ECOOC. 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 (6.2) plus précisément (notations traditionnelles) :

$$H\psi_{n\vec{k}}(\vec{r}) = \varepsilon_n(\vec{k}) \psi_{n\vec{k}}(\vec{r}), \quad \psi_{n\vec{k}}(\vec{r}) = e^{i\vec{k}\cdot\vec{r}} u_{n\vec{k}}(\vec{r}), \quad u_{n\vec{k}}(\vec{r}) = u_{n\vec{k}}(\vec{r} + \vec{R}). \quad (6.3)$$

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

$$\psi_{n\vec{k}}(\vec{r} + \vec{R}) = e^{i\vec{k}\cdot(\vec{r} + \vec{R})} u_{n\vec{k}}(\vec{r} + \vec{R}) = e^{i\vec{k}\cdot\vec{R}} e^{i\vec{k}\cdot\vec{r}} u_{n\vec{k}}(\vec{r}), \quad (6.4)$$

soit finalement :

$$\boxed{\psi_{n\vec{k}}(\vec{r} + \vec{R}) = e^{i\vec{k}\cdot\vec{R}} \psi_{n\vec{k}}(\vec{r})} \quad (6.5)$$

On utilisera selon les cas la forme (6.3) ou la forme (6.5).

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 (6.3) et (6.5) 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 :

$$\boxed{|\psi_{n\vec{k}}(\vec{r} + \vec{R})|^2 = |\psi_{n\vec{k}}(\vec{r})|^2} \quad (6.6)$$

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 (6.5).

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 !), (6.2) 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.

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)

1^{ère} démonstration

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

$$\boxed{\forall \phi(\vec{r}), \mathcal{T}(\vec{R}) \phi(\vec{r}) = \phi(\vec{r} - \vec{R})} \tag{6.8}$$

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 (6.1) :

$$\boxed{\forall \vec{R} \in \mathcal{B} : [H, \mathcal{T}(\vec{R})] = 0} \tag{6.9}$$

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, (6.9) 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 :

$$H(\mathcal{T}\psi) = \mathcal{T}H\psi = \mathcal{T}\varepsilon\psi = \varepsilon(\mathcal{T}\psi) ; \tag{6.10}$$

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 intuitive :

$$\mathcal{T}(\vec{R})\mathcal{T}(\vec{R}') = \mathcal{T}(\vec{R} + \vec{R}') . \tag{6.11}$$

L'inverse est défini comme :

$$\mathcal{T}^{-1}(\vec{R}) = \mathcal{T}(-\vec{R}) . \tag{6.12}$$

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

$$\mathcal{T}(\vec{R})\mathcal{T}(\vec{R}') = \mathcal{T}(\vec{R} + \vec{R}') = \mathcal{T}(\vec{R}' + \vec{R}) = \mathcal{T}(\vec{R}')\mathcal{T}(\vec{R}) \iff [\mathcal{T}(\vec{R}), \mathcal{T}(\vec{R}')] = 0 \tag{6.13}$$

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 :

$$\mathcal{T}^\dagger(\vec{R}) = \mathcal{T}^{-1}(\vec{R}) = \mathcal{T}(-\vec{R}) . \tag{6.14}$$

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 :

$$H\psi(\vec{r}) = \varepsilon\psi(\vec{r}) , \tag{6.15}$$

⁹La définition (6.8) 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 :

$$\mathcal{T}(\vec{R})\phi(\vec{r}) = \phi(\vec{r} + \vec{R}) \quad \forall \phi(\vec{r}) . \tag{6.7}$$

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.

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 $\vec{p} = -i\hbar\nabla$ 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 (6.8).

¹³Pour la simplicité de l'écriture, on omet les indices $(n\vec{k})$ apparaissant dans les énoncés (6.3) et (6.5) du théorème de Bloch.

$$\mathcal{T}(\vec{R})\psi(\vec{r}) = \tau(\vec{R})\psi(\vec{r}) \quad \forall \vec{R} \in \mathcal{B}, \quad (6.16)$$

où, tout naturellement, la valeur propre τ de $\mathcal{T}(\vec{R})$ dépend de \vec{R} . L'équation (6.15) 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 (6.16) et (6.11), on a :

$$\mathcal{T}(\vec{R})\mathcal{T}(\vec{R}')\psi(\vec{r}) = \tau(\vec{R})\tau(\vec{R}')\psi(\vec{r}) \quad (6.17)$$

et :

$$\mathcal{T}(\vec{R} + \vec{R}')\psi(\vec{r}) = \tau(\vec{R} + \vec{R}')\psi(\vec{r}). \quad (6.18)$$

Les équations (6.17) et (6.18) sont vraies $\forall \psi$ également propre de H , donc pour une base complète. Il en résulte :

$$\tau(\vec{R} + \vec{R}') = \tau(\vec{R})\tau(\vec{R}'). \quad (6.19)$$

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 :

$$\tau(\vec{R}) = e^{-i\Phi(\vec{R})}, \quad (6.20)$$

où Φ est une fonction à valeurs réelles ; (6.19) s'écrit alors :

$$\Phi(\vec{R} + \vec{R}') = \Phi(\vec{R}) + \Phi(\vec{R}'); \quad (6.21)$$

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 :

$$\boxed{\tau(\vec{R}) = e^{-i\vec{k} \cdot \vec{R}}} \quad (6.22)$$

Compte tenu de ce résultat, la relation (6.16) s'écrit maintenant, $\forall \psi$ propre de H et des \mathcal{T} :

$$\mathcal{T}(\vec{R})\psi(\vec{r}) = e^{-i\vec{k} \cdot \vec{R}}\psi(\vec{r}) \iff \psi(\vec{r} - \vec{R}) = e^{-i\vec{k} \cdot \vec{R}}\psi(\vec{r}). \quad (6.23)$$

Ceci n'est autre que la formulation (6.5) du théorème de Bloch, avec le vecteur $-\vec{R}$.

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

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

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.

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 :

$$f(\vec{r} + N_i \vec{a}_i) = f(\vec{r}) \quad \forall \text{ fonction } f \quad (i = 1, 2, 3; N_i \gg 1) . \tag{6.25}$$

Ceci étant admis, on voit que le vecteur \vec{k} apparaissant dans (6.5) est maintenant quantifié à toute petite échelle. En effet, en prenant pour f une onde plane, on a :

$$e^{i\vec{k} \cdot (\vec{r} + N_i \vec{a}_i)} = e^{i\vec{k} \cdot \vec{r}} , \tag{6.26}$$

soit :

$$e^{iN_i \vec{k} \cdot \vec{a}_i} = 1 . \tag{6.27}$$

Cette relation contraint les valeurs de \vec{k} qui, étant un vecteur d'onde, se développe tout naturellement 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 $\vec{a}_i \cdot \vec{b}_j = 2\pi \delta_{ij}$; (6.27) montre que les composantes κ_j satisfont $e^{2\pi i N_j \kappa_j} = 1$, soit :

$$\kappa_j = \text{entier relatif} \times \frac{1}{N_j} . \tag{6.28}$$

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

$$\boxed{\vec{k} = \sum_{j=1}^3 \frac{m_j}{N_j} \vec{b}_j \quad (m_j \in \mathbb{Z})} \tag{6.29}$$

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 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 $\delta^3 k$:

$$\delta^3 k = \frac{1}{N_1 N_2 N_3} |\vec{b}_1 \cdot (\vec{b}_2 \times \vec{b}_3)| \equiv \frac{1}{N} |(\vec{b}_1, \vec{b}_2, \vec{b}_3)| = \frac{1}{N} \frac{(2\pi)^3}{v} , \tag{6.30}$$

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

$$\delta^3 k = \frac{(2\pi)^3}{V} . \tag{6.31}$$

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* :

$$\psi(\vec{r}) = \sum_{\vec{q}} \phi(\vec{q}) e^{i\vec{q} \cdot \vec{r}} , \tag{6.32}$$

¹⁶Noter que la cyclicité impose aussi :

$$\mathcal{T}(\vec{R} = N_i \vec{a}_i) = \mathbf{1} . \tag{6.24}$$

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 $\vec{k} + \vec{K}$ ($\vec{K} \in \vec{\mathcal{B}}$) 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 $\vec{\mathcal{B}}$, soit en se bornant (par exemple) à $1 \leq m_j \leq N_j$.

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

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

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 :

$$\vec{q} = \sum_{j=1}^3 \frac{m_j}{N_j} \vec{b}_j \quad (m_j \in \mathbb{Z}) \iff e^{i N_i \vec{q} \cdot \vec{a}_i} = 1 . \quad (6.33)$$

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 (6.32) est un précurseur de l'intégrale (c'est presque une intégrale²⁰). Il n'y a pas de doute que l'expression (6.32) fournit une fonction satisfaisant les conditions cycliques²¹ ; avec (6.33), on peut écrire en effet :

$$\psi(\vec{r} + N_i \vec{a}_i) = \sum_{\vec{q}} \phi(\vec{q}) e^{i \vec{q} \cdot (\vec{r} + N_i \vec{a}_i)} = \sum_{\vec{q}} \phi(\vec{q}) e^{i \vec{q} \cdot \vec{r}} = \psi(\vec{r}) . \quad (6.34)$$

L'objectif final est de trouver $\psi(\vec{r})$: selon (6.32), 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 :

$$\boxed{V(\vec{r}) = \sum_{\vec{K} \in \tilde{\mathcal{B}}} \mathcal{V}(\vec{K}) e^{-i \vec{K} \cdot \vec{r}}} \quad (6.35)$$

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 :

$$\mathcal{V}(\vec{K}) = \frac{1}{v} \int_{\text{maille élémentaire}} V(\vec{r}) e^{i \vec{K} \cdot \vec{r}} d^3 r , \quad (6.36)$$

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

$$\int_{\text{maille élémentaire}} V(\vec{r}) e^{i \vec{K} \cdot \vec{r}} d^3 r = \sum_{\vec{K}' \in \tilde{\mathcal{B}}} \mathcal{V}(\vec{K}') \int_{\text{maille élémentaire}} e^{i(\vec{K} - \vec{K}') \cdot \vec{r}} d^3 r . \quad (6.37)$$

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 (6.37) est donc :

$$\int_{\text{maille élémentaire}} e^{2\pi i \sum_j m_j x_j} d^3 r \quad (6.38)$$

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 :

$$\int_{\text{maille élémentaire}} e^{i(\vec{K} - \vec{K}') \cdot \vec{r}} d^3 r = v \delta_{\vec{K} \vec{K}'}, \quad (6.39)$$

où $\delta_{\vec{K} \vec{K}'}$ est le symbole de Kronecker. Transposées aux variables espace - vecteur d'onde, les relations (6.35), (6.36) et (6.39) ne sont rien d'autre que ($n, n' \in \mathbb{Z}$) :

$$f(t) = \sum_{n \in \mathbb{Z}} f_n e^{-in\omega t} , \quad f_n = \frac{\omega}{2\pi} \int_{\text{période}} f_n e^{in\omega t} dt , \quad \int_{\text{période}} e^{i(n-n')\omega t} dt = 2\pi \delta_{nn'} , \quad (6.40)$$

pour le couple énergie - pulsation dans le cas d'une fonction de période $T = \frac{2\pi}{\omega}$. Notons qu'avec la définition (6.36), 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 :

²⁰ On peut même la considérer comme une somme de Darboux, à un facteur du genre $\delta^3 q$ près.

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

- On peut (presque) toujours²² choisir

$$\boxed{\mathcal{V}(\vec{K} = 0) = 0} \tag{6.41}$$

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 $\int d^3r V(\vec{r}) \sin(\vec{K} \cdot \vec{r})$, possédant un intégrand impair, est nulle. En pareil cas :

$$\boxed{\mathcal{V}(-\vec{K}) = \mathcal{V}(\vec{K}) = \mathcal{V}(\vec{K})^*} \tag{6.42}$$

En reportant les développements (6.32) et (6.35) dans l'équation aux valeurs propres écrite en représentation- q :

$$\left[\frac{\vec{p}^2}{2m} + V(\vec{r}) \right] \psi(\vec{r}) = \varepsilon \psi(\vec{r}) , \tag{6.43}$$

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

$$\frac{\vec{p}^2}{2m} \sum_{\vec{q}} \phi(\vec{q}) e^{i\vec{q} \cdot \vec{r}} = -\frac{\hbar^2}{2m} \sum_{\vec{q}} (i\vec{q})^2 \phi(\vec{q}) e^{i\vec{q} \cdot \vec{r}} , \tag{6.44}$$

et :

$$V(\vec{r}) \sum_{\vec{q}} \phi(\vec{q}) e^{i\vec{q} \cdot \vec{r}} = \sum_{\vec{q}} \sum_{\vec{K}} \mathcal{V}(\vec{K}) \phi(\vec{q}) e^{i\vec{q} \cdot \vec{r}} e^{-i\vec{K} \cdot \vec{r}} = \sum_{\vec{q}'} \sum_{\vec{K} \in \vec{\mathcal{B}}} \mathcal{V}(\vec{K}) \phi(\vec{q}' + \vec{K}) e^{i\vec{q}' \cdot \vec{r}} . \tag{6.45}$$

En identifiant maintenant dans les deux membres de l'équation aux valeurs propres (6.43) le coefficient de $e^{i\vec{q} \cdot \vec{r}}$, on obtient :

$$\forall \vec{q} : \frac{\hbar^2 \vec{q}^2}{2m} \phi(\vec{q}) + \sum_{\vec{K} \in \vec{\mathcal{B}}} \mathcal{V}(\vec{K}) \phi(\vec{q} + \vec{K}) = \varepsilon \phi(\vec{q}) \iff [\varepsilon - \varepsilon^{(0)}(\vec{q})] \phi(\vec{q}) = \sum_{\vec{K} \in \vec{\mathcal{B}}} \mathcal{V}(\vec{K}) \phi(\vec{q} + \vec{K}) , \tag{6.46}$$

où on a posé :

$$\boxed{\varepsilon^{(0)}(\vec{q}) \stackrel{\text{dét}}{=} \frac{\hbar^2 \vec{q}^2}{2m}} \tag{6.47}$$

Par inspection, on voit que le système linéaire²³ (6.46) couple exclusivement entre elles les composantes $\phi(\vec{q})$ ne différant que d'un vecteur \vec{K} appartenant à $\vec{\mathcal{B}}$. Autrement dit, \vec{q} étant confiné dans une maille primitive de $\vec{\mathcal{B}}$, seuls les vecteurs de la forme $\vec{q} + \vec{K}$, \vec{K} quelconque²⁴ dans $\vec{\mathcal{B}}$, apparaissent dans la sommation de l'expression (6.32). 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 :

$$\psi_{\vec{k}}(\vec{r}) = \sum_{\vec{K} \in \vec{\mathcal{B}}} \phi(\vec{k} + \vec{K}) e^{i(\vec{k} + \vec{K}) \cdot \vec{r}} = e^{i\vec{k} \cdot \vec{r}} \sum_{\vec{K} \in \vec{\mathcal{B}}} \phi(\vec{k} + \vec{K}) e^{i\vec{K} \cdot \vec{r}} . \tag{6.48}$$

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 $\vec{\mathcal{B}}$. \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.

²²Le *presque* est requis puisqu'il existe des cas exceptionnels où cette convention ne peut être utilisée. Par exemple, pour le modèle 1D de Kronig - Penney, défini par un potentiel en peigne de Dirac, $V(x) = g \sum_{n \in \mathbb{Z}} \delta(x - na)$, on a $\int_{\text{maille}} V(x) dx = g$, et on ne peut visiblement pas prendre $g = 0 \dots$

²³Il 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.

²⁴à l'exception de $\vec{K} = 0$ en vertu de (6.41).

Maintenant, la somme située le plus à droite dans (6.48) 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 (6.2) du théorème de Bloch ; (6.48) peut ainsi se récrire :

$$\psi_{\vec{k}}(\vec{r}) = e^{i\vec{k} \cdot \vec{r}} u_{\vec{k}}(\vec{r}) \quad , \quad u_{\vec{k}}(\vec{r}) = \sum_{\vec{K} \in \tilde{\mathcal{B}}} \phi(\vec{k} + \vec{K}) e^{i\vec{K} \cdot \vec{r}} = u_{\vec{k}}(\vec{r} + \vec{R}) \quad \forall \vec{R} \in \mathcal{B} \quad ; \quad (6.49)$$

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

$$u_{\vec{k} + \vec{K}_0}(\vec{r}) = \sum_{\vec{K} \in \tilde{\mathcal{B}}} \phi(\vec{k} + \vec{K}_0 + \vec{K}) e^{i\vec{K} \cdot \vec{r}} = \sum_{\vec{K}' \in \tilde{\mathcal{B}}} \phi(\vec{k} + \vec{K}') e^{i(\vec{K}' - \vec{K}_0) \cdot \vec{r}} = e^{-i\vec{K}_0 \cdot \vec{r}} u_{\vec{k}}(\vec{r}) \quad \forall \vec{K}_0 \in \tilde{\mathcal{B}} \quad . \quad (6.50)$$

6.3 Premières conséquences du théorème de Bloch

6.3.1 Première zone de Brillouin

Quand on adopte les conditions cycliques de Born - von Karman, les valeurs de \vec{k} sont de la forme (6.33) :

$$\boxed{\vec{k} = \sum_{j=1}^3 \frac{n_j}{N_j} \vec{b}_j \quad (n_j \text{ entier})} \quad (6.51)$$

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 physiquement distinctes 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 :

$$n_j = p_j N_j + r_j \quad (r_j < N_j) \quad . \quad (6.52)$$

Reporté dans (6.51), ceci montre que \vec{k} est de la forme :

$$\vec{k} = \sum_{j=1}^3 \frac{p_j N_j + r_j}{N_j} \vec{b}_j = \sum_{j=1}^3 \left(p_j + \frac{r_j}{N_j} \right) \vec{b}_j \equiv \vec{K} + \vec{k}_1 \quad , \quad (6.53)$$

\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 :

$$\psi_{\vec{k}}(\vec{r} + \vec{R}) = e^{i\vec{k} \cdot \vec{R}} \psi_{\vec{k}}(\vec{r}) = e^{i(\vec{k}_1 + \vec{K}) \cdot \vec{R}} \psi_{\vec{k}}(\vec{R}) = e^{i\vec{k}_1 \cdot \vec{R}} \psi_{\vec{k}}(\vec{r}) \quad . \quad (6.54)$$

Par ailleurs, on a toujours :

$$\psi_{\vec{k}_1}(\vec{r} + \vec{R}) = e^{i\vec{k}_1 \cdot \vec{R}} \psi_{\vec{k}_1}(\vec{r}) \quad , \quad (6.55)$$

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 (6.48) 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 :

$$\psi_{\vec{k} + \vec{K}_0}(\vec{r}) = \sum_{\vec{K} \in \tilde{\mathcal{B}}} \phi(\vec{k} + \vec{K}_0 + \vec{K}) e^{i(\vec{k} + \vec{K}_0 + \vec{K}) \cdot \vec{r}} = \sum_{\vec{K}' \in \tilde{\mathcal{B}}} \phi(\vec{k} + \vec{K}') e^{i(\vec{k} + \vec{K}') \cdot \vec{r}} \equiv \psi_{\vec{k}}(\vec{r}) \quad . \quad (6.56)$$

Ceci se voit encore comme suit ; en vertu de (6.50), on a encore :

$$\psi_{\vec{k} + \vec{K}_0}(\vec{r}) = e^{i(\vec{k} + \vec{K}_0) \cdot \vec{r}} u_{\vec{k} + \vec{K}_0}(\vec{r}) = e^{i(\vec{k} + \vec{K}_0) \cdot \vec{r}} e^{-i\vec{K}_0 \cdot \vec{r}} u_{\vec{k}}(\vec{r}) = e^{i\vec{k} \cdot \vec{r}} u_{\vec{k}}(\vec{r}) \equiv \psi_{\vec{k}}(\vec{r}) \quad . \quad (6.57)$$

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 fonctions²⁵. Cette propriété de périodicité de $\psi_{\vec{k}}$ se reporte évidemment sur la valeur propre associée $\varepsilon(\vec{k})$. On retiendra donc :

$$\boxed{\psi_{\vec{k} + \vec{K}}(\vec{r}) = \psi_{\vec{k}}(\vec{r}) \quad , \quad \varepsilon(\vec{k} + \vec{K}) = \varepsilon(\vec{k}) \quad \forall \vec{K} \in \tilde{\mathcal{B}}} \quad (6.58)$$

²⁵Tout 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.

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.

6.3.2 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 :

$$\psi_{\vec{k}}(\vec{r}) = e^{i\vec{k}\cdot\vec{r}} u_{\vec{k}}(\vec{r}) , \tag{6.59}$$

avec :

$$u_{\vec{k}}(\vec{r} + \vec{R}) = u_{\vec{k}}(\vec{r}) . \tag{6.60}$$

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 :

$$\left[-\frac{\hbar^2}{2m} \Delta + V(\vec{r}) \right] \psi_{\vec{k}}(\vec{r}) = \varepsilon(\vec{k}) \psi_{\vec{k}}(\vec{r}) , \tag{6.61}$$

et injectant la forme (6.59), on trouve :

$$-\frac{\hbar^2}{2m} \left[(i\vec{k})^2 u_{\vec{k}} + 2i\vec{k}\cdot\vec{\nabla} u_{\vec{k}} + \Delta u_{\vec{k}} \right] + V(\vec{r}) u_{\vec{k}}(\vec{r}) = \varepsilon(\vec{k}) u_{\vec{k}}(\vec{r}) , \tag{6.62}$$

soit :

$$\frac{\hbar^2}{2m} (-i\vec{\nabla} + \vec{k})^2 u_{\vec{k}}(\vec{r}) + V(\vec{r}) u_{\vec{k}}(\vec{r}) = \varepsilon(\vec{k}) u_{\vec{k}}(\vec{r}) , \tag{6.63}$$

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

$$\frac{1}{2m} (\vec{p} + \hbar\vec{k})^2 u_{\vec{k}}(\vec{r}) + V(\vec{r}) u_{\vec{k}}(\vec{r}) = \varepsilon(\vec{k}) u_{\vec{k}}(\vec{r}) , \tag{6.64}$$

En vertu de la périodicité de $u_{\vec{k}}$, (6.60), il suffit de résoudre (6.63) 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 ne présente pas de saut d'amplitude infinie, ce que l'on suppose. Il en résulte qu'ici les conditions aux limites sont :

$$u_{\vec{k}} \text{ et } \vec{\nabla} u_{\vec{k}} \text{ continues sur la frontière de la maille primitive .} \tag{6.65}$$

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

$$\boxed{H_u(\vec{r}, \vec{p}; \vec{k}) u_{n\vec{k}}(\vec{r}) = \varepsilon_n(\vec{k}) u_{n\vec{k}}(\vec{r})} \tag{6.66}$$

avec, d'après (6.64) :

$$\boxed{H_u(\vec{r}, \vec{p}; \vec{k}) \stackrel{\text{déf}}{=} \frac{1}{2m} (\vec{p} + \hbar\vec{k})^2 + V(\vec{r})} \tag{6.67}$$

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 :

$$\boxed{\varepsilon_n(\vec{k} + \vec{K}) = \varepsilon_n(\vec{k}) \quad \forall \vec{K} \in \tilde{\mathcal{B}}}$$
 (6.68)

6.3.3 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} est²⁶ :

$$i\hbar \frac{d}{dt} \vec{p} = [\vec{p}, H] = -i\hbar \vec{\nabla} V \quad \Longleftrightarrow \quad \frac{d}{dt} \vec{p} = [\vec{p}, H] = -\vec{\nabla} V .$$
 (6.70)

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

$$\vec{p}\psi_{n\vec{k}} = -i\hbar \vec{\nabla} e^{i\vec{k}\cdot\vec{r}} u_{n\vec{k}}(\vec{r}) = -i\hbar [i\vec{k} u_{n\vec{k}} + (\vec{\nabla} u_{n\vec{k}})] e^{i\vec{k}\cdot\vec{r}} = \hbar\vec{k} \psi_{n\vec{k}} - i\hbar e^{i\vec{k}\cdot\vec{r}} (\vec{\nabla} e^{-i\vec{k}\cdot\vec{r}} \psi_{n\vec{k}}) .$$
 (6.71)

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 (6.68)), 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.

6.3.4 Vitesse d'un électron de Bloch

On appelle désormais électron de Bloch un électron décrit par une fonction d'onde de Bloch, soit :

$$\boxed{\psi_{n\vec{k}}(\vec{r}) = e^{i\vec{k}\cdot\vec{r}} u_{n\vec{k}}(\vec{r})}$$
 (6.72)

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 $\psi_{n\vec{k}}$:

$$\langle \vec{v} \rangle_{n\vec{k}} \equiv \langle \psi_{n\vec{k}} | \frac{\vec{p}}{m} | \psi_{n\vec{k}} \rangle = \hbar^{-1} \vec{\nabla}_{\vec{k}} \varepsilon_n \equiv \vec{\nabla}_{\hbar\vec{k}} \varepsilon_n .$$
 (6.73)

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 :

$$H(\lambda)\psi(\lambda) = E(\lambda)\psi(\lambda) \quad \langle \psi(\lambda) | \psi(\lambda) \rangle = 1 ,$$
 (6.74)

on a :

$$\boxed{\frac{dE}{d\lambda} = \langle \psi(\lambda) | \frac{\partial H}{\partial \lambda} | \psi(\lambda) \rangle}$$
 (6.75)

²⁶En laissant tomber l'indice H pour simplifier les notations ; avec les notations précises :

$$\frac{d}{dt} \vec{p}_H = -(\vec{\nabla} V)_H .$$
 (6.69)

En effet :

$$\begin{aligned} \frac{d}{d\lambda} \langle \psi(\lambda) | H(\lambda) | \psi(\lambda) \rangle &= \left\langle \frac{\partial \psi}{\partial \lambda} | H | \psi(\lambda) \right\rangle + \langle \psi(\lambda) | \frac{\partial H}{\partial \lambda} | \psi(\lambda) \rangle + \langle \psi(\lambda) | H(\lambda) | \frac{\partial \psi}{\partial \lambda} \rangle \\ &= E(\lambda) \left(\left\langle \frac{\partial \psi}{\partial \lambda} | \psi(\lambda) \right\rangle + \langle \psi(\lambda) | \frac{\partial \psi}{\partial \lambda} \right) + \langle \psi(\lambda) | \frac{\partial H}{\partial \lambda} | \psi(\lambda) \rangle . \end{aligned} \quad (6.76)$$

La parenthèse est nulle en vertu de la normalisation de ψ , et (6.75) en résulte.

Ici, le Hamiltonien pour la fonction $u_{\vec{k}}$ est H_u donné par l'expression (6.67) ; il dépend du vecteur \vec{k} . Le théorème de Hellmann - Feynman s'écrit ici :

$$\vec{\nabla}_{\vec{k}} \varepsilon_n(\vec{k}) = \langle u_{n\vec{k}} | \vec{\nabla}_{\vec{k}} H_u | u_{n\vec{k}} \rangle = \langle u_{n\vec{k}} | \frac{\hbar}{m} (\vec{p} + \hbar \vec{k}) | u_{n\vec{k}} \rangle = \frac{\hbar^2}{m} \langle u_{n\vec{k}} | -i \vec{\nabla}_{\vec{r}} + \vec{k} | u_{n\vec{k}} \rangle . \quad (6.77)$$

On a :

$$\begin{aligned} (-i \vec{\nabla}_{\vec{r}} + \vec{k}) u_{n\vec{k}} &= (-i \vec{\nabla}_{\vec{r}} + \vec{k}) e^{-i \vec{k} \cdot \vec{r}} \psi_{n\vec{k}}(\vec{r}) = (-i)(-i \vec{k}) e^{-i \vec{k} \cdot \vec{r}} \psi_{n\vec{k}} - i e^{-i \vec{k} \cdot \vec{r}} \vec{\nabla}_{\vec{r}} \psi_{n\vec{k}} + \vec{k} e^{-i \vec{k} \cdot \vec{r}} \psi_{n\vec{k}} \\ &= -i e^{-i \vec{k} \cdot \vec{r}} \vec{\nabla}_{\vec{r}} \psi_{n\vec{k}} . \end{aligned} \quad (6.78)$$

D'où :

$$\vec{\nabla}_{\vec{k}} \varepsilon_n(\vec{k}) = \frac{\hbar^2}{m} (-i) \langle u_{n\vec{k}} | e^{-i \vec{k} \cdot \vec{r}} \vec{\nabla}_{\vec{r}} | \psi_{n\vec{k}} \rangle = -i \frac{\hbar^2}{m} \langle \psi_{n\vec{k}} | \vec{\nabla}_{\vec{r}} | \psi_{n\vec{k}} \rangle = \frac{\hbar}{m} \langle \psi_{n\vec{k}} | \vec{p} | \psi_{n\vec{k}} \rangle , \quad (6.79)$$

soit la relation annoncée :

$$\boxed{\langle \vec{v} \rangle_{n\vec{k}} \stackrel{\text{d'ef}}{=} \langle \psi_{n\vec{k}} | \frac{\vec{p}}{m} | \psi_{n\vec{k}} \rangle = \frac{1}{\hbar} \vec{\nabla}_{\vec{k}} \varepsilon_n(\vec{k})} \quad (6.80)$$

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

$$\langle \vec{v} \rangle = \frac{\partial \varepsilon_n}{\partial (\hbar \vec{k})} \quad \longleftrightarrow \quad \dot{q} = \frac{\partial H}{\partial p} , \quad (6.81)$$

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).

6.3.5 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 :

$$\varepsilon(-\vec{k}) = \varepsilon(\vec{k}) , \quad \vec{\nabla} \varepsilon(\vec{K}/2) = 0 \quad \forall \vec{K} \in \vec{\mathcal{B}} , \quad (6.82)$$

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

La fonction de Bloch satisfait :

$$H \psi_{\vec{k}} = \varepsilon(\vec{k}) \psi_{\vec{k}} . \quad (6.83)$$

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 temps²⁷, 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 :

$$\psi_{\vec{k}}(\vec{r}) = e^{i\vec{k}\cdot\vec{r}} u_{\vec{k}}(\vec{r}) \quad \psi_{\vec{k}}^*(\vec{r}) = e^{-i\vec{k}\cdot\vec{r}} u_{\vec{k}}^*(\vec{r}) . \quad (6.84)$$

$\psi_{\vec{k}}^*(\vec{r})$ 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 (6.84) permet l'identification :

$$u_{-\vec{k}}(\vec{r}) = u_{\vec{k}}^*(\vec{r}) . \quad (6.85)$$

Il en résulte que :

$$\psi_{\vec{k}}^*(\vec{r}) = \psi_{-\vec{k}}(\vec{r}) . \quad (6.86)$$

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

$$H \psi_{-\vec{k}} = \varepsilon(-\vec{k}) \psi_{-\vec{k}} . \quad (6.87)$$

d'où il résulte :

$$\varepsilon(-\vec{k}) = \varepsilon(\vec{k}) . \quad (6.88)$$

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

$$\boxed{\vec{\nabla}\varepsilon(-\vec{k}) = -\vec{\nabla}\varepsilon(\vec{k})} \quad (6.89)$$

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

$$\varepsilon(\vec{k} - \vec{K}) = \varepsilon(\vec{k}) \quad \text{et} \quad \vec{\nabla}\varepsilon(\vec{k} - \vec{K}) = \vec{\nabla}\varepsilon(\vec{k}) . \quad (6.90)$$

Pour les vecteurs d'onde tels que $\vec{k} - \vec{K} = -\vec{k}$, soit $\vec{k} = \vec{K}/2$ (ce sont certains points remarquables de la frontière de la maille de Wigner - Seitz), les équations (6.89) et (6.90) donnent simultanément²⁸ :

$$\vec{\nabla}\varepsilon(-\vec{K}/2) = -\vec{\nabla}\varepsilon(\vec{K}/2) \quad \text{et} \quad \vec{\nabla}\varepsilon(-\vec{K}/2) = \vec{\nabla}\varepsilon(\vec{K}/2) \quad \iff \quad \vec{\nabla}\varepsilon(\vec{K}/2) = 0 . \quad (6.91)$$

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

6.3.6 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 approximation²⁹ –, 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 :

$$\boxed{\varepsilon(\vec{k}) = \varepsilon_F \quad (\text{surface de Fermi})} \quad (6.92)$$

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

²⁷C'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 $t \rightarrow -t$.

²⁸La discussion en cours exclut les cas où $\varepsilon(\vec{k})$ présente des points anguleux.

²⁹A 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 (au sens de Landau - Nozières - Luttinger).

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

$$\varepsilon^{(0)}(\vec{k}) = \frac{\hbar^2 \vec{k}^2}{2m} , \tag{6.93}$$

de sorte que la surface de Fermi a pour équation :

$$\frac{\hbar^2 \vec{k}^2}{2m} = \varepsilon_F . \tag{6.94}$$

C'est donc une sphère, de rayon noté $k_F = \hbar^{-1} \sqrt{2m\varepsilon_F}$. 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 :

$$2 \frac{(4\pi/3) k_F^3}{(2\pi/L)^3} = N \iff \boxed{k_F = \left(3\pi^2 \frac{N}{L^3}\right)^{1/3} \equiv (3\pi^2 \rho)^{1/3}} \tag{6.95}$$

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 occupée³⁰, alors la surface de Fermi est encore définie par l'équation (6.92), 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 $\vec{\mathcal{B}}$ sont identiques, \vec{k} doit être restreint à une maille primitive de $\vec{\mathcal{B}}$, la première zone de Brillouin par exemple. Il est donc utile de préciser :

$$\varepsilon(\vec{k}) = \varepsilon_F \quad \forall \vec{k} \in \text{BZ1} . \tag{6.96}$$

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 :

$$\int_0^{\varepsilon_F} dE 2 \left(\frac{L}{2\pi}\right)^D \int_{\mathbb{R}^D} d^D k \delta(\varepsilon(\vec{k}) - E) = N_e , \tag{6.97}$$

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

$$\int_0^{\varepsilon_F} dE \int_{\mathbb{R}^D} d^D k \delta(\varepsilon(\vec{k}) - E) = 2^{1-D} \pi^D \rho_e , \tag{6.98}$$

où ρ_e est la densité électronique.

Notons enfin que, compte tenu de la périodicité de $\varepsilon(\vec{k})$ dans $\vec{\mathcal{B}}$, la surface de Fermi complète s'obtient par les translations de $\vec{\mathcal{B}}$ de l'objet défini par (6.96) ; 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.

³⁰Quand 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).

Chapitre 7

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.

7.1 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 :

$$\boxed{\psi_{\vec{k}}^{(0)}(\vec{r}) = C e^{i\vec{k}\cdot\vec{r}} \quad \varepsilon^{(0)}(\vec{k}) = \frac{\hbar^2 \vec{k}^2}{2m}} \quad (7.1)$$

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 (7.1), qui sont caractéristiques d'une symétrie de translation continue, celle d'une particule libre (une onde plane telle que (7.1) 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 (7.1), $\hbar\vec{k}$ 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 petit¹. 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.

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

$$\psi_{n\vec{k}-\vec{K}}(\vec{r}) = \psi_{n\vec{k}}(\vec{r}) , \quad \varepsilon_n(\vec{k}-\vec{K}) = \varepsilon_n(\vec{k}) \quad \forall \vec{K} \in \tilde{\mathcal{B}} . \quad (7.2)$$

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 Brillouin².

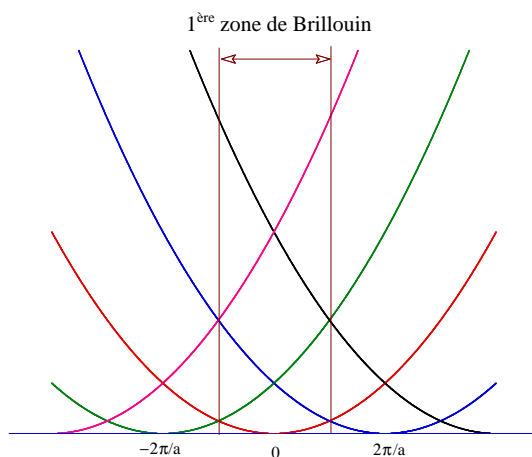


Figure 7.1: Schéma en zones réduites pour un réseau \mathcal{B} unidimensionnel de pas a . Les différentes paraboles représentent les $\varepsilon_{\vec{K}}^{(0)}(\vec{k})$ (voir (7.4)). La première zone de Brillouin s’étend entre $\pm\frac{\pi}{a}$.

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.

Les solutions exactes sont périodiques dans $\tilde{\mathcal{B}}$, ainsi que l’exprime (7.2). D’un autre côté, comme V reste en toute hypothèse petit, on s’attend nécessairement à ce que l’on ait toujours :

$$\varepsilon(\vec{k}) \approx \varepsilon^{(0)}(\vec{k}) . \quad (7.3)$$

¹On 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.

²Dès 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.

La bonne façon de formuler le problème à l'ordre zéro est donc de prolonger par périodisation, conformément à (7.2), 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 7.1). Dans cette représentation, $\forall \vec{k} \in \text{BZ1}$, l'énergie d'un électron en fonction de \vec{k} est infiniment multivaluée³. 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 :

$$\varepsilon_{\vec{K}}^{(0)}(\vec{k}) \stackrel{\text{déf}}{=} \varepsilon^{(0)}(\vec{k} - \vec{K}) = \frac{\hbar^2}{2m}(\vec{k} - \vec{K})^2 \quad \forall \vec{K} \in \tilde{\mathcal{B}} . \quad (7.4)$$

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

$$\varepsilon_{\vec{K}_1}^{(0)}(\vec{k}) = \varepsilon_{\vec{K}_2}^{(0)}(\vec{k}) \iff \varepsilon^{(0)}(\vec{k} - \vec{K}) = \varepsilon^{(0)}(\vec{k}) \quad (7.5)$$

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 (7.5) est satisfaite à chaque fois que :

$$\|\vec{k}\| = \|\vec{k} - \vec{K}\| = \|\vec{K} - \vec{k}\| \quad (7.6)$$

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

$$\varepsilon^{(0)}(\vec{k}) = \varepsilon^{(0)}(\vec{k}') . \quad (7.7)$$

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, (7.7) est équivalente à (7.6).

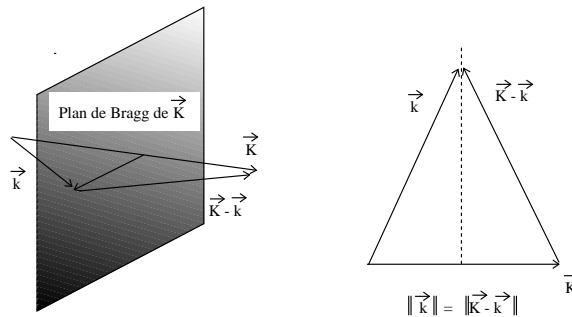


Figure 7.2: Schéma de définition du plan de Bragg d'un vecteur $\vec{K} \in \tilde{\mathcal{B}}$.

L'équation (7.6) 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 Bragg⁴ ; en pareil cas, on pourra toujours poser :

$$\vec{k} = \frac{\vec{K}}{2} + \vec{q} , \quad \vec{q} \perp \vec{K} . \quad (7.8)$$

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 ;

³On 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 $K \in \tilde{\mathcal{B}}$ 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.

⁴sans préjuger d'autres coïncidences par ailleurs, résultant d'une haute symétrie ponctuelle.

- ou bien $\vec{K} = \text{entier impair} \times \vec{b}_0 \equiv (2n_0 + 1)\vec{b}_0$. Alors, $\frac{\vec{k}}{2} = \vec{K}_0 + \frac{1}{2}\vec{b}_0 \notin \vec{\mathcal{B}}$ et la dégénérescence stricte se produit en un certain point de la frontière de la maille de Wigner - Seitz⁵. On parle alors de dégénérescence en bord de zone.

Comme le paramètre \vec{k} varie continûment⁶, 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 (7.9)). Il faut donc construire un schéma de calcul plus subtil que celui énoncé par la théorie des perturbations "académique"⁷.

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 rapport à une énergie typique ; compte tenu des paramètres physiques disponibles, cette énergie ne peut être que⁸ $\frac{\hbar^2}{ma^2}$. Au total, le bon critère s'énonce par une condition du type :

$$|\varepsilon_{\vec{K}}^{(0)}(\vec{k}) - \varepsilon_{\vec{K}'}^{(0)}(\vec{k})| \lesssim |V| \ll \frac{\hbar^2}{ma^2} \quad (7.9)$$

et le traitement de perturbation doit considérer *simultanément* tous les états d'ordre zéro satisfaisant (7.9).

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 (7.9). 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 $\vec{\mathcal{B}}$, 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 :

$$\psi_{\vec{k}}(\vec{r}) = e^{i\vec{k}\cdot\vec{r}} u_{\vec{k}}(\vec{r}) , \quad (7.10)$$

⁵définie comme le polyèdre dont les faces sont les plans de Bragg de tous les vecteurs réciproques joignant l'origine de $\vec{\mathcal{B}}$ 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 $\vec{\mathcal{B}}$ 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).

⁶À une petite quantification près imposée par les conditions cycliques.

⁷On 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.

⁸ a est l'ordre de grandeur de la dimension linéaire d'une maille primitive.

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 :

$$u_{\vec{k}}(\vec{r}) = \sum_{\vec{K} \in \tilde{\mathcal{B}}} U_{\vec{k}}(\vec{k}) e^{-i\vec{K} \cdot \vec{r}} . \quad (7.11)$$

Il en résulte que l'on peut écrire :

$$\psi_{\vec{k}}(\vec{r}) = \sum_{\vec{K} \in \tilde{\mathcal{B}}} U_{\vec{k}}(\vec{k}) e^{i(\vec{k}-\vec{K}) \cdot \vec{r}} . \quad (7.12)$$

On sait que deux états de Bloch \vec{k} et $\vec{k} + \vec{K}_0$, $\vec{K}_0 \in \tilde{\mathcal{B}}$, 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, (7.12) où on remplace \vec{k} par $\vec{k} + \vec{K}_0$ donne :

$$\psi_{\vec{k}+\vec{K}_0}(\vec{r}) = \sum_{\vec{K} \in \tilde{\mathcal{B}}} U_{\vec{k}+\vec{K}_0}(\vec{k}) e^{i(\vec{k}+\vec{K}_0-\vec{K}) \cdot \vec{r}} ; \quad (7.13)$$

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

$$\psi_{\vec{k}+\vec{K}_0}(\vec{r}) = \sum_{\vec{K}' \in \tilde{\mathcal{B}}} U_{\vec{k}+\vec{K}_0}(\vec{K}' + \vec{K}_0) e^{i(\vec{k}-\vec{K}') \cdot \vec{r}} \equiv \sum_{\vec{K} \in \tilde{\mathcal{B}}} U_{\vec{k}+\vec{K}_0}(\vec{K} + \vec{K}_0) e^{i(\vec{k}-\vec{K}) \cdot \vec{r}} ; \quad (7.14)$$

comme $\psi_{\vec{k}} = \psi_{\vec{k}+\vec{K}_0}$, l'identification terme à terme de (7.12) et (7.14) donne :

$$U_{\vec{k}+\vec{K}_0}(\vec{K} + \vec{K}_0) = U_{\vec{k}}(\vec{k}) . \quad (7.15)$$

Ceci signifie que la fonction $U_{\vec{k}}(\vec{k})$ est invariante dans toute translation de $\tilde{\mathcal{B}}$ et ne dépend donc que de $\vec{k} - \vec{K}$. En posant désormais $U_{\vec{k}}(\vec{k}) = c_{\vec{k}-\vec{K}}$, (7.12) prend la forme :

$$\boxed{\psi_{\vec{k}}(\vec{r}) = \sum_{\vec{K} \in \tilde{\mathcal{B}}} c_{\vec{k}-\vec{K}} e^{i(\vec{k}-\vec{K}) \cdot \vec{r}}} \quad (7.16)$$

$\psi_{\vec{k}}(\vec{r})$ 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 (7.2). En termes de *kets*, on peut écrire :

$$|\psi_{\vec{k}}\rangle = \sum_{\vec{K} \in \tilde{\mathcal{B}}} c_{\vec{k}-\vec{K}} |f_{\vec{k}-\vec{K}}\rangle , \quad \langle \vec{r} | f_{\vec{k}-\vec{K}}\rangle = e^{i(\vec{k}-\vec{K}) \cdot \vec{r}} \equiv \psi_{\vec{k}-\vec{K}}^{(0)}(\vec{r}) , \quad (7.17)$$

les $\psi_{\vec{k}-\vec{K}}^{(0)}(\vec{r})$ étant 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 :

$$-\frac{\hbar^2}{2m} \Delta \psi_{\vec{k}}(\vec{r}) + V(\vec{r}) \psi_{\vec{k}}(\vec{r}) = \varepsilon(\vec{k}) \psi_{\vec{k}}(\vec{r}) . \quad (7.18)$$

D'autre part, $\mathcal{V}(\vec{K})$ désignant toujours les composantes de Fourier⁹ du potentiel périodique $V(\vec{r})$:

$$V(\vec{r}) = \sum_{\vec{K} \in \tilde{\mathcal{B}}} \mathcal{V}(\vec{K}) e^{-i\vec{K} \cdot \vec{r}} \iff \mathcal{V}(\vec{K}) = \frac{1}{v} \int_{\text{maille élémentaire de } \mathcal{B}} d^3r V(\vec{r}) e^{i\vec{K} \cdot \vec{r}} . \quad (7.19)$$

En reportant les développements (7.16) et (7.19) dans (7.18), on trouve :

⁹ $\mathcal{V}(\vec{K})$ est homogène à une énergie.

$$\begin{aligned}
& - \sum_{\vec{K} \in \tilde{\mathcal{B}}} \frac{\hbar^2}{2m} [i(\vec{k} - \vec{K})]^2 c_{\vec{k}-\vec{K}} e^{i(\vec{k}-\vec{K}) \cdot \vec{r}} + \sum_{\vec{K}' \in \tilde{\mathcal{B}}, \vec{K}'' \in \tilde{\mathcal{B}}} \mathcal{V}(\vec{K}') c_{\vec{k}-\vec{K}''} e^{-i\vec{K}' \cdot \vec{r}} e^{i(\vec{k}-\vec{K}'') \cdot \vec{r}} = \\
& \sum_{\vec{K} \in \tilde{\mathcal{B}}} \varepsilon(\vec{k}) c_{\vec{k}-\vec{K}} e^{i(\vec{k}-\vec{K}) \cdot \vec{r}} , \quad (7.20)
\end{aligned}$$

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

$$\sum_{\vec{K} \in \tilde{\mathcal{B}}} [\varepsilon^{(0)}(\vec{k} - \vec{K}) - \varepsilon(\vec{k})] c_{\vec{k}-\vec{K}} e^{i(\vec{k}-\vec{K}) \cdot \vec{r}} + \sum_{\vec{K}, \vec{K}' \in \tilde{\mathcal{B}}^2} \mathcal{V}(\vec{K}') c_{\vec{k}-\vec{K}+\vec{K}'} e^{i(\vec{k}-\vec{K}) \cdot \vec{r}} = 0 . \quad (7.21)$$

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 $\psi_{\vec{k}}$:

$$[\varepsilon^{(0)}(\vec{k} - \vec{K}) - \varepsilon(\vec{k})] c_{\vec{k}-\vec{K}} + \sum_{\vec{K}' \in \tilde{\mathcal{B}}} \mathcal{V}(\vec{K}') c_{\vec{k}-\vec{K}+\vec{K}'} = 0 , \quad \forall \vec{K} \in \tilde{\mathcal{B}} , \quad (7.22)$$

ou, de façon équivalente :

$$[\varepsilon^{(0)}(\vec{k} - \vec{K}) - \varepsilon(\vec{k})] c_{\vec{k}-\vec{K}} + \sum_{\vec{K}'' \in \tilde{\mathcal{B}}} \mathcal{V}(\vec{K} - \vec{K}'') c_{\vec{k}-\vec{K}''} = 0 , \quad \forall \vec{K} \in \tilde{\mathcal{B}} . \quad (7.23)$$

\vec{k} étant fixé, il y a une telle équation pour chaque valeur de \vec{K} : il s'agit bien d'un *système* linéaire.

En choisissant convenablement le zéro d'énergie potentielle¹⁰ – 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 (7.23) 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 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 :

$$\int_{\mathbb{R}^3} d^3r e^{-i(\vec{k}-\vec{K}) \cdot \vec{r}} V(\vec{r}) e^{i(\vec{k}-\vec{K}) \cdot \vec{r}} = \int_{\mathbb{R}^3} d^3r V(\vec{r}) = 0 . \quad (7.24)$$

Jusqu'à présent, aucune approximation n'a été faite : le système (7.23) est exact, il est tout juste une reformulation de l'équation aux valeurs propres (7.18) dans l'espace de Fourier. Inexploitable en tant que tel dans le cas général, on en effectue maintenant l'analyse perturbative.

Le système (7.23) s'écrit aussi :

$$\boxed{\left\{ [\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k} - \vec{K})] c_{\vec{k}-\vec{K}} = \sum_{\vec{K}' \in \tilde{\mathcal{B}}, \vec{K}' \neq \vec{K}} \mathcal{V}(\vec{K} - \vec{K}') c_{\vec{k}-\vec{K}'} \right\}_{\vec{K} \in \tilde{\mathcal{B}}}} \quad (7.25)$$

Si $V \equiv 0$, il se réduit à :

$$[\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k} - \vec{K})] c_{\vec{k}-\vec{K}} = 0 , \quad \forall \vec{K} \in \tilde{\mathcal{B}} . \quad (7.26)$$

L'équation (7.26) 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 (7.25) en exhibant à part la ligne $\vec{K} = \vec{K}_1$:

$$[\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k} - \vec{K}_1)] c_{\vec{k}-\vec{K}_1} = \sum_{\vec{K}' \in \tilde{\mathcal{B}}, \vec{K}' \neq \vec{K}_1} \mathcal{V}(\vec{K}_1 - \vec{K}') c_{\vec{k}-\vec{K}'} , \quad (7.27)$$

¹⁰ce que l'on peut faire presque toujours.

$$[\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k} - \vec{K})] c_{\vec{k}-\vec{K}} = \sum_{\vec{K}' \in \tilde{\mathcal{B}}, \vec{K}' \neq \vec{K}} \mathcal{V}(\vec{K} - \vec{K}') c_{\vec{k}-\vec{K}'} , \quad \forall \vec{K} \in \tilde{\mathcal{B}}, \quad \vec{K} \neq \vec{K}_1 . \quad (7.28)$$

Quand $V \rightarrow 0$, un seul crochet $[\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k} - \vec{K})]$ figurant aux premiers membres du système (7.25) s'annule, celui de (7.27), 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 (7.28) 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 (7.27) – qui ne contient que les $c_{\vec{k}-\vec{K}'}$, $\vec{K}' \neq \vec{K}_1$ – est au moins du second ordre¹¹. En définitive, la différence d'énergie $\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k} - \vec{K}_1)$ est du second ordre¹² 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 (7.11) 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 :

$$|\varepsilon^{(0)}(\vec{k} - \vec{K}_i) - \varepsilon^{(0)}(\vec{k} - \vec{K}_j)| \lesssim |V| \quad (i, j = 1, 2, \dots, g) . \quad (7.29)$$

Mettons à part toutes les lignes où ces \vec{K}_i sont au premier membre de (7.25) :

$$[\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k} - \vec{K}_i)] c_{\vec{k}-\vec{K}_i} = \sum_{\vec{K}' \in \tilde{\mathcal{B}}, \vec{K}' \neq \vec{K}_i} \mathcal{V}(\vec{K}_i - \vec{K}') c_{\vec{k}-\vec{K}'} , \quad \vec{K}_i = \vec{K}_1, \vec{K}_2, \dots, \vec{K}_g , \quad (7.30)$$

$$[\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k} - \vec{K})] c_{\vec{k}-\vec{K}} = \sum_{\vec{K}' \in \tilde{\mathcal{B}}, \vec{K}' \neq \vec{K}} \mathcal{V}(\vec{K} - \vec{K}') c_{\vec{k}-\vec{K}'} , \quad \forall \vec{K} \in \tilde{\mathcal{B}}, \quad \vec{K} \neq \vec{K}_i, \quad i = 1, 2, \dots, g . \quad (7.31)$$

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 (7.30) 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 (7.31), 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 (7.25) les seules équations où ces \vec{K}_i figurent au premier membre (il y en a g), il vient :

$$[\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k} - \vec{K}_i)] c_{\vec{k}-\vec{K}_i} = \sum_{\vec{K}' \in \tilde{\mathcal{B}}, \vec{K}' \neq \vec{K}_i} \mathcal{V}(\vec{K}_i - \vec{K}') c_{\vec{k}-\vec{K}'} , \quad \forall \vec{K} \in \{\vec{K}_i\}_{1 \leq i \leq g} . \quad (7.32)$$

Au second membre de (7.32), 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 (7.32) tronqué aux g lignes pertinentes peut donc s'écrire :

$$[\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k} - \vec{K}_i)] c_{\vec{k}-\vec{K}_i} = \sum_{\vec{K}_j \in \{\vec{K}_i\}, j \neq i} \mathcal{V}(\vec{K}_i - \vec{K}_j) c_{\vec{k}-\vec{K}_j} + \mathcal{O}(V^2) \quad (i = 1, 2, \dots, g) , \quad (7.33)$$

où tous les coefficients $c_{\vec{k}-\vec{K}_j}$ explicitement écrits sont $\mathcal{O}(V^0)$. Ceci étant, (7.33) 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 (7.33), et on est réduit à devoir résoudre :

$$\boxed{\left\{ [\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k} - \vec{K}_i)] c_{\vec{k}-\vec{K}_i} = \sum_{\vec{K}_j \in \{\vec{K}_i\}, j \neq i} \mathcal{V}(\vec{K}_i - \vec{K}_j) c_{\vec{k}-\vec{K}_j} \right\}_{i=1, 2, \dots, g}} \quad (7.34)$$

¹¹Comme 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 .

¹²Bien sûr, l'absence de correction d'énergie au 1^{er} ordre est une conséquence du calage de $V(\vec{r})$ assurant que $\mathcal{V}(\vec{K}) = 0$.

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 (7.34), se réduit alors au couple d'équations :

$$[\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k})] c_{\vec{k}} - \mathcal{V}(-\vec{K}) c_{\vec{k}-\vec{K}} = 0 , \quad (7.35)$$

$$-\mathcal{V}(\vec{K}) c_{\vec{k}} + [\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k} - \vec{K})] c_{\vec{k}-\vec{K}} = 0 . \quad (7.36)$$

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

$$\boxed{\varepsilon_{\pm}(\vec{k}) = \Sigma\varepsilon \pm \sqrt{[\Delta\varepsilon]^2 + |\mathcal{V}(\vec{K})|^2} \equiv \Sigma\varepsilon \pm h(\vec{k})} \quad (7.37)$$

où on a posé :

$$\Sigma\varepsilon = \frac{1}{2}[\varepsilon^{(0)}(\vec{k}) + \varepsilon^{(0)}(\vec{k} - \vec{K})] , \quad \Delta\varepsilon = \frac{1}{2}[\varepsilon^{(0)}(\vec{k}) - \varepsilon^{(0)}(\vec{k} - \vec{K})] . \quad (7.38)$$

L'expression (7.37) donne donc la levée de dégénérescence dans le voisinage d'un point de l'espace réciproque où $\varepsilon^{(0)}(\vec{k}) = \varepsilon^{(0)}(\vec{k} - \vec{K})$, vérifiée si $\|\vec{k}\| = \|\vec{k} - \vec{K}\|$, égalité 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 (7.37) donne :

$$\boxed{\varepsilon_{\pm}(\vec{k} \in \mathcal{P}_{\frac{\vec{K}}{2}}) = \varepsilon^{(0)}(\vec{k}) \pm |\mathcal{V}(\vec{K})|} \quad (7.39)$$

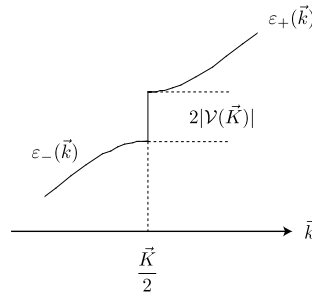
Ainsi, les deux énergies qui coïncidaient à l'ordre zéro quand l'extrémité de \vec{k} est dans un plan de Bragg sont maintenant distinctes : le potentiel périodique a ouvert un *gap* d'énergie, la valeur propre $\varepsilon(\vec{k})$ différant nettement de $\varepsilon^{(0)}(\vec{k})$ - au moins sur le plan qualitatif ; 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 $\mathcal{O}(V^2)$ près.

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

$$\boxed{\varepsilon_{\pm}(\vec{k} = \frac{\vec{K}}{2}) = \varepsilon^{(0)}(\vec{k} = \frac{\vec{K}}{2}) \pm |\mathcal{V}(\vec{K})|} \quad (7.40)$$

(voir fig. 7.3). Il existe un vecteur primitif \vec{b}_0 de $\vec{\mathcal{B}}$ 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

Figure 7.3: Schéma de la levée de dégénérescence pour $\vec{k} = \frac{\vec{K}}{2}$.

périodisée, on voit qu'elle se transforme en une suite d'arcs disjoints, les (petites) discontinuités survenant pour $\vec{k} = \pm\frac{\vec{K}}{2}, \pm\vec{K}, \pm3\frac{\vec{K}}{2}, \dots$. 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 (7.37) en choisissant $\vec{K}_1 = 0$ (alors $\vec{K} = \vec{K}_2$) et, pour la commodité, mesurons \vec{k} relativement à $\frac{\vec{K}}{2}$:

$$\vec{k} = \frac{1}{2}\vec{K} + \vec{q}, \quad (7.41)$$

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 :

$$\Sigma\varepsilon = \frac{1}{2} \frac{\hbar^2}{2m} \left[(\vec{q} + \frac{1}{2}\vec{K})^2 + (\vec{q} - \frac{1}{2}\vec{K})^2 \right] = \varepsilon^{(0)}\left(\frac{\vec{K}}{2}\right) + \varepsilon^{(0)}(\vec{q}), \quad \Delta\varepsilon = \frac{\hbar^2}{m} \vec{q} \cdot \vec{K}. \quad (7.42)$$

Les valeurs propres approchées données par (7.37) s'écrivent maintenant :

$$\varepsilon_{\pm}(\vec{k}) = \varepsilon^{(0)}\left(\frac{\vec{K}}{2}\right) + \varepsilon^{(0)}(\vec{q}) \pm \sqrt{\left(\frac{\hbar^2}{2m}\right)^2 (\vec{K} \cdot \vec{q})^2 + |\mathcal{V}(\vec{K})|^2}. \quad (7.43)$$

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 :

$$\varepsilon_{\pm}(\vec{k} \in \mathcal{P}_{\frac{\vec{K}}{2}}) = \varepsilon^{(0)}\left(\frac{\vec{K}}{2}\right) + \varepsilon^{(0)}(\vec{q}) \pm |\mathcal{V}(\vec{K})| = \varepsilon^{(0)}\left(\frac{\vec{K}}{2}\right) + \frac{\hbar^2}{2m} (\vec{k}^2 - \vec{k} \cdot \vec{K} + \frac{\vec{K}^2}{2}) \pm |\mathcal{V}(\vec{K})|, \quad (7.44)$$

en agrément avec (7.39). De ceci il résulte :

$$(\vec{\nabla}_{\vec{k}} \varepsilon_{\pm})_{\vec{k} \in \mathcal{P}_{\frac{\vec{K}}{2}}} = \frac{\hbar^2}{m} (\vec{k} - \frac{\vec{K}}{2}), \quad (7.45)$$

quantité qui s'annule bien en $\vec{k} = \frac{\vec{K}}{2}$. Le gradient de ε_{\pm} donne la direction de la normale à la surface $\varepsilon_{\pm}(\vec{k}) = \text{C}^{\text{ste}}$. 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 *isolé*.

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

$$\varepsilon_{\pm}(\vec{k} \parallel \vec{K}) = \varepsilon^{(0)}\left(\frac{\vec{K}}{2}\right) + \varepsilon^{(0)}(\vec{q}) \pm \sqrt{\left(\frac{\hbar^2 \vec{K} q}{2m}\right)^2 + |\mathcal{V}(\vec{K})|^2}. \quad (7.46)$$

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 :

$$\begin{aligned} \varepsilon_{\pm}(\vec{k} \parallel \vec{K}, \vec{k} \simeq \frac{\vec{K}}{2}) &\simeq \varepsilon^{(0)}\left(\frac{\vec{K}}{2}\right) + \frac{\hbar^2 q^2}{2m} \pm |\mathcal{V}(\vec{K})| \left[1 + \frac{[\hbar^2 K q / (2m)]^2}{2|\mathcal{V}(\vec{K})|^2}\right] = \\ &\varepsilon^{(0)}\left(\frac{\vec{K}}{2}\right) \pm |\mathcal{V}(\vec{K})| + \frac{\hbar^2 q^2}{2m} \left(1 \pm \frac{\hbar^2 K^2}{4m|\mathcal{V}(\vec{K})|}\right). \end{aligned} \quad (7.47)$$

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, quel que soit \vec{K} :

$$|\mathcal{V}(\vec{K})| \ll \frac{\hbar^2}{ma^2}. \quad (7.48)$$

Il en résulte que le signe du coefficient de q^2 dans (7.47) 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 :

$$m_{\text{eff}} = \hbar^{-2} \left[\frac{\partial^2 \varepsilon}{\partial q^2} \right]^{-1} = \frac{m}{1 \pm \frac{\hbar^2 K^2}{4m} |\mathcal{V}(\vec{K})|}, \quad (7.49)$$

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 :

$$\boxed{m_{\text{eff } uv} = \hbar^{-2} \left[\frac{\partial^2 \varepsilon}{\partial k_u \partial k_v} \right]^{-1} \quad (u, v = x, y, z)} \quad (7.50)$$

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'inversion¹³ :

$$V(-\vec{r}) = V(\vec{r}), \quad (7.51)$$

ce qui, allié au fait que V est réel, assure que $\mathcal{V}(\vec{K}) \in \mathbb{R}$. Les deux équations (7.35) et (7.36) peuvent être considérées comme la forme matricielle de l'équation aux vecteurs propres :

$$H_{\text{eff}} |\psi_{\vec{k}}\rangle = \varepsilon(\vec{k}) |\psi_{\vec{k}}\rangle \quad H_{\text{eff}} = \Sigma \varepsilon \mathbf{1} + \Delta \varepsilon \sigma_z + \mathcal{V}(\vec{k}) \sigma_x, \quad (7.52)$$

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 :

$$H_{\text{eff}} - \Sigma \varepsilon \mathbf{1} = h(\vec{k}) (\cos 2\theta \sigma_z + \sin 2\theta \sigma_x) = h(\vec{k}) e^{-i\theta \sigma_y} \sigma_z e^{+i\theta \sigma_y}, \quad (7.53)$$

où $h(\vec{k}) = \sqrt{(\Delta \varepsilon)^2 + |\mathcal{V}(\vec{K})|^2}$ (voir (7.37)) et avec¹⁴ :

$$\boxed{\tan 2\theta = \frac{\mathcal{V}(\vec{K})}{\Delta \varepsilon / 2}} \quad (7.54)$$

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

$$|\psi_{\vec{k}+}\rangle = e^{-i\theta \sigma_y} |f_{\vec{k}}\rangle, \quad |\psi_{\vec{k}-}\rangle = e^{-i\theta \sigma_y} |f_{\vec{k}-\vec{K}}\rangle. \quad (7.55)$$

¹³C'est toujours vrai pour une structure cristalline sans motif, avec des objets à symétrie sphérique.

¹⁴L'angle θ dépend de \vec{k} , mais cette dépendance est sous-entendue pour simplifier l'écriture.

En effet, on a par exemple :

$$H_{\text{eff}}|\psi_{\vec{k}+}\rangle = h(\vec{k}) e^{-i\theta\sigma_y} \sigma_z e^{+i\theta\sigma_y} e^{-i\theta\sigma_y} |f_{\vec{k}}\rangle = h(\vec{k}) e^{-i\theta\sigma_y} \sigma_z |f_{\vec{k}}\rangle = +h(\vec{k}) e^{-i\theta\sigma_y} |f_{\vec{k}}\rangle \equiv +h(\vec{k}) |\psi_{\vec{k}+}\rangle, \quad (7.56)$$

et de même :

$$H_{\text{eff}}|\psi_{\vec{k}-}\rangle = h(\vec{k}) e^{-i\theta\sigma_y} \sigma_z e^{+i\theta\sigma_y} e^{-i\theta\sigma_y} |f_{\vec{k}-\vec{K}}\rangle = h(\vec{k}) e^{-i\theta\sigma_y} \sigma_z |f_{\vec{k}-\vec{K}}\rangle = -h(\vec{k}) e^{-i\theta\sigma_y} |f_{\vec{k}-\vec{K}}\rangle \equiv -h(\vec{k}) |\psi_{\vec{k}-}\rangle. \quad (7.57)$$

De (7.55) on déduit la forme explicite des fonctions de Bloch approchées $\psi_{\vec{k}\pm}$:

$$\psi_{\vec{k}+}(\vec{r}) = \cos\theta f_{\vec{k}}(\vec{r}) + \sin\theta f_{\vec{k}-\vec{K}}(\vec{r}) = e^{i\vec{k}\cdot\vec{r}} (\cos\theta + \sin\theta e^{-i\vec{K}\cdot\vec{r}}) \equiv e^{i\vec{k}\cdot\vec{r}} u_{\vec{k}+}(\vec{r}), \quad (7.58)$$

$$\psi_{\vec{k}-}(\vec{r}) = -\sin\theta f_{\vec{k}}(\vec{r}) + \cos\theta f_{\vec{k}-\vec{K}}(\vec{r}) = e^{i\vec{k}\cdot\vec{r}} (-\sin\theta + \cos\theta e^{-i\vec{K}\cdot\vec{r}}) \equiv e^{i\vec{k}\cdot\vec{r}} u_{\vec{k}-}(\vec{r}). \quad (7.59)$$

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 (7.58) et (7.59) 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 :

$$\psi_{\vec{k}+}(\vec{r}) = \sqrt{2} e^{i(\vec{k} - \frac{\vec{K}}{2})\cdot\vec{r}} \cos \frac{\vec{K}\cdot\vec{r}}{2}, \quad (7.60)$$

$$\psi_{\vec{k}-}(\vec{r}) = -i\sqrt{2} e^{i(\vec{k} - \frac{\vec{K}}{2})\cdot\vec{r}} \sin \frac{\vec{K}\cdot\vec{r}}{2}. \quad (7.61)$$

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 $\psi_{\vec{k}\pm}(\vec{r})$, (7.58) et (7.59), satisfont le théorème de Bloch (les $u_{\vec{k}\pm}$ sont périodiques dans \mathcal{B}).

7.2 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 :

$$\boxed{\psi_{\nu\vec{R}}(\vec{r}) = \mathcal{T}(\vec{R}) \psi_{\nu}(\vec{r}) = \psi_{\nu}(\vec{r} - \vec{R})} \quad (7.62)$$

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 :

$$\psi_{\nu, \vec{R} \neq \vec{0}}(\vec{r}) = \sum_{\nu'} C_{\nu\nu'}(\vec{r}) \psi_{\nu'}(\vec{r}) . \quad (7.63)$$

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 opérateur¹⁵. 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 :

$$\phi(\vec{r}) = \sum_{\nu=1}^g c_{\nu} \psi_{\nu}(\vec{r}) , \quad (7.64)$$

avec :

$$H_0 \psi_{\nu}(\vec{r}) = E_{\nu} \psi_{\nu}(\vec{r}) . \quad (7.65)$$

Les coefficients c_{ν} dans (7.64) 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 d'un site et placé dans une orbitale $\psi_{\nu, \vec{R}}$ (ou une fonction du type ϕ ci-dessus), fortement localisée. \vec{R} variant, 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 même¹⁶. 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 :

$$\phi_{\vec{R}}(\vec{r}) = \mathcal{T}(\vec{R}) \phi(\vec{r}) = \phi(\vec{r} - \vec{R}) ; \quad (7.66)$$

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

$$\langle \phi_{\vec{R}} | \phi_{\vec{R}'} \rangle \neq 0 \quad \forall \vec{R}, \vec{R}' . \quad (7.67)$$

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 :

$$\boxed{\mathcal{T}(\vec{R}) \psi_{\vec{k}}(\vec{r}) = e^{-i\vec{k} \cdot \vec{R}} \psi_{\vec{k}}(\vec{r}) \iff \psi_{\vec{k}}(\vec{r} + \vec{R}) = e^{+i\vec{k} \cdot \vec{R}} \psi_{\vec{k}}(\vec{r})} \quad (7.68)$$

¹⁵L'inclusion de tels états délocalisés (étendus) ferait repartir clairement en direction de l'approximation des électrons presque libres.

¹⁶Le problème est donc très fortement dégénéré.

Il existe des combinaisons linéaires des $\phi_{\vec{R}}(\vec{r})$ satisfaisant les égalités (7.68). En effet, posons :

$$\psi_{\vec{k}}(\vec{r}) = \sum_{\vec{R}' \in \mathcal{B}} A_{\vec{k}}(\vec{R}') \phi_{\vec{R}'}(\vec{r}) = \sum_{\vec{R}' \in \mathcal{B}} A_{\vec{k}}(\vec{R}') \phi(\vec{r} - \vec{R}') . \quad (7.69)$$

En reportant dans (7.68), il vient :

$$\sum_{\vec{R}' \in \mathcal{B}} A_{\vec{k}}(\vec{R}') \phi(\vec{r} + \vec{R} - \vec{R}') = e^{i\vec{k} \cdot \vec{R}} \sum_{\vec{R}' \in \mathcal{B}} A_{\vec{k}}(\vec{R}') \phi(\vec{r} - \vec{R}') . \quad (7.70)$$

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

$$\sum_{\vec{R}'' \in \mathcal{B}} A_{\vec{k}}(\vec{R} + \vec{R}'') \phi(\vec{r} - \vec{R}'') = e^{i\vec{k} \cdot \vec{R}} \sum_{\vec{R}' \in \mathcal{B}} A_{\vec{k}}(\vec{R}') \phi(\vec{r} - \vec{R}') . \quad (7.71)$$

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

$$A_{\vec{k}}(\vec{R} + \vec{R}') = e^{i\vec{k} \cdot \vec{R}} A_{\vec{k}}(\vec{R}') , \quad (7.72)$$

ce qui fournit la solution :

$$\boxed{A_{\vec{k}}(\vec{r}) = C e^{i\vec{k} \cdot \vec{R}} \quad (C = \text{constante arbitraire})} \quad (7.73)$$

En conclusion, étant donné un jeu de fonctions $\phi_{\vec{R}}$ déduites les unes des autres par les translations¹⁷, toute fonction du type¹⁸ :

$$\psi_{\vec{k}}(\vec{r}) = C \sum_{\vec{R} \in \mathcal{B}} e^{i\vec{k} \cdot \vec{R}} \phi_{\vec{R}}(\vec{r}) \quad (7.75)$$

satisfait le théorème de Bloch¹⁹, quelle que soit la constante arbitraire C , choisie par normalisation²⁰. 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 facilement²¹ :

$$\boxed{N |C|^2 \sum_{\vec{R}} e^{i\vec{k} \cdot \vec{R}} \langle \phi_{\vec{R}=\vec{0}} | \phi_{\vec{R}} \rangle = 1} \quad (7.77)$$

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 :

$$\psi_{n\vec{k}}(\vec{r}) = C \sum_{\vec{R} \in \mathcal{B}} e^{i\vec{k} \cdot \vec{R}} \phi_{n\vec{R}}(\vec{r}) \quad (7.78)$$

¹⁷Noter 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.

¹⁸Il est intéressant de remarquer que, finalement, la fonction de bonne symétrie (7.75) peut être obtenue directement en faisant agir l'opérateur \mathbf{T} :

$$\mathbf{T} = \sum_{\vec{R} \in \mathcal{B}} e^{i\vec{k} \cdot \vec{R}} \mathcal{T}(\vec{R}) \quad (7.74)$$

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 $Y_1 \stackrel{\text{déf}}{=} N!^{-1} \sum_{\lambda} \varepsilon_{\lambda} P_{\lambda}$ à une fonction de symétrie *quelconque* pour obtenir la fonction de symétrie voulue.

¹⁹ $\psi_{\vec{k}}$ 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" $\phi_{\nu\vec{R}}$.

²⁰Si 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.

²¹Au total, la fonction de Bloch normalisée est, pour une seule dimension d'espace et quand on néglige les produits scalaires $\langle \phi_n | \phi_{n'} \rangle$, $n \neq n'$:

$$\psi_{\vec{k}}(x) = \frac{1}{\sqrt{N}} \sum_{n \in \mathbb{Z}} e^{i n k a} \phi_n(x) . \quad (7.76)$$

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 :

$$\mathcal{T}(\vec{R})\psi_{n\vec{k}}(\vec{r}) = e^{-i\vec{k}\cdot\vec{R}}\psi_{n\vec{k}}(\vec{r}) . \quad (7.79)$$

En effet, H commute avec tous les opérateurs $\mathcal{T}(\vec{R})$:

$$[H, \mathcal{T}(\vec{R})] = 0 \iff H\mathcal{T}(\vec{R}) = \mathcal{T}(\vec{R})H . \quad (7.80)$$

Prenons l'élément de matrice de (7.80) entre deux fonctions $\psi_{n\vec{k}}$ et $\psi_{n'\vec{k}'}$:

$$\langle\psi_{n\vec{k}}|H\mathcal{T}(\vec{R})|\psi_{n'\vec{k}'}\rangle = \langle\psi_{n\vec{k}}|\mathcal{T}(\vec{R})H|\psi_{n'\vec{k}'}\rangle . \quad (7.81)$$

En vertu de (7.79), le premier membre est :

$$e^{-i\vec{k}'\cdot\vec{R}}\langle\psi_{n\vec{k}}|H|\psi_{n'\vec{k}'}\rangle . \quad (7.82)$$

Par définition de l'adjoint d'un opérateur, le second membre de (7.81) est :

$$(|\psi_{n\vec{k}}\rangle, \mathcal{T}(\vec{R})H|\psi_{n'\vec{k}'}\rangle) = ([\mathcal{T}(\vec{R})]^\dagger|\psi_{n\vec{k}}\rangle, H|\psi_{n'\vec{k}'}\rangle) . \quad (7.83)$$

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

$$(\mathcal{T}(-\vec{R})|\psi_{n\vec{k}}\rangle, H|\psi_{n'\vec{k}'}\rangle) = (e^{-i\vec{k}\cdot(-\vec{r})}|\psi_{n\vec{k}}\rangle, H|\psi_{n'\vec{k}'}\rangle) = e^{-i\vec{k}\cdot\vec{R}}(|\psi_{n\vec{k}}\rangle, H|\psi_{n'\vec{k}'}\rangle) , \quad (7.84)$$

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 (7.82) et (7.84), il vient :

$$(e^{-i\vec{k}'\cdot\vec{R}} - e^{-i\vec{k}\cdot\vec{R}})\langle\psi_{n\vec{k}}|H|\psi_{n'\vec{k}'}\rangle = 0 , \quad (7.85)$$

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 \vec{k} mais de n différents :

$$\langle\psi_{n\vec{k}}|H|\psi_{n'\vec{k}}\rangle \neq 0 . \quad (7.86)$$

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 (7.64) se réduit à ψ_1 :

$$\phi(\vec{r}) = \psi_1(\vec{r}) , \quad H_0\psi_1 = E_1\psi_1 . \quad (7.87)$$

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 (7.85), 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 :

$$\psi_{\vec{k}} = C \sum_{\vec{R} \in \mathcal{B}} e^{i\vec{k}\cdot\vec{R}} \psi_1(\vec{r} - \vec{R}) , \quad (7.88)$$

d'où :

$$\varepsilon(\vec{k}) = |C|^2 \sum_{\vec{R}} \sum_{\vec{R}'} e^{-i\vec{k}\cdot\vec{R}'} e^{i\vec{k}\cdot\vec{R}} \langle\psi_{1\vec{R}'}|H|\psi_{1\vec{R}}\rangle . \quad (7.89)$$

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

$$\langle \psi_1 |_{\vec{R}'} | H | \psi_1 |_{\vec{R}} \rangle = \int d^3r \psi_1^*(\vec{r} - \vec{R}') H \psi_1(\vec{r} - \vec{R}) = \int d^3r \psi_1^*(\vec{r}) H \psi_1(\vec{r} - \vec{R} + \vec{R}') \equiv \langle \psi_1 | H | \psi_1 |_{\vec{R} - \vec{R}'} \rangle , \quad (7.90)$$

d'où :

$$\varepsilon(\vec{k}) = |C|^2 \sum_{\vec{R}} \sum_{\vec{R}'} e^{i\vec{k} \cdot (\vec{R} - \vec{R}')} \langle \psi_1 | H | \psi_1 |_{\vec{R} - \vec{R}'} \rangle = |C|^2 \sum_{\vec{R}} \sum_{\vec{R}''} e^{i\vec{k} \cdot \vec{R}''} \langle \psi_1 | H | \psi_1 |_{\vec{R}''} \rangle . \quad (7.91)$$

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

$$\varepsilon(\vec{k}) = N |C|^2 \sum_{\vec{R}''} e^{i\vec{k} \cdot \vec{R}''} \langle \psi_1 | H | \psi_1 |_{\vec{R}''} \rangle . \quad (7.92)$$

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

$$H = \frac{\vec{p}^2}{2m} + \sum_{\vec{R} \in \mathcal{B}} v(\vec{r} - \vec{R}) , \quad (7.93)$$

où $v(\vec{r})$ est l'énergie potentielle d'un électron soumis à l'attraction du seul nœud situé à l'origine²² ; en isolant le terme $\vec{R} = 0$:

$$H = \left[\frac{\vec{p}^2}{2m} + v(\vec{r}) \right] + \sum_{\vec{R} \in \mathcal{B}^*} v(\vec{r} - \vec{R}) \equiv H_0 + \Delta V , \quad (7.94)$$

\mathcal{B}^* désignant le réseau privé du site à l'origine. Maintenant, l'élément de matrice du Hamiltonien apparaissant dans (7.92) est :

$$\langle \psi_1 | H_0 + \Delta V | \psi_1 |_{\vec{R}} \rangle = E_1 \langle \psi_1 | \psi_1 |_{\vec{R}} \rangle + \langle \psi_1 | \Delta V | \psi_1 |_{\vec{R}} \rangle , \quad (7.95)$$

d'où la nouvelle forme de (7.91) compte tenu de (7.77) :

$$\varepsilon(\vec{k}) = \frac{\sum_{\vec{R}} e^{i\vec{k} \cdot \vec{R}} (E_1 \langle \psi_1 | \psi_1 |_{\vec{R}} \rangle + \langle \psi_1 | \Delta V | \psi_1 |_{\vec{R}} \rangle)}{\sum_{\vec{R}} e^{i\vec{k} \cdot \vec{R}} \langle \psi_1 | \psi_1 |_{\vec{R}} \rangle} = E_1 + \frac{\sum_{\vec{R}} e^{i\vec{k} \cdot \vec{R}} \langle \psi_1 | \Delta V | \psi_1 |_{\vec{R}} \rangle}{\sum_{\vec{R}} e^{i\vec{k} \cdot \vec{R}} \langle \psi_1 | \psi_1 |_{\vec{R}} \rangle} . \quad (7.96)$$

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 :

$$\varepsilon(\vec{k}) = E_1 + \frac{\langle \psi_1 | \Delta V | \psi_1 \rangle + \sum_{\langle \vec{0}, \vec{R} \rangle} e^{i\vec{k} \cdot \vec{R}} \langle \psi_1 | \Delta V | \psi_1 |_{\vec{R}} \rangle}{1 + \sum_{\langle \vec{0}, \vec{R} \rangle} e^{i\vec{k} \cdot \vec{R}} \langle \psi_1 | \psi_1 |_{\vec{R}} \rangle} . \quad (7.97)$$

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 :

$$\varepsilon(\vec{k}) = E_1 + \langle \psi_1 | \Delta V | \psi_1 \rangle + \sum_{\langle \vec{0}, \vec{R} \rangle} e^{i\vec{k} \cdot \vec{R}} \langle \psi_1 | \Delta V | \psi_1 |_{\vec{R}} \rangle . \quad (7.98)$$

Le second terme peut être absorbé dans une redéfinition²³ de E_1 , notée ε_0 ; avec ces diverses approximations, le résultat final retenu est :

$$\boxed{\varepsilon(\vec{k}) = \varepsilon_0 + \sum_{\langle \vec{0}, \vec{R} \rangle} e^{i\vec{k} \cdot \vec{R}} \langle \psi_1 | \Delta V | \psi_1 |_{\vec{R}} \rangle} \quad (7.99)$$

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 .

²²On obtient bien le potentiel du réseau en effectuant sur $v(\vec{r})$ toutes les translations du réseau de Bravais.

²³ou, pour faire pédant, une "renormalisation".

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 (7.99) donne :

$$\varepsilon(\vec{k}) = \varepsilon_0 + 2w (\cos k_x a + \cos k_y a + \cos k_z a) . \quad (7.100)$$

Cette expression est visiblement périodique ($\varepsilon(\vec{k} + \vec{K}) = \varepsilon(\vec{k})$). Elle définit une bande d'énergie s'étendant entre les deux valeurs²⁴ $\varepsilon_0 \pm 6|w|$; mesurée à la bonne aune, la dispersion est effectivement faible (bande étroite), puisque $|w|$ est réputé petit par rapport aux différences d'énergie $|E_\nu - E_{\nu'}|$.

◆ *Remarque*

D'une façon générale (et inversement), à partir de tout état de Bloch exact²⁵ $\psi_{\vec{k}}$, on peut indépendamment définir sa transformée de Fourier par :

$$\psi_{\vec{k}}(\vec{r}) = \frac{1}{\sqrt{N}} \sum_{\vec{R} \in \mathcal{B}} e^{i\vec{k} \cdot \vec{R}} \phi_{\vec{R}}^{(W)}(\vec{r}) \iff \phi_{\vec{R}}^{(W)}(\vec{r}) = \frac{1}{\sqrt{N}} \sum_{\vec{k} \in BZ1} e^{-i\vec{k} \cdot \vec{R}} \psi_{\vec{k}}(\vec{r}) . \quad (7.101)$$

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 :

$$\langle \phi_{\vec{R}}^{(W)} | \phi_{\vec{R}'}^{(W)} \rangle = \delta_{\vec{R} \vec{R}'} , \quad (7.102)$$

et constituent une autre base possible de représentation ; l'orthogonalité des fonctions de Wannier résulte du fait que la transformation (7.101) 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 (7.66).◆

7.3 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 :

$$\text{Nombre d'états d'énergie comprise entre } E_0 \text{ et } E_0 + \delta E = \mathcal{D}(E_0) \delta E . \quad (7.103)$$

$\mathcal{D}(E_0) \delta E$ 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 ($\delta^3 k = (\frac{2\pi}{L})^3$) :

$$\boxed{\sum_{\text{états}} \equiv 2 \sum_{\vec{k}} \rightarrow 2 \left(\frac{L}{2\pi}\right)^3 \int d^3 k} \quad (7.104)$$

²⁴Le 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 21) ; le changement de signe de w peut s'obtenir par un rephasage $(-1)^n$ de chaque ϕ_n , ce qui remplace simplement k par $k - \frac{\pi}{a} (\frac{2\pi}{a})$: 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. . .

²⁵c'est-à-dire obtenu en résolvant exactement l'équation pour $u_{\vec{k}}$.

En conséquence, la densité d'états de la bande d'énergie $\varepsilon_n(\vec{k})$ est donnée par²⁶ :

$$\mathcal{D}_n(E_0) = 2L^3 \int_{\mathbb{R}^3} \frac{d^3k}{(2\pi)^3} \delta(\varepsilon_n(\vec{k}) - E_0) \quad ([\mathcal{D}] = E^{-1}) . \quad (7.105)$$

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 monotone²⁷ :

$$\delta(f(x) - f_0) = \frac{1}{|f'(x_i)|} \delta(x - x_i) . \quad (7.107)$$

À trois dimensions, cette relation se généralise en²⁸ :

$$\delta(f(\vec{x}) - f_0) = \frac{1}{\|\vec{\nabla} f(\vec{x}_0)\|} \delta((\vec{x} - \vec{x}_0) \cdot \vec{n}) , \quad f(\vec{x}_0) = f_0 , \quad (7.111)$$

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 (7.105) s'écrit :

$$\mathcal{D}_n(E_0) = L^3 \int_{\mathbb{R}^3} \frac{d^3k}{4\pi^3} \frac{1}{\|\vec{\nabla} \varepsilon_n(\vec{k}_0)\|} \delta((\vec{k} - \vec{k}_0) \cdot \vec{n}) , \quad \varepsilon_n(\vec{k}_0) = E_0 . \quad (7.112)$$

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 (7.112) pour la fonction $\mathcal{D}_n(E)$ est donc :

$$\boxed{\mathcal{D}_n(E) = L^3 \int_{S_E} \frac{d^2k}{4\pi^3} \frac{1}{\|\vec{\nabla} \varepsilon_n(\vec{k})\|}} \quad (7.113)$$

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*, $\vec{\nabla} \varepsilon_n$ 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 :

$$\mathcal{D}(E) = \sum_n \mathcal{D}_n(E) , \quad (7.114)$$

de sorte que l'allure générale de $\mathcal{D}(\mathcal{E})$ est celle indiquée sur la figure 7.4, 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

²⁶La densité d'états est parfois définie par unité de volume, ce qui revient à poser $L = 1$ dans la définition (7.105).

²⁷Si la fonction n'est pas monotone, l'égalité $f(x) = f_0$ peut être satisfaite pour plusieurs valeurs x_i de x ; alors :

$$\delta(f(x) - f_0) = \sum_i \frac{1}{|f'(x_i)|} \delta(x - x_i) \quad f(x_i) = f_0 \quad \forall i . \quad (7.106)$$

²⁸En effet, le développement de Taylor au voisinage de $\vec{x}_0 \in S_{f_0}$ s'écrit :

$$f(\vec{x}) = f_0 + (\vec{x} - \vec{x}_0) \cdot \vec{\nabla} f(\vec{x}_0) + \mathcal{O}((\vec{x} - \vec{x}_0)^2) . \quad (7.108)$$

Par définition, $\vec{\nabla} f$ est perpendiculaire à la surface S_{f_0} , donc dirigé suivant la normale \vec{n} ; il en résulte que (7.108) se réécrit comme :

$$f(\vec{x}) = f_0 + \|\vec{x} - \vec{x}_0\| \|\vec{\nabla} f(\vec{x}_0)\| + \mathcal{O}((\vec{x} - \vec{x}_0)^2) , \quad (7.109)$$

ce qui donne :

$$\delta(f(\vec{x}) - f_0) = \delta(\|\vec{\nabla} f(\vec{x}_0)\| (\vec{x} - \vec{x}_0) \cdot \vec{n}) , \quad (7.110)$$

et compte tenu de (7.107), on trouve (7.111). On vérifie sans peine l'homogénéité de (7.111), sachant que $[\delta(\vec{x} \cdot \vec{n})] = L^{-1}$ (alors que $[\delta(\vec{x})] = L^{-3}$).

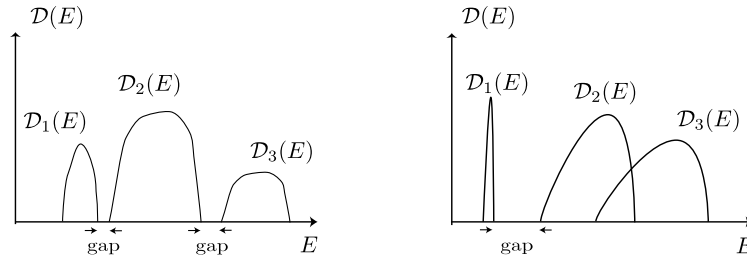


Figure 7.4: Illustration schématique de la densité d'états. À droite, l'exemple de deux bandes qui se recouvrent.

recouvrent en énergie (voir fig. 7.4, 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.

$\mathcal{D}(E)$ 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 :

$$\mathcal{A} = \sum_{\text{états}} A(E) \rightarrow 2 \left(\frac{L}{2\pi} \right)^3 \int d^3k A(E), \quad (7.115)$$

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 :

$$\mathcal{A} = 2 \left(\frac{L}{2\pi} \right)^3 \int d^3k \delta(\varepsilon(\vec{k}) - E) A(E). \quad (7.116)$$

Avec les règles habituelles concernant la fonction de Dirac rappelées plus haut, (7.116) devient :

$$2 \left(\frac{L}{2\pi} \right)^3 \int dE \int_{S_E} d^2k \frac{1}{\|\vec{\nabla}_{\varepsilon}(\vec{k})\|} A(E), \quad (7.117)$$

Au total, on a donc :

$$\mathcal{A} = \int dE \mathcal{D}(E) A(E) \equiv \int dE \tilde{A}(E), \quad (7.118)$$

Les distinctions quantitatives des deux situations extrêmes étudiées plus haut peuvent s'illustrer comme indiqué sur la figure 7.5 : 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.

On est maintenant en mesure de comprendre – en raisonnant à température nulle²⁹ – 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

²⁹Pour 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 $Y(\varepsilon_F - E)$.

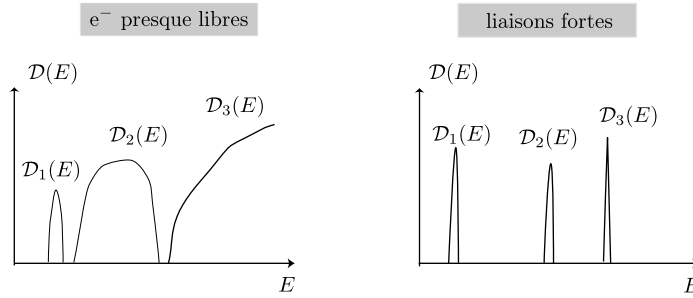


Figure 7.5: Illustration schématique des différences de la densité d'états pour les approximations des électrons presque libres (à gauche) et des liaisons fortes (à droite).

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 conduction³⁰. 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 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'ambiante³¹), 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 (intrinsèque³²). La résistivité d'un semi-conducteur *décroit* quand la température augmente, ce qui est le comportement inverse de celui observé dans les métaux.
2. la plus haute bande est incomplètement remplie, ce qui situe le niveau de Fermi ε_F à l'intérieur de cette bande³³ 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 base³⁴ (cristal monoatomique) et si chaque atome fournit un seul électron ($n_e = 1$), alors la dernière bande est à moitié remplie³⁵ – 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 isolant³⁶

³⁰l'avant-dernière bande complètement remplie s'appelle bande de valence.

³¹La dilatation du réseau avec la température donne une dépendance en température aux énergies des *gaps*.

³²Cette 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.

³³ 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*.

³⁴On rappelle que la structure cristalline d'un solide est complètement définie en donnant le réseau de Bravais \mathcal{B} et l'objet physique (atome, molécule, ion, ...) situé aux nœuds de ce réseau. Si l'on veut éviter une référence physique, on appelle simplement *motif* (ou *base*) l'objet situé à chaque nœud.

³⁵On parle alors de *demi-remplissage*.

³⁶Usuellement, 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.

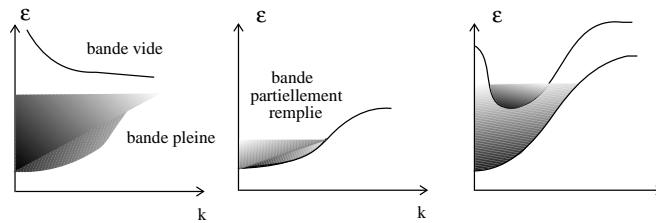


Figure 7.6: Illustration schématique selon la théorie des bandes (théorie à *une* particule) de la différence entre isolant (à gauche), métal (au milieu) et semi-métal (à droite).

L'existence de recouvrement³⁷ 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.

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 $\vec{v} = \hbar^{-1} \vec{\nabla}_{\vec{k}} \epsilon$: quand la bande est plate, la vitesse est nulle.

Dans les cas moins extrêmes, les grandeurs relatives des composantes de $\vec{\nabla}_{\vec{k}} \epsilon$ 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 supérieure³⁸ 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 diffuse³⁹ 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.

³⁷Un 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).

³⁸environ 2.4 fois plus grande.

³⁹On parle alors usuellement d'états *étendus*.

◆ *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 statique⁴⁰, coïncidant pour cette situation de demi-remplissage, au réseau de Bravais sous-jacent du solide⁴¹. Il est physiquement évident qu'un tel solide ne 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 réminiscent 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 antiferromagnétique⁴², 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 considérée⁴³, on voit qu'il y a une

⁴⁰Dans le régime des faibles densités ($r_s \gg a_0$), c'est le réseau de Wigner.

⁴¹Quand 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*).

⁴²Cette affirmation, sommaire et désinvolte, laisse de côté de grandes difficultés. En effet, la structure antiferromagnétique ne s'impose pas en l'état actuel des choses puisque tout basculement de spins conservant $M_S = 0$ ne coûte rien ; on peut même imaginer une structure en domaines (des groupes de spins tous alignés, soit vers le haut, soit vers le bas). En définitive, dans la limite $\alpha = 0$, la dégénérescence est infinie. Savoir ce que fait réellement le système à α petit est une question d'une grande difficulté. Le fondamental du Hamiltonien de Hubbard a été obtenu en 1968 par Lieb et Wu ; dans le strict cadre du modèle de Hubbard, il n'y a pas de transition de Mott.

⁴³Après 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.

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 possible⁴⁴. 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. ♦

⁴⁴Visiblement, l'énergie de la bande pleine avec les *gaps* est plus basse que celle de la bande non déformée.

Chapitre 8

Vibrations d'un solide ordonné

Jusqu'à présent, les ions (atomes, cœurs, ions, *etc.*) ont été considérés comme fixes, situés aux nœuds d'un réseau de Bravais \mathcal{B} . Ce cadre restrictif doit être dépassé, pour plusieurs raisons, dont les plus importantes sont les suivantes :

- les vibrations du réseau mettent en jeu des degrés de liberté intervenant dans la chaleur spécifique C_V . Leur quantification permet de rendre compte du comportement basse température de celle-ci, l'un des premiers succès de la quantification (Einstein, 1905), dans le droit fil de l'idée-maîtresse de Planck (1900) : contrairement à ce qu'affirme la loi (classique) de Dulong et Petit, C_V n'est pas constante et contient un terme $\propto T^3$. Comme toujours, ce terme est essentiel à basse température, $k_B T \lesssim \hbar \omega_{\text{vibr}}$, où ω_{vibr} est une pulsation typique des vibrations du réseau.
- La conduction dans un métal relève usuellement de la loi d'Ohm, laquelle ne peut être comprise sans l'intervention d'une dissipation d'énergie. Les vibrations y jouent un rôle majeur : le mouvement des électrons est un peu un jeu de billard où ceux-ci effectuent une sorte de marche au hasard en raison des collisions inélastiques¹ sur les ions (cœurs) du réseau. La dissipation est assurée par le caractère inélastique des collisions, rendu possible par la mise en mouvement des cœurs.
- Les spectres de diffraction – dont la théorie a été faite comme si les centres diffuseurs étaient fixes – sont évidemment altérés par les vibrations, donnant des élargissements spectraux dont l'analyse apporte de précieuses informations. En outre, il s'agit d'un problème où la dimensionnalité du réseau joue un rôle important.

La question des vibrations des centres massifs se pose en soi, indépendamment de l'arrangement spatial de ces derniers (cristallisé, vitreux, amorphe, *etc.*). Pour simplifier, on ne considère dans la suite que les vibrations d'un cristal parfait, complètement défini par la donnée de sa structure (réseau de Bravais et motif).

8.1 L'approximation harmonique

La considération des vibrations repose cependant sur l'idée que le cristal existe, et notamment que l'énergie (cinétique et potentielle) due à la vibration reste petite par rapport à l'énergie de cohésion. En pratique, ceci contient une hypothèse de petites oscillations : l'amplitude des vibrations doit rester petite devant les distances moyennes, assurant que l'énergie de vibration (potentielle et cinétique) est faible comparée à l'énergie statique d'interaction conduisant à l'énergie de cohésion² (dans le cas contraire, cela signifierait que le cristal est proche

¹Ceci ne préjuge en rien de l'existence d'autres causes de dissipation : impuretés chimiques ou structurales, dislocations,...

²Dans l'hypothèse harmonique qui va être faite, les énergies cinétique et potentielle de vibration sont égales entre elles, chacune représentant très exactement la moitié de l'énergie totale de vibration. On utilise parfois le mot *élastique* au lieu de *harmonique*, par opposition à *plastique* ; dans ce dernier cas, une fois la contrainte relâchée, chaque atome ne revient pas à son point de départ.

de son point de fusion). Dans toute la suite, on maintient donc l'existence d'un réseau de Bravais de référence – les nœuds de ce réseau sont alors les positions moyennes (d'équilibre) des objets qui vibrent – et on fait l'hypothèse de petites oscillations. Notant $\vec{r}(\vec{R})$ la position instantanée de l'ion³ vibrant près du nœud \vec{R} , on peut toujours écrire :

$$\vec{r}(\vec{R}) = \vec{R} + \vec{u}(\vec{R}) , \quad (8.1)$$

\vec{u} désignant l'écart à la position d'équilibre (qui est aussi la position moyenne) \vec{R} . Le Hamiltonien du réseau est :

$$H = \sum_{\vec{R} \in \mathcal{B}} \frac{\vec{P}_{\vec{r}(\vec{R})}^2}{2M} + \frac{1}{2} \sum_{\vec{R}, \vec{R}' \in \mathcal{B}} V(\vec{r}(\vec{R}) - \vec{r}(\vec{R}')) \equiv T + U \quad (8.2)$$

Compte tenu de (8.1), l'énergie potentielle s'écrit :

$$U = \frac{1}{2} \sum_{\vec{R}, \vec{R}' \in \mathcal{B}} V(\vec{R} - \vec{R}' + \vec{u}(\vec{R}) - \vec{u}(\vec{R}')) , \quad (8.3)$$

et, dans l'hypothèse où les écarts sont faibles, le développement de Taylor au voisinage des positions d'équilibre \vec{R} et \vec{R}' donne :

$$U = U_{\text{eq}} + U_{\text{h}} , \quad (8.4)$$

avec :

$$U_{\text{eq}} = \frac{1}{2} \sum_{\vec{R}, \vec{R}' \in \mathcal{B}} V(\vec{R} - \vec{R}') , \quad (8.5)$$

qui correspond au réseau immobile, et :

$$U_{\text{h}} = \frac{1}{2} \sum_{\vec{R}, \vec{R}' \in \mathcal{B}} \sum_{\alpha, \beta=x, y, z} \frac{1}{2!} [\vec{u}(\vec{R}) - \vec{u}(\vec{R}')]_{\alpha} K_{\alpha\beta}(\vec{R} - \vec{R}') [\vec{u}(\vec{R}) - \vec{u}(\vec{R}')]_{\beta} . \quad (8.6)$$

Dans cette expression, les $K_{\alpha\beta}$ sont les dérivées secondes :

$$K_{\alpha\beta} = \left(\frac{\partial^2 V}{\partial r_{\alpha} \partial r_{\beta}} \right)_{\vec{R}, \vec{R}'} , \quad (8.7)$$

calculées aux positions d'équilibre du réseau d'ions fixes ; la notation $K_{\alpha\beta}$ évoque une constante de raideur de ressort parfait ($[K_{\alpha\beta}] = \text{MT}^{-2}$). U_{eq} est le terme donnant lieu à l'énergie de cohésion dans l'approximation des cœurs fixes. Arrêter le développement au second ordre inclus n'est manifestement sensé que dans l'hypothèse acceptée de petites oscillations. En réarrangeant les sommations, le terme harmonique⁴ U_{h} donné par (8.6) peut toujours être mis sous la forme :

$$U_{\text{h}} = \frac{1}{2} \sum_{\vec{R}, \vec{R}' \in \mathcal{B}} \sum_{\alpha, \beta=x, y, z} u_{\alpha}(\vec{R}) Q_{\alpha\beta}(\vec{R} - \vec{R}') u_{\beta}(\vec{R}') \quad (8.8)$$

où les $Q_{\alpha\beta}$ sont des combinaisons linéaires des constantes de raideur $K_{\alpha\beta}$. En outre, si on donne le même déplacement \vec{u}_0 à tous les atomes (translation en bloc), l'énergie potentielle ne change pas. Il en résulte que :

$$0 = \frac{1}{2} \sum_{\vec{R}, \vec{R}' \in \mathcal{B}} \sum_{\alpha, \beta=x, y, z} u_{0,\alpha}(\vec{R}) Q_{\alpha\beta}(\vec{R} - \vec{R}') u_{0,\beta}(\vec{R}') , \quad (8.9)$$

soit :

$$\sum_{\alpha, \beta=x, y, z} u_{0,\beta}(\vec{R}') u_{0,\alpha}(\vec{R}) \sum_{\vec{R}, \vec{R}' \in \mathcal{B}} Q_{\alpha\beta}(\vec{R} - \vec{R}') = 0 \quad \forall \vec{u}_0(\vec{R}) . \quad (8.10)$$

³Pour simplifier les notations, on suppose ici le réseau sans base. La généralisation, à ce stade, est un pur jeu d'écriture. Le cas d'un réseau avec base est examiné dans la suite.

⁴Le terme U_{eq} ne joue aucun rôle en ce qui concerne les vibrations, et sera omis dans la suite. Il reste que c'est ce terme qui, précisément, fixe la configuration d'équilibre, et donne l'énergie de cohésion dans l'hypothèse d'ions fixes.

La dernière somme est $N \sum_{\vec{R}} Q_{\alpha\beta}(\vec{R})$; d'où :

$$\boxed{\sum_{\vec{R} \in \mathcal{B}} Q_{\alpha\beta}(\vec{R}) = 0 \quad \forall \alpha, \beta} \quad (8.11)$$

Une énergie potentielle étant définie, il est possible d'écrire les équations du mouvement (classiques⁵) pour chaque coordonnée $\vec{r}(\vec{R})$, étant entendu que $\dot{\vec{r}}(\vec{R}) = \dot{\vec{u}}(\vec{R})$. On a :

$$\begin{aligned} M \ddot{u}_\alpha(\vec{R}) &= -\frac{\partial}{\partial u_\alpha(\vec{R})} \frac{1}{2} \sum_{\vec{R}', \vec{R}''} \sum_{\alpha', \beta'} u_{\alpha'}(\vec{R}') Q_{\alpha'\beta'}(\vec{R}' - \vec{R}'') u_{\beta'}(\vec{R}'') \\ &= -\frac{1}{2} \sum_{\vec{R}', \vec{R}''} \sum_{\alpha', \beta'} [\delta_{\alpha'\alpha} \delta_{\vec{R}' \vec{R}} u_{\beta'}(\vec{R}'') + \delta_{\beta'\alpha} \delta_{\vec{R}'' \vec{R}} u_{\alpha'}(\vec{R}')] Q_{\alpha'\beta'}(\vec{R}' - \vec{R}'') . \end{aligned} \quad (8.12)$$

En jouant avec le nom des indices muets de sommation, ceci se réécrit :

$$\boxed{M \ddot{u}_\alpha(\vec{R}) = -\frac{1}{2} \sum_{\vec{R}'} \sum_{\beta} [Q_{\alpha\beta}(\vec{R} - \vec{R}') + Q_{\beta\alpha}(\vec{R}' - \vec{R})] u_\beta(\vec{R}')} \quad (8.13)$$

$Q_{\alpha\beta}$ est une matrice symétrique, puisque l'ordre des dérivations incluses dans (8.8) – voir aussi (8.7) – est sans importance. Par ailleurs, tout réseau de Bravais a un centre d'inversion⁶ : $Q_{\alpha\beta}(\vec{R})$ est donc invariant dans le changement $\vec{R} \rightarrow -\vec{R}$; au total :

$$Q_{\alpha\beta}(\vec{R}) = Q_{\beta\alpha}(\vec{R}) = Q_{\alpha\beta}(-\vec{R}) . \quad (8.14)$$

Compte tenu de ces symétries, les équations du mouvement (8.13) deviennent :

$$M \ddot{u}_\alpha(\vec{R}, t) = - \sum_{\vec{R}'} \sum_{\beta=x,y,z} Q_{\alpha\beta}(\vec{R} - \vec{R}') u_\beta(\vec{R}', t) . \quad (8.15)$$

8.2 Modes normaux d'un réseau unidimensionnel

Afin d'illustrer les idées précédentes, considérons un réseau unidimensionnel de paramètre a , dont les atomes peuvent vibrer le long du réseau⁷. En vertu du caractère unidimensionnel, chaque déplacement \vec{u} n'a qu'une composante et la matrice $K_{\alpha\beta}$ se réduit à un scalaire. La position de chaque ion est⁸ r_j :

$$r_j = R_j + u_j , \quad R_j = ja \quad (j \in \mathbb{Z}) . \quad (8.16)$$

et (8.6) devient :

$$U_h = \frac{1}{2} \sum_{j, j' \in \mathbb{Z}} \frac{1}{2!} K(R_j - R_{j'}) (u_j - u_{j'})^2 . \quad (8.17)$$

Même dans ce cas très simple, le calcul effectif de la fonction $K(R)$ à partir des premiers principes est en général très difficile, pour ne pas dire impossible. L'usage est de s'en remettre à une nouvelle approximation *ad hoc*, consistant à ne considérer que des termes entre premiers voisins, ce qui n'introduit ici qu'un seul paramètre $K(\pm a)$, noté⁹ simplement K :

$$K(R_j - R_{j'}) = \begin{cases} K & \text{si } |j - j'| = 1 \\ 0 & \text{autrement} \end{cases} . \quad (8.18)$$

⁵On verra en détail dans la section 8.2 que les équations classiques et quantiques sont formellement identiques.

⁶Il en va donc de même pour tout cristal monoatomique formé d'objets ponctuels ou à symétrie sphérique (scalaires, par opposition à vecteurs, ...)

⁷Les résultats sont inchangés si tous les atomes vibrent suivant une direction commune quelconque.

⁸ r_j n'est pas une longueur, c'est une abscisse $\in \mathbb{R}$.

⁹À ne pas confondre avec un vecteur de $\vec{\mathcal{B}}$!

Il n'existe pas vraiment de justification théorique d'une telle approximation, ce d'autant plus que les interactions élastiques sont à longue portée, surtout en basse dimensionnalité. En fait, la validation de ce schéma minimal – à $d = 1$ comme en dimension supérieure – repose sur ses succès quand on le confronte aux résultats expérimentaux. Ceci étant admis, et comme les deux termes $j' = j \pm 1$ sont en fait égaux entre eux (ce qui fait sauter un facteur $\frac{1}{2}$), le potentiel harmonique se réduit à :

$$U_h = \frac{1}{2} K \sum_{j \in \mathbb{Z}} (u_j - u_{j+1})^2 \quad (8.19)$$

L'interprétation de cette expression est élémentaire ; les équations de Lagrange classiques avec une telle énergie potentielle conduisent à l'équation fondamentale de la Dynamique :

$$M \ddot{u}_j = - \frac{\partial U_h}{\partial u_j} = -K (2u_j - u_{j-1} - u_{j+1}) \iff \ddot{u}_j = -\omega_0^2 (2u_j - u_{j-1} - u_{j+1}) \quad (8.20)$$

où on a posé :

$$K = M \omega_0^2 \quad \omega_0 = \sqrt{\frac{K}{M}} > 0 . \quad (8.21)$$

D'un autre point de vue, la force harmonique de rappel pour l'ion situé en ja résulte des deux forces exercées par ses deux voisins situés en $(j \pm 1)a$; elle vaut $-K(u_j - u_{j-1}) - K(u_j - u_{j+1})$: c'est très exactement ce qui figure au second membre de (8.20). Autrement dit, dans l'approximation en cours, tout se passe comme si on avait une chaîne de boules identiques reliées par des ressorts parfaits ayant tous la même constante de raideur K : on est finalement réduit à trouver les modes propres (classiques) d'un ensemble d'oscillateurs couplés, au nombre de N s'il y a N ions.

Si l'on regarde les choses d'un point de vue quantique, l'interprétation est la même. Le Hamiltonien de la chaîne d'ions est :

$$H = \sum_{j \in \mathbb{Z}} \frac{P_j^2}{2M} + \frac{1}{2} K \sum_{j \in \mathbb{Z}} (u_j - u_{j+1})^2 , \quad (8.22)$$

où P_j est le moment conjugué (impulsion) de l'ion situé au voisinage du point ja . Compte tenu de $[u_i, P_j] = i\hbar \delta_{ij}$, les équations de Heisenberg sont :

$$i\hbar \frac{d}{dt} u_{j,H} = [u_j, H]_H = \sum_{j' \in \mathbb{Z}} \frac{1}{2M} [u_j, P_{j'}^2]_H = \frac{i\hbar}{M} P_{j,H} \iff \dot{u}_{j,H} = \frac{P_{j,H}}{M} , \quad (8.23)$$

et :

$$\begin{aligned} i\hbar \frac{d}{dt} P_{j,H} &= [P_j, H]_H = \frac{1}{2} K \sum_{j' \in \mathbb{Z}} [P_j, (u_{j'} - u_{j'+1})^2]_H = \frac{1}{2} K [P_j, (u_j - u_{j+1})^2 + (u_{j-1} - u_j)^2]_H \\ &= -i\hbar K (u_j - u_{j+1} + u_j - u_{j-1}) \iff \dot{P}_{j,H} = -K (2u_j - u_{j+1} - u_{j-1})_H . \end{aligned} \quad (8.24)$$

En rapprochant (8.23) et (8.24), on trouve :

$$\ddot{u}_{j,H} = -\omega_0^2 (2u_j - u_{j+1} - u_{j-1})_H \quad (8.25)$$

qui n'est autre que la version quantique de (8.20). La seule distinction entre ces deux équations est la nature des objets qui y figurent : des coordonnées u_j au sens classique, ou des observables $u_{j,H}$ en représentation de Heisenberg. Dans un cas comme dans l'autre, on est en face d'un système linéaire de N inconnues, qui se traite toujours de la même façon, puisque, en définitive, tous ces objets commutent toujours entre eux : l'algèbre reste donc "classique".

Finalement, notant x_j soit u_j , soit $u_{j,H}$, il convient de résoudre le système linéaire homogène :

$$\ddot{x}_j = -\omega_0^2 (2x_j - x_{j+1} - x_{j-1}) \quad (8.26)$$

où, à ce stade, la nature précise des inconnues x_j est techniquement sans importance. Pour éviter des difficultés inessentiels, on procède comme d'habitude en considérant dans un premier temps un réseau de taille finie (N atomes) muni des conditions aux limites commodes précisées en temps utile. Dès lors, (8.26) est un système linéaire $N \times N$ du second ordre en temps, dont la résolution exige la connaissance de $2N$ conditions initiales. Dans le cas classique, il s'agit des positions et des vitesses initiales des atomes de la chaîne ; dans le cas quantique, ce sont les $u_{j\text{H}}(t=0) \equiv u_j$ et des $P_{j\text{H}}(t=0) \equiv P_j$, c'est-à-dire les opérateurs au sens de Schrödinger. Quoiqu'il en soit, la méthode de résolution consiste à effectuer une transformée de Fourier en temps, en posant :

$$x_j(t) = \int_{-\infty}^{+\infty} d\omega X_j(\omega) e^{-i\omega t} . \quad (8.27)$$

En vertu de l'hermiticité – ou de la réalité – de x_j , on doit avoir :

$$X_j(-\omega) = X_j(\omega)^\dagger \quad \text{ou} \quad X_j(-\omega) = X_j(\omega)^* . \quad (8.28)$$

Le report dans (8.26) fournit :

$$\int_{-\infty}^{+\infty} d\omega [\omega^2 X_j - \omega_0^2 (2X_j - X_{j+1} - X_{j-1})] e^{-i\omega t} = 0 \quad \forall \omega , \quad (8.29)$$

ce qui impose aux X_j de satisfaire :

$$\omega^2 X_j - \omega_0^2 (2X_j - X_{j+1} - X_{j-1}) = 0 \quad (j \in \mathbb{Z}) . \quad (8.30)$$

Il s'agit d'un système linéaire homogène, qui n'a de solution non triviale que si son déterminant est nul. Cette condition fixe une fois pour toutes les seules valeurs possibles de ω^2 , en tant que zéros d'un polynôme de degré N , $\mathcal{P}_N(\omega^2)$. À ce stade – même si le calcul explicite n'est pas encore fait –, les valeurs possibles de ω^2 sont formellement déterminées¹⁰ par l'équation $\mathcal{P}_N(\omega^2) = 0$.

Par ailleurs, le système homogène (8.30) est une équation aux différences finies, dont les solutions sont de la forme $X_j = (\text{quelque chose})^j$ – et on peut toujours dire que (quelque chose) est une exponentielle. Posons donc :

$$X_j = C e^{ij\phi} ; \quad (8.31)$$

ϕ est une quantité réelle, sinon on aurait d'un côté ($j > 0$) ou de l'autre ($j < 0$) des divergences exponentielles inacceptables physiquement, tandis que C est un facteur indéterminé¹¹. Alors (8.30) devient :

$$[\omega^2 - \omega_0^2 (2 - e^{i\phi} - e^{-i\phi})] e^{ij\phi} = 0 \quad \iff \quad \omega^2 - 4\omega_0^2 \sin^2 \frac{\phi}{2} = 0 . \quad (8.32)$$

ω étant formellement déjà déterminé, ceci fixe ϕ complètement, ou *vice versa*. En pratique¹², il est plus commode d'obtenir d'abord ϕ , en choisissant des conditions aux limites agréables dont la nature précise est sans importance pour les résultats physiques tant que $N \gg 1$; comme d'habitude, on prend des conditions cycliques identifiant X_{N+1} et X_1 :

$$X_{N+1} = X_1 \quad \iff \quad e^{i(N+1)\phi} = e^{i\phi} \quad \iff \quad e^{iN\phi} = 1 . \quad (8.33)$$

Comme déjà vu à plusieurs reprises, ceci quantifie les valeurs de ϕ , un angle défini à 2π près ; ϕ est sans dimension, X_j est relatif au site situé en ja : il est donc naturel d'introduire un vecteur d'onde k , ici à une dimension¹³, k , tel que $\phi = ka$. Alors :

$$e^{iNka} = 1 \quad \iff \quad k = \text{entier} \times \frac{2\pi}{Na} \quad \iff \quad \boxed{k = n \frac{2\pi}{Na} \equiv k_n} \quad (8.34)$$

¹⁰Il resterait à prouver que ce polynôme a non seulement tous ses zéros réels mais encore qu'ils sont tous positifs ou nuls. La réalité est assurée par le fait que \mathcal{P}_N est associé à une matrice symétrique (hermitique réelle). La positivité peut se démontrer soit par récurrence en N , soit en utilisant des propriétés générales (pour une chaîne ouverte – non refermée à la Born - von Karman –, \mathcal{P}_N est un polynôme bien connu, appelé polynôme de Gegenbauer).

¹¹Le système (8.30) étant *homogène*, les X_j s'en trouvent fixés à un facteur multiplicatif près

¹²Les deux procédures sont strictement équivalentes pour la raison suivante. Appliquer les conditions cycliques, c'est finalement jouer implicitement avec les opérateurs de rotation d'angle entier $\frac{2\pi}{N}$, qui représentent des symétries de la chaîne refermée par l'identification $N + 1 = 1$. Ces opérateurs forment un groupe abélien et, ce faisant, on diagonalise complètement *ipso facto* la matrice associée au système linéaire (8.30).

¹³ k est donc défini à $\frac{2\pi}{a}$ près.

Puisqu'il y a N modes propres au total, n prend N valeurs, par exemple¹⁴ $1, 2, \dots, N$; les valeurs de ω sont alors explicites :

$$\omega = \pm\omega(k) \ , \quad \omega(k) = 2\omega_0 \left| \sin \frac{ka}{2} \right| \quad k \in \{k_n\}_n \quad (8.35)$$

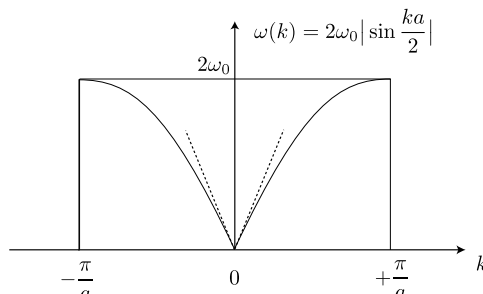


Figure 8.1: Relation de dispersion pour un réseau unidimensionnel avec un couplage entre premiers voisins. La pente des lignes en pointillés est la vitesse $c \stackrel{\text{d\'ef}}{=} a\omega_0$.

Au total, il n'y a bien que N valeurs distinctes pour k , ce qui donne également N valeurs distinctes pour les deux branches $\pm\omega(k)$, eu égard à la parité de $\omega(k)$ (pour chaque valeur de k , la dégénérescence est égale à 2). L'usage ordinaire¹⁵ consiste à faire varier k symétriquement autour de l'origine, essentiellement entre $\pm \frac{N}{2} \frac{2\pi}{Na} = \pm \frac{\pi}{a}$. La variation de la fonction $\omega(k)$ avec k constitue la loi de dispersion pour les modes de vibration ; pour un k donné, les X_j sont donnés par¹⁶ :

$$X_j(\omega(k)) = C e^{ikja} \ . \quad (8.36)$$

Il est instructif d'examiner en détail la collection des déplacements u_j résultant d'un état initial précis. Il suffit pour cela d'inverser la relation de Fourier (8.27) ; les valeurs possibles de ω sont en fait discrètes¹⁷, de sorte que l'intégrale dans (8.27) est à ce stade une somme \sum_k . Compte tenu des deux branches $\pm\omega(k)$ de la loi de dispersion, on peut toujours écrire $x_j(t)$ sous la forme :

$$x_j(t) = \sum_k X_j(\omega(k)) (\alpha_k + e^{i\omega(k)t} + \alpha_k - e^{-i\omega(k)t}) \ , \quad (8.37)$$

soit :

$$x_j(t) = \sum_k C e^{ikja} (\alpha_k + e^{i\omega(k)t} + \alpha_k - e^{-i\omega(k)t}) \ , \quad (8.38)$$

où les $\alpha_{k\pm}$ sont deux constantes d'intégration, à k fixé. Au total, il faut trouver $2N$ constantes $C\alpha_{k\pm}$, ce qui est possible en utilisant les $2N$ conditions initiales¹⁸. Raisonnons d'abord dans le cadre classique ; les x_j sont les u_j et selon (8.38) :

$$u_j(t) = \sum_k C e^{ikja} (\alpha_k + e^{i\omega(k)t} + \alpha_k - e^{-i\omega(k)t}) \ , \quad (8.39)$$

et les moments sont ($P_j \equiv M\dot{r}_j = M\dot{u}_j$) :

$$P_j(t) = M \sum_k i\omega(k) C e^{ikja} (\alpha_k + e^{i\omega(k)t} - \alpha_k - e^{-i\omega(k)t}) \ . \quad (8.40)$$

Les $2N$ conditions initiales s'écrivent :

$$u_j(0) = \sum_{k \in [-\frac{\pi}{a}, +\frac{\pi}{a}]} C e^{ikja} (\alpha_k + + \alpha_k -) \ , \quad (8.41)$$

¹⁴On pourrait tout autant prendre $n = -\frac{N}{2}, -\frac{N}{2} + 1, \dots, +\frac{N}{2}$, en supposant N pair – hypothèse non-restrictive dans la limite $N \gg 1$.

¹⁵ $\omega(k)$ est de toute façon une fonction périodique de k , avec la (plus petite) période $\frac{2\pi}{a}$.

¹⁶La constante multiplicative C vient du fait que le système (8.30) est homogène. C peut dépendre de k , mais cette dépendance est sous-entendue par simplicité.

¹⁷Ce sont les N solutions de $\mathcal{P}_N(\omega^2) = 0$.

¹⁸qui, dans le cas quantique, contiennent les opérateurs au sens de Schrödinger.

$$P_j(0) = M \sum_{k \in [-\frac{\pi}{a}, +\frac{\pi}{a}]} i\omega(k) C e^{ikja} (\alpha_{k+} - \alpha_{k-}) . \quad (8.42)$$

Pour simplifier, on suppose toutes les vitesses initiales nulles ; (8.42) dit alors que $\alpha_{k+} = \alpha_{k-} = \alpha_k$ ce qui donne pour (8.39) :

$$u_j(t) = \sum_{k \in [-\frac{\pi}{a}, +\frac{\pi}{a}]} 2\alpha_k C e^{ikja} \cos \omega(k)t , \quad (8.43)$$

avec maintenant :

$$u_j(0) = \sum_{k \in [-\frac{\pi}{a}, +\frac{\pi}{a}]} 2\alpha_k C e^{ikja} . \quad (8.44)$$

Cette dernière relation s'inverse facilement en utilisant¹⁹ :

$$\sum_{j=1}^N e^{i(k-k')ja} = N \delta_{kk'} \quad (N \gg 1) , \quad (8.45)$$

où, tant que N est fini, $\delta_{kk'}$ est le *symbole* de Kronecker. On trouve ainsi :

$$2N\alpha_k C = \sum_{j=1}^N u_j(0) e^{-ikja} . \quad (8.46)$$

Finalement, les déplacements pour une chaîne classique abandonnée hors de son point d'équilibre sans vitesse initiale pour les atomes :

$$u_j(t) = \frac{1}{N} \sum_{j'=1}^N u_{j'}(0) \sum_{k \in [-\frac{\pi}{a}, +\frac{\pi}{a}]} e^{ik(j-j')a} \cos \omega(k)t , \quad (8.47)$$

dans la limite $N \gg 1$, la somme peut être remplacée par une intégrale :

$$\sum_{k \in [-\frac{\pi}{a}, +\frac{\pi}{a}]} \rightarrow \frac{L}{2\pi} \int_{-\frac{\pi}{a}}^{+\frac{\pi}{a}} dk , \quad (8.48)$$

et comme $L = Na$, on obtient :

$$u_j(t) = \sum_{j'=1}^N u_{j'}(0) \frac{a}{2\pi} \int_{-\frac{\pi}{a}}^{+\frac{\pi}{a}} dk e^{ik(j-j')a} \cos \omega(k)t . \quad (8.49)$$

Compte tenu de la dépendance de ω en k (voir (8.35)), l'intégrale porte sur des exponentielles du genre $e^{ik(j-j')a} e^{\pm 2i\omega_0 t \sin(ka/2)}$; le changement de variable $ka/2 = \phi$ fait apparaître des intégrales $e^{2i(j-j')\phi} e^{\pm 2i\omega_0 t \sin \phi}$, caractéristiques de la définition des fonctions de Bessel²⁰. Au total, $u_j(t)$ est une combinaison linéaire de ces fonctions spéciales :

$$u_j(t) = \sum_{j'=1}^N u_{j'}(0) J_{2(j-j')}(2\omega_0 t) \quad (8.51)$$

En particulier, si un seul atome est déplacé au départ de u_0 , celui en j_0 , on a $u_j(0) = u_0 \delta_{jj_0}$ et :

$$u_j(t) = u_0 J_{2(j-j_0)}(2\omega_0 t) ; \quad (8.52)$$

aux grands temps $t \gg \omega_0^{-1}$, on a à peu près :

$$u_j(t) \sim \frac{u_0}{\sqrt{\pi\omega_0 t}} \cos \left[2\omega_0 t - (j - j_0) \frac{\pi}{2} - \frac{\pi}{4} \right] ; \quad (8.53)$$

¹⁹Cette relation a déjà été utilisée à propos de l'approximation des liaisons fortes et aussi pour établir la condition de Bragg.

²⁰La fonction de Bessel ordinaire J_n , n entier, est l'intégrale :

$$J_n(x) \stackrel{\text{déf}}{=} \frac{1}{2\pi} \int_0^{2\pi} d\phi e^{-in\phi} e^{ix \sin \phi} \quad (n \in \mathbb{Z}) . \quad (8.50)$$

Pour $x \gg 1$, la méthode du col donne $J_n(x) \sim \sqrt{\frac{2}{\pi x}} \cos(x - n\frac{\pi}{2} - \frac{\pi}{4})$.

à t grand, le champ des déplacements a en gros l'allure d'une sinusoïde, avec une amplitude $\sim t^{-\frac{1}{2}}$. Clairement, le "paquet d'ondes" de la déformation initiale s'élargit et se dilue le long de la chaîne ; notamment, le déplacement initial en j_0 s'évanouit et tend vers zéro comme $t^{-\frac{1}{2}}$.

En $t = 0$, le second membre de (8.47) est²¹ :

$$\frac{1}{N} \sum_{j'} u_{j'}(0) \sum_k e^{ik(j-j')a} = \frac{1}{N} \sum_{j'} u_{j'}(0) N \delta_{jj'} = u_j(0) , \quad (8.54)$$

comme il se doit. Par ailleurs, si on déplace initialement chaque atome de la même quantité $u(0)$, (8.47) donne :

$$u_j(t) = \frac{u(0)}{N} \sum_{j'=1}^N \sum_k e^{ik(j-j')a} \cos \omega(k)t = \frac{u(0)}{N} \sum_k e^{ikja} N \delta_{ka0} \cos \omega(k)t = u(0) . \quad (8.55)$$

Dans ces circonstances, chaque atome garde sa position initiale à tout temps, ce qui est bien normal : un tel état de départ est en fait une translation en bloc du réseau, sans déformation, n'induisant aucune force de rappel et ne provoquant donc aucun mouvement ultérieur. Ceci permet de retrouver le fait que le mode $k = 0$ est en réalité associé à un déplacement (translation) en bloc.

D'une façon générale, les modes de petit k ($ka \ll 1$) correspondent à des déformations dynamiques $u_j(t)$ variant lentement en fonction de l'entier j . Autrement dit, les déformations correspondantes ont une variation spatiale lente à l'échelle du pas du réseau (si les déplacements initiaux $u_{j'}(0)$ possèdent cette propriété, il en ira évidemment de même à tout temps). Dès lors, ce sont les modes de grande longueur d'onde qui sont prépondérants : on ne peut plus bien sûr se contenter strictement de $k = 0$, mais, à l'inverse, le *voisinage* de $k = 0$ suffit. Alors, dans (8.49), on peut remplacer $\omega(k)$ par son développement limité (voir (8.35)) :

$$\omega(k) = 2\omega_0 \frac{|k|a}{2} + \dots \equiv |k|c + \dots \quad (8.56)$$

où a été introduite une certaine vitesse c :

$$c \stackrel{\text{déf}}{=} a \omega_0 , \quad (8.57)$$

qui n'est autre que la vitesse du son (vitesse de groupe et vitesse de phase coïncident dans la limite du solide continu²²). De la sorte, en posant $x = ja$, $x' = j'a$, $u(x = ja, t) = u_j(t)$, (8.49) devient :

$$u(x, t) \simeq \frac{a}{2\pi} \int_{-\frac{\pi}{a}}^{+\frac{\pi}{a}} dk \sum_{j'=1}^N u(x', 0) e^{ik(x-x')} \cos kct . \quad (8.58)$$

Dans cette optique à grande échelle, tout se passe comme si le pas a du réseau était un "infiniment petit". Il est donc licite de remplacer $a \sum_{j'}$ par $\int dx'$; (8.58) a donc le même sens que²³ :

$$u(x, t) \simeq \int_{-\frac{\pi}{a}}^{+\frac{\pi}{a}} \frac{dk}{2\pi} \int dx' u(x', 0) e^{ik(x-x')} \cos kct . \quad (8.60)$$

Il s'agit de détails ; l'une ou l'autre de ces dernières expressions montre que la déformation initiale se propage le long du réseau sous la forme d'un paquet d'ondes dispersé suivant la loi (8.35), dont vitesse de phase

²¹Il en va de même si on fait $t = 0$ au second membre de (8.49).

²²Dans la limite $ka \rightarrow 0$, la nature discrète du réseau s'efface, toutes les longueurs d'onde du champ des déplacements étant très grandes devant le pas a du réseau.

²³La linéarisation en k de la loi de dispersion – et la limite $a \rightarrow 0$ – rend élémentaire l'intégration en k de (8.60), qui fait apparaître $\delta(x - x' \pm ct)$, d'où :

$$u(x, t) = \frac{1}{2} [u(x - ct, t = 0) + u(x + ct, t = 0)] ; \quad (8.59)$$

cette expression est bien la solution de l'équation de propagation $(c^{-2}\partial_{t^2} - \partial_{x^2})u = 0$ et représente la dynamique d'une onde dans un milieu non dispersif (toutes les fréquences avancent à la même vitesse) ; dans ces conditions, tout paquet d'ondes se propage en bloc, sans déformation. Retomber sur l'équation de propagation ne doit pas surprendre : le deuxième membre de (8.26) n'est rien d'autre qu'une dérivée seconde *discrète* par rapport à l'espace.

et vitesse de groupe coïncident dans la limite des grandes longueurs d'onde. Une branche de vibration qui tend vers zéro linéairement en k est appelée branche acoustique²⁴ :

$$\boxed{\lim_{k \rightarrow 0} \omega(k) = 0 \quad \text{et} \quad \lim_{k \rightarrow 0} \frac{\omega(k)}{k} = c \equiv C^{\text{ste}} \iff \text{branche acoustique}} \quad (8.61)$$

D'un autre côté, si les déplacements initiaux varient vite avec j , il faut conserver la loi de dispersion (8.35) dans toute son intégrité. La vitesse de groupe associée à la loi de dispersion (8.35) est :

$$v_g(k) = \frac{d\omega(k)}{dk} = a\omega_0 \cos \frac{ka}{2} . \quad (8.62)$$

Elle s'annule lorsque $ka = \pi$, c'est-à-dire en bord de zone. Un paquet d'ondes ayant une amplitude maximum en $k = k_0 = \frac{\pi}{a}$ ne se propage donc pas et fait du sur-place. Pour comprendre ce que ceci signifie en termes de déplacements, choisissons un état initial où :

$$u_j(t = 0) = (-1)^j u_0 . \quad (8.63)$$

Ceci représente un état où les sites pairs sont tous écartés d'une même quantité u_0 , tous les sites impairs étant écartés de $-u_0$; visiblement, c'est bien un cas où le champ des écarts initiaux varie vite à l'échelle du réseau (il est nul en moyenne sur deux mailles adjacentes). Le report de (8.63) dans (8.49) fait apparaître la somme :

$$\sum_{j'=1}^N e^{-ij'ka} u_{j'}(0) = u_0 \sum_{j'=1}^N e^{-ij'ka} (-1)^{j'} = u_0 \sum_{j'=1}^N e^{ij'(\pi-ka)} = u_0 N \delta_{ka\pi} . \quad (8.64)$$

Insérée dans une intégrale sur k , $N \delta_{ka\pi}$ se comporte comme une fonction de Dirac, $C \delta(ka - \pi)$ – soit $\frac{C}{a} \delta(k - \frac{\pi}{a})$; en outre, il est facile de voir²⁵ que $C = 2\pi$. L'expression de $u_j(t)$ selon (8.49) est alors :

$$u_j(t) = u_0 \int_{-\frac{\pi}{a}}^{+\frac{\pi}{a}} dk e^{ijka} \delta(k - \frac{\pi}{a}) \cos \omega(k)t = u_0 e^{ij\pi} \cos 2\omega_0 t , \quad (8.65)$$

où l'équation (8.35) a été utilisée ($\omega(k = \frac{\pi}{a}) = 2\omega_0$). Finalement :

$$u_j(t) = u_j(0) \cos 2\omega_0 t . \quad (8.66)$$

Tous les sites vibrent donc à la même fréquence $2\omega_0$ et deux sites voisins restent en opposition de phase, tout comme ils l'étaient²⁶ à $t = 0$. Si l'on regarde le "paquet d'ondes" réduit à deux tels sites voisins, on voit en effet qu'il reste en place : il ne bouge pas et ne se déforme pas.

Le cas quantique se traite de la même façon ; en prenant les valeurs moyennes des opérateurs à la Heisenberg sur un état initial des atomes où toutes les vitesses sont nulles, on obtient exactement les mêmes expressions, $u_j(t)$ donné par (8.49) étant alors la valeur moyenne (quantique) du déplacement à l'instant t quand les noyaux sont immobiles (en moyenne) dans l'état initial.

8.3 Modes normaux d'un réseau unidimensionnel à deux atomes par maille

On a vu dans la section 8.2 que les modes $k \approx 0$ d'une chaîne monoatomique correspondent à un champ de déplacement de grande longueur d'onde. À la limite où $k = 0$, ceci donne un mode de translation en bloc d'énergie nulle : il n'est donc pas surprenant que la pulsation $\omega(k)$ s'annule en $k = 0$.

²⁴L'origine de la terminologie sera claire par la suite.

²⁵Le pas en k est $\delta k = \frac{2\pi}{a}$ et $L = Na$.

²⁶Ainsi, la symétrie "locale" présente dans l'état de départ est préservée à tout instant ultérieur. Ce n'est pas surprenant : aucun terme dans le Hamiltonien n'est susceptible de la briser.

Par ailleurs, rien n'interdit de considérer le réseau précédent comme un réseau de maille $a' = 2a$, avec une base à deux atomes séparés de la même distance a , donc constituée de molécules diatomiques (“dimères”) ; dans cette optique, le plus petit vecteur du réseau réciproque est de norme²⁷ $K = \frac{\pi}{2a}$. En suivant le même procédé que pour les électrons dans un potentiel périodique, on peut récrire les résultats de la section 8.2 dans un schéma en zones réduites, où k varie maintenant entre $\pm \frac{\pi}{2a}$. La première branche qui apparaît, notée désormais ω_{ac} (acoustique) est donc identique à (8.35), mais est maintenant bornée entre $\pm \frac{\pi}{2a}$:

$$\omega_{ac}(k) = 2\omega_0 \left| \sin \frac{ka}{2} \right|, \quad k \in \left[-\frac{\pi}{2a}, +\frac{\pi}{2a} \right]. \quad (8.67)$$

L'arc de (8.35) (*resp.* son symétrique) initialement situé entre $\frac{\pi}{a}$ et $\frac{\pi}{2a}$ (*resp.* $-\frac{\pi}{a}$ et $-\frac{\pi}{2a}$) doit être décalé vers la gauche (*resp.* vers la droite) d'un multiple de $K = \frac{\pi}{2a}$, de façon à se retrouver entre $\pm \frac{\pi}{2a}$: on voit tout de suite qu'il faut le translater de $2\frac{\pi}{2a} \equiv 2K$; notant ω_{opt} cette branche, appelée branche optique²⁸, on a :

$$\omega_{opt}(k) = 2\omega_0 \left| \sin \frac{(k \mp \frac{\pi}{a})a}{2} \right| \equiv 2\omega_0 \left| \cos \frac{ka}{2} \right|, \quad k \in \left[-\frac{\pi}{2a}, +\frac{\pi}{2a} \right]. \quad (8.68)$$

Dans le schéma en zones réduites, la pulsation est maintenant *bivaluée*, les deux branches se rejoignant en $\pm \frac{\pi}{2a}$. Notamment, il y a deux modes $k = 0$. Celui de la branche acoustique représente toujours la translation en bloc et a toujours une énergie nulle ; en revanche, celui de la branche optique correspond à une vibration en opposition de phase des deux membres du dimère, tous les dimères vibrant en phase (voir (8.66)) ; son énergie n'est pas nulle, évidemment. Un mode optique se différencie notamment d'un mode acoustique par le fait que sa pulsation ne tend pas vers zéro aux grandes longueurs d'onde puisqu'il représente, dans le schéma en zones réduites, le mode où la vitesse de groupe s'annule et où deux atomes voisins vibrent en opposition de phase.

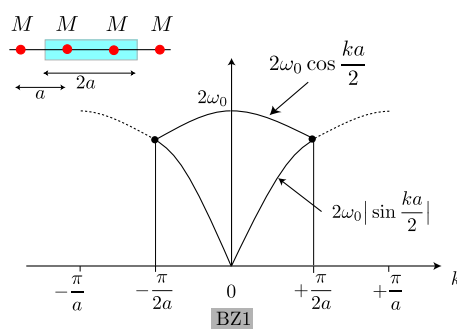


Figure 8.2: Relation de dispersion (8.35) dans le schéma en zones réduites pour le réseau de maille double avec deux atomes *identiques* par maille.

La description qui précède peut paraître artificielle lorsque les deux partenaires du dimère sont identiques, mais elle permet une transition vers le cas où ils sont, de fait, différents, une situation qui correspond à un vrai cristal unidimensionnel avec deux atomes par maille. Dans l'immédiat, on va se borner à différencier un atome sur deux du réseau de pas a (par leur masse, par exemple). Le traitement sera finalement étendu à une base diatomique où les partenaires sont à une distance d l'un de l'autre (en pratique, $d < a$).

On peut deviner ce qui va se passer, en raisonnant dans un cas limite. Puisque les atomes sont différents, les fréquences propres vont être différentes à l'intérieur d'un dimère (dont les deux masses sont différentes) ou entre deux dimères distincts (qui ont la même masse). À titre d'exemple, si les deux atomes sont seulement isotopes l'un de l'autre (masses M_1 et M_2 , $M_1 < M_2$), le potentiel harmonique – qui ne dépend que de la distribution électronique périphérique – sera le même sur tous les sites, donc la constante de raideur K sera la même d'un bout à l'autre du réseau. En revanche, on aura maintenant deux pulsations propres :

$$\omega_0 = \sqrt{\frac{K}{M_1}} \quad \Omega_0 = \sqrt{\frac{K}{M_2}} < \omega_0. \quad (8.69)$$

D'une façon générale, deux branches vont donc apparaître si deux masses sont en présence.

²⁷Ne pas confondre avec la constante de raideur !

²⁸À nouveau, l'origine de la terminologie sera claire par la suite.

Il est facile de prévoir l'allure des lois de dispersion dans le cas extrême où $M_1 \ll M_2$. En effet, dans ce cas, chaque dimère a une masse à peu près égale à M_2 ; le réseau de dimères a un pas égal à $2a$. Donc, en considérant chacun d'entre eux comme une grosse particule, on voit qu'il existe des modes de type acoustique, la loi étant approximativement²⁹ donnée par :

$$\omega_{\text{ac}}(k) \simeq 2 \frac{\Omega_0}{\sqrt{2}} \left| \sin \frac{k(2a)}{2} \right|, \quad k \in \left[-\frac{\pi}{2a}, +\frac{\pi}{2a} \right] \quad (M_1 \ll M_2); \quad (8.70)$$

c'est bien la pulsation $\frac{\Omega_0}{\sqrt{2}}$ qui doit être insérée dans (8.68), puisque chaque grosse particule a une masse à peu près égale à M_2 et que chacune des deux grosses masses est reliée à la petite par un ressort de constante K . Le ressort effectif³⁰ entre les deux grosses masses a donc une constante K_{eff} telle que $K_{\text{eff}}^{-1} = K^{-1} + K^{-1} = 2K^{-1}$, soit $K_{\text{eff}} = \frac{K}{2}$. La fréquence d'oscillation est donc $\Omega_{\text{eff}}^2 = \frac{K}{2} M_2$, soit $\Omega_{\text{eff}} = \frac{\Omega_0}{\sqrt{2}}$. La branche part de 0 en $k = 0$ et atteint sa valeur maximale $\sqrt{2}\Omega_0$ en $k = \frac{\pi}{2a}$.

Ceci définit une première classe de modes de vibration. Il existe aussi un mode de vibration intradimère, représentant cette fois une oscillation rapide de la petite masse M_1 rappelée par la grande masse M_2 quasiment immobile. En considérant un seul dimère, cette oscillation se fait strictement à la pulsation³¹ $\sqrt{2}\omega_0 \gg \Omega_0$, et a donc une grande énergie³². En raison du couplage mécanique entre les petites masses³³, l'énergie correspondante dépendra de k , mais on s'attend à ce qu'elle soit peu dispersée autour de ω_0 – elle est donc relativement plate – et, en tout cas ressemble à un mode optique, tendant vers une valeur finie en $k = 0$.

En définitive, dans le cas de deux masses très différentes, $M_1 \ll M_2$, on attend deux branches, l'une optique, l'autre acoustique, qui ne se rejoignent plus en $k = \frac{\pi}{2a}$: un *gap* s'est ouvert en bord de zone par comparaison au cas $M_1 = M_2$ traité comme ci-dessus (deux atomes identiques par maille) – tout comme dans le cas des électrons presque libres.

Le calcul exact détaillé est facile, avec M_1 et M_2 quelconques (mais toujours $M_1 < M_2$ pour fixer les idées). On repart des équations du mouvement (8.20) adaptées au cas analysé maintenant ; supposons les petites masses M_1 situées sur les sites impairs, les grandes masses M_2 sur les sites pairs. Alors :

$$M_1 \ddot{u}_{2j+1} = - \frac{\partial U_{\text{h}}}{\partial u_{2j+1}} = -K (2u_{2j+1} - u_{2j} - u_{2j+2}), \quad (8.71)$$

et :

$$M_2 \ddot{u}_{2j} = - \frac{\partial U_{\text{h}}}{\partial u_{2j}} = -K (2u_{2j} - u_{2j-1} - u_{2j+1}). \quad (8.72)$$

Avec la même convention que précédemment (x_j est soit un déplacement classique $u(t)$, soit la représentation de Heisenberg $u_{\text{H}}(t)$), on obtient le double système (c'est la généralisation du système unique (8.26)) :

$$\ddot{x}_{2j+1} = -\omega_0^2 (2x_{2j+1} - x_{2j} - x_{2j+2}), \quad (8.73)$$

et :

$$\ddot{x}_{2j} = -\Omega_0^2 (2x_{2j} - x_{2j-1} - x_{2j+1}). \quad (8.74)$$

L'analyse de Fourier donne maintenant :

$$\omega^2 X_{2j+1} - \omega_0^2 (2X_{2j+1} - X_{2j} - X_{2j+2}) = 0, \quad (8.75)$$

et :

$$\omega^2 X_{2j} - \Omega_0^2 (2X_{2j} - X_{2j-1} - X_{2j+1}) = 0. \quad (8.76)$$

²⁹On va évidemment trouver des corrections, puisque le rapport $\frac{M_2}{M_1}$ n'est pas infini. Noter que $M_1 \ll M_2$ est équivalent à $\Omega_0 \ll \omega_0$.

³⁰Quand on met des ressorts en série, ce sont les inverses des constantes de raideur qui s'additionnent (au total, c'est plus "mou") ; au contraire, si les ressorts sont mis en parallèle, c'est plus "dur" et les constantes de raideur s'additionnent. Une chaîne 1-d infinie a une compressibilité infinie (elle est infiniment molle).

³¹Pour une masse M_1 liée de part et d'autre à deux masses inertes ($u_{j\pm 1} = 0$), l'équation est $M_1 \ddot{u}_j = -2K u_j$, d'où le facteur $\sqrt{2}$.

³²La relation entre haute fréquence et haute énergie tient tant dans le cas classique que dans le cas quantique. Dans le cas quantique, elle vient de $E = \hbar\omega$, évidemment. Dans le cas classique, elle vient du fait que l'énergie mécanique d'un oscillateur écarté de x_0 de sa position initiale est égale à $\frac{1}{2}m\omega_0^2 x_0^2$, d'autant plus grande que ω_0 est grand.

³³Comme les grosses masses M_2 ne sont pas totalement inertes, il existe évidemment un couplage résiduel entre les petites masses M_1 , médié par les masses M_2 .

Tout naturellement, les petites masses sont couplées aux grandes masses et, à nouveau, il s'agit d'équations aux différences finies dont les solutions sont les puissances d'exponentielles complexes. Les X sont des amplitudes de Fourier des déplacements ; comme les déplacements sont d'autant plus faibles que l'inertie est grande, et qu'il y a deux masses, il convient d'introduire deux constantes de proportionnalité C_1 et C_2 pour les sites impairs et pairs respectivement. On pose donc :

$$X_{2j+1} = C_1 e^{i(2j+1)ka} , \quad X_{2j} = C_2 e^{i2jka} . \quad (8.77)$$

Le report dans (8.75) et (8.76) donne le système :

$$(\omega^2 - 2\omega_0^2) C_1 + 2\omega_0^2 \cos ka C_2 = 0 , \quad (8.78)$$

$$2\Omega_0^2 \cos ka C_1 + (\omega^2 - 2\Omega_0^2) C_2 = 0 . \quad (8.79)$$

La solution est non-triviale si et seulement si le déterminant est nul, soit :

$$\omega^4 - 2(\omega_0^2 + \Omega_0^2)\omega^2 + 4\omega_0^2\Omega_0^2 \sin^2 ka = 0 . \quad (8.80)$$

Les solutions en ω^2 sont :

$$\omega^2 = \omega_0^2 + \Omega_0^2 \pm \sqrt{(\omega_0^2 + \Omega_0^2)^2 - 4\omega_0^2\Omega_0^2 \sin^2 ka} . \quad (8.81)$$

La quantité sous la racine carrée est aussi $(\omega_0^2 - \Omega_0^2)^2 + 4\omega_0^2\Omega_0^2 \cos^2 ka$: elle est bien positive). Par ailleurs, ω^2 est visiblement positive quel que soit le signe au second membre. L'équation (8.81) donne notamment les deux branches positives :

$$\omega_{\pm}(k) = \left[\omega_0^2 + \Omega_0^2 \pm \sqrt{(\omega_0^2 + \Omega_0^2)^2 - 4\omega_0^2\Omega_0^2 \sin^2 ka} \right]^{1/2} , \quad (8.82)$$

ou encore :

$$\boxed{\omega_{\pm}(k) = \left[\omega_0^2 + \Omega_0^2 \pm \sqrt{(\omega_0^2 - \Omega_0^2)^2 + 4\omega_0^2\Omega_0^2 \cos^2 ka} \right]^{1/2}} \quad (8.83)$$

En $k = 0$, elles valent $[\omega_0^2 + \Omega_0^2 \pm (\omega_0^2 + \Omega_0^2)]^{1/2} = 0, \sqrt{2(\omega_0^2 + \Omega_0^2)}$; plus précisément, un développement limité donne ($ka \simeq 0$) :

$$\omega_-(k) \simeq \sqrt{2 \frac{\omega_0^2 \Omega_0^2}{\omega_0^2 + \Omega_0^2}} |k|a , \quad \omega_+(k) \simeq \sqrt{2(\omega_0^2 + \Omega_0^2)} \left[1 - \frac{\omega_0^2 \Omega_0^2}{2(\omega_0^2 + \Omega_0^2)^2} k^2 a^2 \right] \quad (8.84)$$

ω_- est la branche acoustique (elle s'annule linéairement en k), ω_+ est la branche optique, parabolique près de $k = 0$.

Pour examiner ce qui se passe près du bord de zone, $k \simeq \frac{\pi}{2a}$, on part de :

$$\omega_{\pm} = \omega_0^2 + \Omega_0^2 \pm \sqrt{(\omega_0^2 - \Omega_0^2)^2 + 4\omega_0^2\Omega_0^2 \sin^2(\frac{\pi}{2} - ka)} ; \quad (8.85)$$

avec $\omega_0 > \Omega_0$; le développement limité donne :

$$\omega_+(k) \simeq \sqrt{2} \omega_0 \left[1 + \frac{\Omega_0^2}{2(\omega_0^2 - \Omega_0^2)} \left(\frac{\pi}{2} - ka\right)^2 \right] , \quad (8.86)$$

$$\omega_-(k) \simeq \sqrt{2} \Omega_0 \left[1 - \frac{\omega_0^2}{2(\omega_0^2 - \Omega_0^2)} \left(\frac{\pi}{2} - ka\right)^2 \right] . \quad (8.87)$$

Les deux branches sont donc à peu près paraboliques en bord de zone et séparées par le *gap* :

$$\boxed{(\omega_+ - \omega_-)_{\frac{\pi}{2a}} = \sqrt{2}(\omega_0 - \Omega_0)} \quad (8.88)$$

qui s'annule bien si on en revient à deux masses identiques. À l'inverse, si $M_1 \ll M_2$ (soit $\omega_0 \gg \Omega_0$), (8.82) montre que la branche optique est peu dispersée ; en effet, on trouve alors :

$$\omega_+(k) \simeq \sqrt{2\omega_0^2} \left[1 + \frac{\Omega_0^2}{\omega_0^2} \cos^2 ka \right] . \quad (8.89)$$

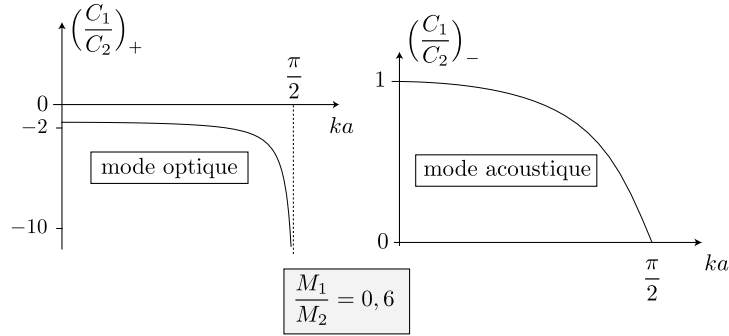


Figure 8.3: Représentation schématique des déplacements à un instant donné pour un réseau à deux atomes par maille.

Comme $\frac{\Omega_0}{\omega_0} \ll 1$, la variation induite par $\cos^2 ka$ est faible : la bande optique est presque plate.

Une fois obtenues les valeurs propres, il est possible d'analyser les mouvements respectifs des différents atomes dans chacun des modes obtenus (voir fig. ??). Il suffit pour cela d'examiner le rapport $\frac{C_1}{C_2}$, donné (par exemple) par (8.78) :

$$\left(\frac{C_1}{C_2}\right)_{\pm} = -\frac{2\omega_0^2 \cos ka}{\omega_{\pm}^2 - 2\omega_0^2} = \frac{2 \cos ka}{1 - \mu \mp \sqrt{(1 + \mu)^2 - 4\mu \sin^2 ka}} , \quad \mu = \frac{\Omega_0^2}{\omega_0^2} = \frac{M_1}{M_2} \leq 1 . \quad (8.90)$$

En utilisant les développements limités de $\omega_{\pm}(k)$, il est facile de voir que :

- $k \rightarrow 0$

$$\text{mode optique : } \left(\frac{C_1}{C_2}\right)_{+} \simeq -\frac{1}{\mu} \left[1 - \frac{1 - \mu}{2(1 + \mu)} (ka)^2\right] = -\frac{M_2}{M_1} [1 + \mathcal{O}(k^2 a^2)] . \quad (8.91)$$

Les deux atomes vibrent en opposition de phase – ce que l'on attend d'un mode optique, avec une amplitude inversement proportionnelle à leur masse ; on a $C_1 > C_2$ si $M_1 < M_2$ (inertie). Pour le mode acoustique ω_- , on trouve tout naturellement que les deux amplitudes sont du même ordre de grandeur par rapport au rapport des masses :

$$\text{mode acoustique : } \left(\frac{C_1}{C_2}\right)_{-} \simeq 1 - \frac{M_2 - M_1}{2(M_1 + M_2)} (ka)^2 . \quad (8.92)$$

- $k \rightarrow \frac{\pi}{2a}$

$\omega_{+}^2 - 2\omega_0^2$ s'annule quadratiquement par rapport à $\epsilon \stackrel{\text{déf}}{=} k - \frac{\pi}{2a} < 0$. Au contraire, le numérateur de (8.90) s'annule linéairement. Il en résulte que le rapport $(\frac{C_1}{C_2})_{+}$ diverge ; très précisément, un développement limité donne :

$$\text{mode optique : } \left(\frac{C_1}{C_2}\right)_{+} \simeq \left(\frac{M_2}{M_1} - 1\right) \left(k - \frac{\pi}{2a}\right)^{-1} \rightarrow -\infty . \quad (8.93)$$

Dans le mode optique, près du bord de zone, l'atome lourd est presque immobile et seul l'atome léger vibre avec une amplitude finie (comme le montre le signe $-$, le mouvement est toujours en opposition de phase). C'est un mode de haute énergie, correspondant pratiquement à un mode de vibration intramoléculaire.

Pour ω_- (mode acoustique), il en va tout autrement. Maintenant, le dénominateur a une limite finie. Il en résulte :

$$\text{mode acoustique : } \left(\frac{C_1}{C_2}\right)_{-} \simeq \frac{M_2}{M_2 - M_1} \left(\frac{\pi}{2} - ka\right) > 0 . \quad (8.94)$$

Cette fois, c'est l'atome massif qui a une grande amplitude, cependant que, relativement, l'atome léger est à peu près fixe ; l'énergie de vibration est nettement plus petite que pour le mode optique.

La figure 8.4 résume ces derniers résultats.

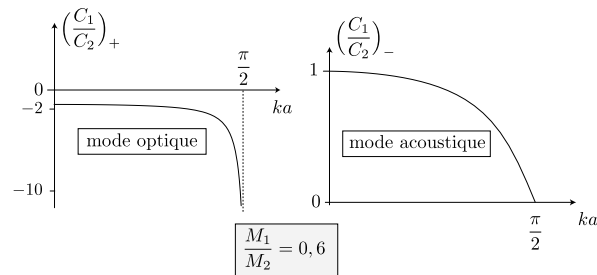


Figure 8.4: Rapport des amplitudes de vibration $\frac{C_1}{C_2}$ des deux masses M_1 et M_2 (voir (8.90)) ; les courbes sont tracées pour $\mu \stackrel{\text{déf}}{=} \frac{M_1}{M_2} = 0,6$.

Le cas précédent a été analysé dans la lignée du réseau monoatomique, en alternant les masses et en définissant donc un réseau de maille deux fois plus grande avec une base. Comme dernier exemple, on considère un réseau de paramètre³⁴ a , avec une base diatomique constituée de deux atomes de même masse M séparés (à l'équilibre) de la distance $d < a$ (molécule homonucléaire) (voir fig. 8.5).

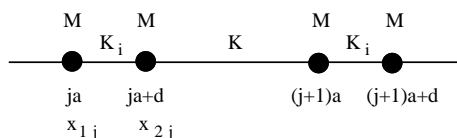


Figure 8.5: Chaîne linéaire à deux atomes par maille (même masse M).

Il est clair physiquement que les constantes de raideur sont maintenant différentes : K_i pour le ressort entre les deux atomes séparés de d (vibration intramoléculaire), K pour deux atomes appartenant à deux molécules adjacentes ; afin de fixer les idées, on suppose $K_i > K$, ce qui est physiquement bien naturel. On désigne par $x_{1,j}$ l'écart de l'atome qui vibre autour de ja , $x_{2,j}$ l'écart de celui qui est près de $ja + d$. Les équations du mouvement sont :

$$M\ddot{x}_{1,j} = -K(x_{1,j} - x_{2,j-1}) - K_i(x_{1,j} - x_{2,j}) = -(K + K_i)x_{1,j} + Kx_{2,j-1} + K_ix_{2,j} , \quad (8.95)$$

$$M\ddot{x}_{2,j} = -K_i(x_{2,j} - x_{1,j}) - K(x_{2,j} - x_{1,j+1}) = -(K_i + K)x_{2,j} + K_ix_{1,j} + Kx_{1,j+1} . \quad (8.96)$$

Pour les composantes de Fourier, on a :

$$[M\omega^2 - (K + K_i)] X_{1,j} + K X_{2,j-1} + K_i X_{2,j} = 0 , \quad (8.97)$$

et :

$$[M\omega^2 - (K_i + K)] X_{2,j} + K_i X_{1,j} + K X_{1,j+1} = 0 . \quad (8.98)$$

Posant :

$$\omega_0^2 = \frac{K_i + K}{M} , \quad \Omega_0^2 = \frac{K}{M} , \quad \Omega_i^2 = \frac{K_i}{M} , \quad \Omega_0^2 + \Omega_i^2 = 2\omega_0^2 , \quad (8.99)$$

il vient :

$$(\omega^2 - 2\omega_0^2) X_{1,j} + \Omega_0^2 X_{2,j-1} + \Omega_i^2 X_{2,j} = 0 , \quad (8.100)$$

et :

$$(\omega^2 - 2\omega_0^2) X_{2,j} + \Omega_i^2 X_{1,j} + \Omega_0^2 X_{1,j+1} = 0 . \quad (8.101)$$

Avec :

$$X_{1,j} = C_1 e^{ijk a} , \quad X_{2,j} = C_2 e^{ijk a} , \quad (8.102)$$

les deux équations (8.100) et (8.101) donnent le système homogène :

$$(\omega^2 - 2\omega_0^2) C_1 + \Omega_i^2 C_2 e^{-ika} + \Omega_0^2 C_2 = 0 , \quad (\omega^2 - 2\omega_0^2) C_2 + \Omega_i^2 C_1 + \Omega_0^2 C_1 e^{ika} = 0 . \quad (8.103)$$

³⁴Si on veut comparer les expressions obtenues ci-dessous avec celles trouvées pour le réseau à deux masses, il faut se souvenir qu'alors la maille avait pour côté $2a$.

La condition d'annulation du déterminant est :

$$\omega^2 = \Omega_0^2 + \Omega_i^2 \pm (\Omega_0^4 + 2\Omega_0^2\Omega_i^2 \cos ka + \Omega_i^4)^{1/2} = \frac{1}{M} \left[K + K_i \pm [(K + K_i)^2 - 4KK_i \sin^2 \frac{ka}{2}]^{1/2} \right], \quad (8.104)$$

et donne à nouveau une branche acoustique et une branche optique. Au voisinage des points remarquables, on a :

- $k \rightarrow 0$

$$\omega_+ \simeq \sqrt{2 \frac{K + K_i}{M}} \left[1 - \frac{KK_i}{8(K + K_i)^2} k^2 a^2 \right], \quad \omega_- \simeq \sqrt{\frac{KK_i}{2M(K + K_i)}} ka. \quad (8.105)$$

- $k \rightarrow \frac{\pi}{a}$ et si $K_i > K$

$$\omega_+ \simeq \sqrt{2 \frac{K_i}{M}} \left[1 + \frac{K}{2(K_i - K)} (\pi - ka)^2 \right], \quad \omega_- \simeq \sqrt{2 \frac{K}{M}} \left[1 - \frac{K_i}{2(K_i - K)} (\pi - ka)^2 \right]. \quad (8.106)$$

Le gap en bord de zone est donc :

$$\boxed{(\omega_+ - \omega_-)_{k=\frac{\pi}{a}} = \sqrt{2} (\Omega_i - \Omega_0)} \quad (8.107)$$

Notons que dans tous les cas, les déplacements s'écrivent finalement comme des combinaisons linéaires d'ondes planes $e^{i[\vec{k} \cdot n\vec{a} - \omega(\vec{k})t]}$ où $n\vec{a}$ est un vecteur de \mathcal{B} et où $\omega(\vec{k})$ est la loi de dispersion.

8.4 Modes normaux d'un réseau tridimensionnel

Les vibrations d'un réseau tridimensionnel s'obtiennent exactement comme précédemment, en tenant compte simplement du fait que le déplacement de chaque atome est déterminé par un vecteur $\vec{u}(\vec{R})$.

8.4.1 Réseau sans base

Lorsque le réseau est sans base, les composantes de $\vec{u}(\vec{R})$ obéissent aux équations du mouvement (8.15). Tout comme avant, des solutions particulières (complètes) peuvent s'écrire sous la forme d'ondes planes, dont il faut maintenant de surcroît préciser la polarisation $\vec{\varepsilon}$:

$$\vec{f}(\vec{R}, t) = \vec{\varepsilon} e^{i(\vec{k} \cdot \vec{R} - \omega t)}. \quad (8.108)$$

Pour chaque composante ε_α de $\vec{\varepsilon}$, le report de cette expression dans (8.15) donne :

$$-M\omega^2 \varepsilon_\alpha e^{i\vec{k} \cdot \vec{R}} = - \sum_{\beta} \sum_{\vec{R}'} Q_{\alpha\beta}(\vec{R}') \varepsilon_\beta e^{i\vec{k} \cdot \vec{R}'}. \quad (8.109)$$

Pour un réseau macroscopique, ceci se récrit comme suit :

$$M\omega^2 \varepsilon_\alpha = \sum_{\beta} \sum_{\vec{R}''} Q_{\alpha\beta}(\vec{R}'') \varepsilon_\beta e^{-i\vec{k} \cdot \vec{R}''} \equiv \sum_{\beta} Q_{\alpha\beta}(\vec{k}) \varepsilon_\beta, \quad (8.110)$$

ce que l'on peut aussi écrire sous forme matricielle :

$$M\omega^2 \vec{\varepsilon} = \mathcal{Q}(\vec{k}) \vec{\varepsilon}. \quad (8.111)$$

$\mathcal{Q}(\vec{k})$ est appelé matrice dynamique. Si N désigne toujours le nombre de nœuds (atomes), il existe maintenant $3N$ modes normaux distincts³⁵. En vertu des symétries de la matrice Q (voir (8.14)), on a :

$$\mathcal{Q}(\vec{k}) \equiv \frac{1}{2} \left[\sum_{\vec{R}} Q(\vec{R}) e^{-i\vec{k}\cdot\vec{R}} + \sum_{\vec{R}} Q(-\vec{R}) e^{+i\vec{k}\cdot\vec{R}} \right] = \sum_{\vec{R}} Q(\vec{R}) \cos \vec{k}\cdot\vec{R} . \quad (8.112)$$

Par ailleurs, compte tenu de (8.11), ceci vaut aussi :

$$\mathcal{Q}(\vec{k}) = \sum_{\vec{R}} Q(\vec{R}) \cos \vec{k}\cdot\vec{R} - \sum_{\vec{R}} Q(\vec{R}) = -2 \sum_{\vec{R}} Q(\vec{R}) \sin^2 \frac{1}{2} \vec{k}\cdot\vec{R} . \quad (8.113)$$

$\mathcal{Q}(\vec{k})$ est donc une combinaison réelle de matrices $Q(\vec{R})$ symétriques réelles de dimension 3, c'est donc une matrice symétrique réelle de dimension 3. En tant que telle, elle a 3 vecteurs propres réels (orthogonaux) satisfaisant :

$$\mathcal{Q}(\vec{k}) \vec{\varepsilon}_\lambda = q_\lambda(\vec{k}) \vec{\varepsilon}_\lambda \quad (\lambda = 1, 2, 3) ; \quad (8.114)$$

dans la suite, on supposera éventuellement que ces vecteurs sont de surcroît normalisés, de sorte que :

$$\vec{\varepsilon}_\lambda \cdot \vec{\varepsilon}_{\lambda'} = \delta_{\lambda\lambda'} . \quad (8.115)$$

Il y a trois polarisations distinctes (et orthogonales) pour un même vecteur³⁶ \vec{k} . Les pulsations propres $\omega_\lambda(\vec{k})$ sont données par :

$$M \omega_\lambda^2(\vec{k}) = q_\lambda(\vec{k}) \iff \omega_\lambda(\vec{k}) = \sqrt{\frac{q_\lambda(\vec{k})}{M}} . \quad (8.116)$$

Tout comme le réseau monoatomique unidimensionnel, les trois branches s'annulent linéairement en $\vec{k} = 0$. Dans cette limite :

$$\mathcal{Q}(\vec{k}) \simeq -\frac{1}{2} \sum_{\vec{R}} Q(\vec{R}) (\vec{k}\cdot\vec{R})^2 \equiv -\frac{1}{2} k^2 \sum_{\vec{R}} Q(\vec{R}) (\vec{\kappa}\cdot\vec{R})^2 , \quad k \rightarrow 0 \quad (8.117)$$

où $\vec{\kappa}$ est le vecteur unitaire porté par \vec{k} . Conformément à (8.116), les pulsations sont dans cette limite approximativement données par les racines carrées des valeurs propres de la matrice $-\frac{1}{2M} \sum_{\vec{R}} Q(\vec{R}) (\vec{\kappa}\cdot\vec{R})^2$; elles dépendent de la direction de \vec{k} par l'intermédiaire de $\vec{\kappa}$. Les notant $c_\lambda(\vec{\kappa})$, il vient :

$$\boxed{\omega_\lambda(\vec{k}) \simeq |k| c_\lambda(\vec{\kappa}) \quad ka \rightarrow 0} \quad (8.118)$$

$c_\lambda(\vec{\kappa})$ est la vitesse (anisotrope) de propagation du son dans le cristal. En l'absence de base, il n'existe que des modes de type acoustique.

8.4.2 Réseau avec base

Lorsque le réseau contient une base ("molécule") à n atomes, il existe au total $3nN$ modes de vibrations, c'est-à-dire $3n$ modes pour chaque valeur de \vec{k} . Parmi ces derniers, il y en a forcément 3 dont la fréquence tend vers zéro à petit k , donnant à la limite la translation d'énergie nulle. Il reste donc $3(n-1)$ modes, qui sont de nature optique, leur fréquence ayant une valeur finie en $\vec{k} = 0$. Ils sont en fait issus des $3(n-1)$ modes de vibration-rotation d'une molécule³⁷, qui présentent une dispersion en \vec{k} compte tenu de l'interaction entre les molécules constituant le cristal.

³⁵On admet que le cristal ne peut pas tourner – pas de degrés de liberté de rotation. Les trois modes de translation sont inclus dans la suite ($\vec{k} = \vec{0}$). De toute façon, avec $N \gg 1$, $3N$ ou $3N - 6$ font le même effet.

³⁶La disposition relative des polarisations $\vec{\varepsilon}_\lambda$ et de \vec{k} dépend de la symétrie (ponctuelle) du cristal. Pour un cristal isotrope, pour chaque \vec{k} , on peut toujours choisir les solutions de sorte qu'il existe un mode longitudinal et deux modes transverses. Lorsque le cristal est anisotrope, il n'en va plus ainsi, sauf si \vec{k} est lui-même invariant dans certaines opérations de symétrie.

³⁷Rappelons qu'une molécule à n atomes possède $3n$ degrés de liberté nucléaires. 3 d'entre eux représentent une translation en bloc de la molécule, tous les autres sont associés au mouvement du squelette moléculaire et peuvent se subdiviser en deux : la rotation en bloc de la molécule (autour d'une configuration fixée des noyaux, usuellement la configuration stable) qui exige le recours à 3 angles d'Euler pour une molécule quelconque, 2 seulement si la molécule est linéaire. En conséquence, il reste $3n - 6$ ou $3n - 5$ degrés de liberté de vibration, qui correspondent à des déformations de la molécule.

8.5 Quantification des vibrations de réseau

La quantification des vibrations harmoniques du réseau est immédiate. Il suffit de revenir au Hamiltonien (8.2) où U est remplacé par la forme quadratique (8.8). Les $3N$ composantes $u_\alpha(\vec{R})$ – pour un réseau sans base – peuvent être exprimées à l’aide des déplacements normaux $\xi_\lambda(\vec{k})$ qui diagonalisent la matrice dynamique (8.111). La transformation correspondante est unitaire, de sorte que la partie cinétique T de (8.2) est invariante en forme et devient :

$$T = \sum_{\lambda, \vec{k}} \frac{\Pi_{\lambda \vec{k}}^2}{2M} , \quad \Pi_{\lambda \vec{k}} = -i\hbar \frac{\partial}{\partial \xi_{\lambda \vec{k}}} , \quad (8.119)$$

cependant que la partie potentielle devient par construction :

$$U_h = \sum_{\lambda, \vec{k}} \frac{1}{2} M \omega_\lambda^2(\vec{k}) \xi_\lambda^2(\vec{k}) . \quad (8.120)$$

Les degrés de liberté normaux $\xi_\lambda(\vec{k})$ et les moments conjugués $\Pi_{\lambda \vec{k}}$ satisfont les relations de commutation canoniques :

$$[\xi_{\lambda \vec{k}}, \Pi_{\lambda' \vec{k}'}] = i\hbar \delta_{\lambda \lambda'} \delta_{\vec{k} \vec{k}'} , \quad (8.121)$$

en conséquence du fait que le passage des $\{u\}$ aux $\{\xi\}$ s’effectue par une transformation unitaire.

Il en résulte que H est maintenant une *somme* de $3N$ Hamiltoniens d’oscillateurs harmoniques :

$$H = \sum_{\lambda, \vec{k}} \left[\frac{\Pi_{\lambda \vec{k}}^2}{2M} + \frac{1}{2} M \omega_\lambda^2(\vec{k}) \xi_\lambda^2(\vec{k}) \right] . \quad (8.122)$$

L’additivité de H signifie que chacun des oscillateurs est indépendant de tous les autres. De ce fait, la construction des états de vibration du cristal s’effectue en formant les produits des fonctions propres d’oscillateurs harmoniques $\psi_{\lambda \vec{k}}$, associées aux énergies $E_\lambda(\vec{k}) = \hbar \omega_\lambda(\vec{k})$. Chaque *quantum* d’excitation est appelé phonon ; si le mode (λ, \vec{k}) est au fondamental, il contient zéro phonon, s’il est dans le premier état excité, il y a un phonon, *etc.* Les phonons sont au champ des vibrations de réseau ce que les photons sont au champ électromagnétique³⁸. Notons qu’un phonon acoustique de tout petit \vec{k} a, comme le photon, une relation de dispersion linéaire.

En définitive, l’état vibrationnel du cristal est complètement spécifié en donnant le nombre $n_{\lambda \vec{k}}$ de phonons dans chaque mode. Comme rien n’interdit un état de vibration arbitrairement excité pour un mode donné, $n_{\lambda \vec{k}}$ pouvant prendre n’importe quelle valeur $\in \mathbb{N}$, on peut caser un nombre arbitraire de phonons dans un seul mode : du point de vue de la statistique, les phonons sont des bosons³⁹. Si le mode (λ, \vec{k}) contient $n_{\lambda \vec{k}}$ phonons, il a l’énergie $\hbar \omega_\lambda(\vec{k})(n_{\lambda \vec{k}} + \frac{1}{2})$. Pour une distribution des nombres d’occupation $n_{\lambda \vec{k}}$ sur l’ensemble des modes, l’énergie de vibration totale du réseau est :

$$E = \sum_{\lambda, \vec{k}} \hbar \omega_\lambda(\vec{k}) \left(n_{\lambda \vec{k}} + \frac{1}{2} \right) . \quad (8.125)$$

Le demi-quantum est une simple constante additive⁴⁰, délaissée dans la suite puisque seules comptent les *différences* d’énergie. L’expression retenue pour l’énergie totale de vibration est ainsi :

$$E = \sum_{\lambda, \vec{k}} n_{\lambda \vec{k}} \hbar \omega_\lambda(\vec{k}) . \quad (8.126)$$

³⁸Toutefois, au contraire du champ électromagnétique, le champ de phonons contient un mode propagatif longitudinal.

³⁹Ceci se voit encore plus directement en introduisant les opérateurs d’annihilation et de création pour chaque mode. Ces opérateurs satisfont :

$$[b_{\lambda \vec{k}}, b_{\lambda' \vec{k}'}^\dagger] = \delta_{\lambda \lambda'} \delta_{\vec{k} \vec{k}'} . \quad (8.123)$$

La commutation de tels opérateurs est caractéristique des bosons. Pour des fermions, on a des relations d’anticommutation :

$$\{c_{\lambda \vec{k}}, c_{\lambda' \vec{k}'}^\dagger\} = \delta_{\lambda \lambda'} \delta_{\vec{k} \vec{k}'} , \quad (8.124)$$

où $\{A, B\} \equiv AB + BA$.

⁴⁰éventuellement infinie. . .

8.6 Contribution des phonons à la chaleur spécifique

Pour trouver la chaleur spécifique due aux phonons, il suffit d'écrire l'expression de la valeur moyenne de l'énergie du réseau en équilibre thermique à la température T . La distribution des nombres moyens d'occupation est celle de bosons, donnée par la statistique de Bose - Einstein :

$$n_{\lambda \vec{k}} = \frac{1}{e^{\beta \hbar \omega_{\lambda}(\vec{k})} - 1}, \quad \beta = \frac{1}{k_B T}. \quad (8.127)$$

Il en résulte immédiatement que l'énergie moyenne (au sens de la statistique quantique) à la température T est donnée par (8.125) où l'expression (8.127) est utilisée. Si V désigne le volume de l'échantillon, la chaleur spécifique due aux phonons est⁴¹ :

$$c_{V \text{ ph}} = \frac{1}{V} \frac{\partial E}{\partial T} = \frac{1}{V} \frac{\partial}{\partial T} \sum_{\lambda, \vec{k}} n_{\lambda \vec{k}} \hbar \omega_{\lambda}(\vec{k}) = \frac{1}{V} \frac{\partial}{\partial T} \sum_{\lambda, \vec{k}} \frac{\hbar \omega_{\lambda}(\vec{k})}{e^{\beta \hbar \omega_{\lambda}(\vec{k})} - 1}. \quad (8.128)$$

À très haute température ($\beta \hbar \omega_{\lambda}(\vec{k}) \ll 1$) :

$$c_{V \text{ ph}} \simeq \frac{1}{V} \frac{\partial}{\partial T} \sum_{\lambda, \vec{k}} \hbar \omega_{\lambda}(\vec{k}) (\beta \hbar \omega_{\lambda}(\vec{k}))^{-1} = \frac{1}{V} \frac{\partial}{\partial T} \sum_{\lambda, \vec{k}} k_B T = \frac{3N}{V} k_B. \quad (8.129)$$

C'est le résultat classique, appelé loi de Dulong et Petit. Comme toujours, la limite haute température efface le *quantum* porté par \hbar et la physique devient classique⁴². Les corrections quantiques sont faciles à écrire à partir de (8.128), mais sont souvent non-significatives car du même ordre de grandeur que les anharmonicités, ignorées d'un bout à l'autre.

L'autre limite, basse température, est évidemment plus intéressante puisque révélatrice des effets quantiques. Il est plus commode de discuter la forme intégrale de (8.128) ; comme $\delta \vec{k} = (\frac{2\pi}{L})^3 = \frac{(2\pi)^3}{V}$, il vient :

$$c_{V \text{ ph}} = \frac{\partial}{\partial T} \sum_{\lambda} \int_{\text{BZ1}} \frac{d^3 k}{(2\pi)^3} \frac{\hbar \omega_{\lambda}(\vec{k})}{e^{\beta \hbar \omega_{\lambda}(\vec{k})} - 1}. \quad (8.130)$$

En se plaçant à température assez basse, on peut négliger tous les modes optiques éventuels⁴³, qui donnent des contributions exponentiellement petites par rapport à celle des modes acoustiques. Par ailleurs, en supposant de plus $k_B T \ll \hbar k c_{\lambda}$ pour toutes les branches acoustiques retenues, on peut remplacer les vraies lois de dispersion $\hbar \omega_{\lambda}(\vec{k})$ par leurs expressions approchées à petit \vec{k} : à nouveau, la substitution n'introduit que des erreurs exponentiellement petites⁴⁴. Enfin, dans le droit-fil, on rejette à l'infini les bornes d'intégration. En conséquence, une expression admissible de $c_{V \text{ ph}}$ est :

$$c_{V \text{ ph}} = \frac{\partial}{\partial T} \sum_{\text{acoustiques}} \int_{\mathbb{R}^3} \frac{d^3 k}{(2\pi)^3} \frac{\hbar k c_{\text{ac}}(\vec{k})}{e^{\beta \hbar k c_{\text{ac}}(\vec{k})} - 1}, \quad (8.131)$$

où la somme ne porte que sur les modes acoustiques, caractérisés par les différentes vitesses $c_{\text{ac}}(\vec{k})$, celles-ci ne dépendant que de la direction de \vec{k} , pas de son module. En séparant les dépendances radiale et angulaire, chaque intégrale de la somme se transforme comme suit ($d^3 k = k^2 dk d\Omega$) :

$$\int_0^{+\infty} \frac{k^2 dk}{2\pi^2} \int_{4\pi} \frac{d\Omega(\vec{k})}{4\pi} \frac{\hbar k c_{\text{ac}}(\vec{k})}{e^{\beta \hbar k c_{\text{ac}}(\vec{k})} - 1}; \quad (8.132)$$

comme $c_{\text{ac}}(\vec{k})$ ne dépend que de l'orientation de \vec{k} , on peut extraire la fraction contenant la température de l'intégrale sur $d\Omega$. En posant $x = \beta \hbar k c_{\text{ac}}(\vec{k})$, il vient ainsi :

$$\frac{1}{2\pi^2 \beta^4 \hbar^3} \int_0^{+\infty} \frac{x^3 dx}{e^x - 1} \int_{4\pi} \frac{d\Omega(\vec{k})}{4\pi} \frac{1}{[c_{\text{ac}}(\vec{k})]^3}. \quad (8.133)$$

⁴¹On remarque que la constante additive d'écart entre (8.125) et (8.126) ne joue aucun rôle pour $c_{V \text{ ph}}$.

⁴²Quoique, comme toujours, l'approche est singulière.

⁴³présents s'il existe un motif.

⁴⁴L'exponentielle $e^{\beta \hbar \omega_{\lambda}(\vec{k})}$ du dénominateur agit comme une fonction de coupure aux grands k et ne retient que les vecteurs d'onde tels que $\hbar \omega_{\lambda}(\vec{k}) \lesssim k_B T$.

C'est le $\beta^{-4} \propto T^4$ qui va donner par dérivation $c_{V \text{ ph}} \propto T^3$. Dans les détails, on peut s'y prendre comme suit ; en introduisant la moyenne angulaire de $\frac{1}{c_{\text{ac}}^3}$:

$$\frac{1}{c^3} \equiv \frac{1}{3} \sum_{\text{acoustiques}} \int_{4\pi} \frac{d\Omega(\vec{k})}{4\pi} \frac{1}{[c_{\text{ac}}(\vec{k})]^3}, \quad (8.134)$$

et en utilisant :

$$\int_0^{+\infty} \frac{x^3 dx}{e^x - 1} = 3! \sum_{n=1}^{+\infty} \frac{1}{n^4} = \frac{\pi^4}{15}, \quad (8.135)$$

on trouve finalement :

$$c_{V \text{ ph}} = \frac{\pi^2 k_{\text{B}}^4}{10 \hbar^3 c^3} \frac{\partial}{\partial T} T^4 = \frac{2\pi^2}{5} \left(\frac{k_{\text{B}} T}{\hbar c} \right)^3 k_{\text{B}} \quad (8.136)$$

À basse température, la contribution des phonons à la chaleur spécifique varie bien essentiellement comme T^3 . Toutefois, ce comportement n'est observé qu'à *très* basse température : comme les lois de dispersion s'éloignent assez vite de la linéarité, il existe en général un grand intervalle de température où, la loi de Dulong et Petit étant déjà invalide, la loi en T^3 n'est pas pour autant vérifiée.

Pour un métal, cette variation typique en T^3 s'ajoute à la contribution du gaz de Fermi constitué par les électrons, qui donne un comportement linéaire en température pour un métal : seuls les électrons situés au niveau de Fermi sont susceptibles d'être excités thermiquement, et leur nombre est d'ordre $\mathcal{D}(\varepsilon_{\text{F}}) k_{\text{B}} T$. La variation d'énergie par fluctuation thermique est donc d'ordre $\mathcal{D}(\varepsilon_{\text{F}}) k_{\text{B}} T \times k_{\text{B}} T$; par dérivation, on en déduit $c_{V \text{ el}} \propto T$. Au total, la chaleur spécifique à basse température est une loi du genre :

$$c_V = (\gamma T + AT^3) k_{\text{B}} \quad (8.137)$$

où γ est appelé "constante de Sommerfeld" (qui contient la température de Fermi T_{F} du gaz d'électrons dégénéré).

8.7 Manifestation des phonons dans les spectres de diffraction

La théorie élémentaire de la diffraction exposée dans un chapitre précédent repose sur l'hypothèse d'un cristal parfait (pas de défauts, ni structurels, ni chimiques), dont les objets physiques sont immobiles. La conclusion majeure alors obtenue, en ne considérant que la diffusion élastique, est l'apparition d'une tache ponctuelle de diffraction à chaque fois que le transfert de moment $\vec{q} = \vec{k}_i - \vec{k}_f$ coïncide avec un vecteur \vec{K} du réseau réciproque.

Bien évidemment, cette conclusion doit être reconsidérée si les objets diffuseurs ont la liberté de se mouvoir, et notamment s'ils effectuent des petites oscillations autour d'une position d'équilibre, définie par le réseau de Bravais. Cette nécessité s'impose quand on se souvient que c'est précisément l'ordre à longue distance, parfaitement déterminé pour un réseau sans défauts et immobile, qui produit l'effet coopératif : l'action collective d'un grand nombre de diffuseurs produit une hypersélectivité donnant lieu à des taches ponctuelles de diffraction. Par ailleurs, lorsqu'il s'agit de diffusion neutronique, il y a maintenant possibilité de diffusion *inélastique* avec création ou annihilation de phonons⁴⁵, puisque les bons neutrons ont une énergie thermique et que, dès lors, on peut satisfaire les deux lois de conservation énergie et impulsion⁴⁶. Afin d'en rester au niveau élémentaire, on continue toutefois à ne considérer que la diffusion élastique.

L'objet central est l'amplitude de diffusion, dont on a vu qu'elle contenait le facteur décisif :

$$S = \sum_{\vec{R} \in \mathcal{B}} e^{i\vec{q} \cdot \vec{R}}, \quad \vec{q} = \vec{k}_i - \vec{k}_f. \quad (8.138)$$

⁴⁵En ce qui concerne l'irradiation X, on continue à laisser de côté les diffusions inélastiques à très haute énergies (~ 10 keV).

⁴⁶D'ailleurs, la diffusion inélastique des neutrons est une méthode puissante pour obtenir expérimentalement les lois de dispersion des phonons.

Dans le cas où les atomes effectuent des petites vibrations, compte tenu des notations introduites en (8.1), on doit maintenant écrire :

$$S = \sum_{\vec{R} \in \mathcal{B}} e^{i\vec{q} \cdot [\vec{R} + \vec{u}(\vec{R})]} \quad (8.139)$$

L'intensité de diffraction est proportionnelle à $|S|^2$

$$|S|^2 = \sum_{\vec{R}, \vec{R}'} e^{i\vec{q} \cdot [\vec{R} + \vec{u}(\vec{R})]} e^{-i\vec{q} \cdot [\vec{R}' + \vec{u}(\vec{R}')] } = \sum_{\vec{R}, \vec{R}'} e^{i\vec{q} \cdot (\vec{R} - \vec{R}')} e^{i\vec{q} \cdot [\vec{u}(\vec{R}) - \vec{u}(\vec{R}')] } ; \quad (8.140)$$

ceci est l'amplitude pour une configuration donnée des écarts mais ce que l'on observe, c'est l'amplitude *moyenne* à une température donnée T . Ainsi, pour un réseau en équilibre thermique, la quantité importante à trouver est :

$$\langle |S|^2 \rangle = \sum_{\vec{R}, \vec{R}'} e^{i\vec{q} \cdot (\vec{R} - \vec{R}')} \langle e^{i\vec{q} \cdot [\vec{u}(\vec{R}) - \vec{u}(\vec{R}')] } \rangle , \quad (8.141)$$

où les crochets représentent la moyenne prise au sens de la statistique (classique à très haute température, forcément quantique autrement). La moyenne de droite est évidemment invariante dans toute translation du réseau, de sorte que :

$$\langle |S|^2 \rangle = N \sum_{\vec{R}} e^{-i\vec{q} \cdot \vec{R}} \langle e^{i\vec{q} \cdot \vec{u}(\vec{0})} e^{-i\vec{q} \cdot \vec{u}(\vec{R})} \rangle ; \quad (8.142)$$

en définitive, il faut en premier lieu obtenir la quantité :

$$Q \stackrel{\text{def}}{=} \langle e^{i\vec{q} \cdot \vec{u}(\vec{0})} e^{-i\vec{q} \cdot \vec{u}(\vec{R})} \rangle \quad (8.143)$$

puis effectuer la sommation sur \vec{R} .

On sait que la coordonnée et le moment conjugué d'un oscillateur harmonique peuvent s'exprimer en combinaison linéaire des opérateurs de création et d'annihilation b^\dagger et b . Les déplacements $\vec{u}(\vec{R})$ – et leurs moments conjugués – sont donc des combinaisons linéaires des opérateurs $b_{\vec{k}\lambda}^\dagger$ et $b_{\vec{k}\lambda}$ associés aux modes normaux déterminés plus haut. Il en résulte que Q est de la forme :

$$Q = \langle e^A e^B \rangle , \quad (8.144)$$

où A et B sont des formes linéaires des $b_{\vec{k}\lambda}^\dagger$ et $b_{\vec{k}\lambda}$. Ces opérateurs satisfont :

$$[b_{\vec{k}\lambda}, b_{\vec{k}'\lambda'}^\dagger] = \delta_{\vec{k}\vec{k}'} \delta_{\lambda\lambda'} , \quad (8.145)$$

de sorte que l'on peut remplacer sans complication l'exponentielle d'une somme par le produit des exponentielles. Q apparaît alors sous la forme :

$$Q = \langle \prod_{\vec{k}, \lambda} e^{A_{\vec{k}\lambda}} e^{B_{\vec{k}\lambda}} \rangle . \quad (8.146)$$

Maintenant, $A_{\vec{k}\lambda}$ et $B_{\vec{k}\lambda}$ sont des combinaisons linéaires des deux seuls opérateurs $b_{\vec{k}\lambda}^\dagger$ et $b_{\vec{k}\lambda}$. En tant que tels, ils relèvent de la formule de Glauber (voir [7], p. 376 ou [8], complément B_{II}) :

$$e^{A_{\vec{k}\lambda} + B_{\vec{k}\lambda}} = e^{A_{\vec{k}\lambda}} e^{B_{\vec{k}\lambda}} e^{-\frac{1}{2} [A_{\vec{k}\lambda}, B_{\vec{k}\lambda}]} , \quad (8.147)$$

où le commutateur est un simple nombre (qui, notamment, est insensible aux opérations de moyennes). Lue à l'envers, cette relation donne :

$$e^{A_{\vec{k}\lambda}} e^{B_{\vec{k}\lambda}} = e^{A_{\vec{k}\lambda} + B_{\vec{k}\lambda}} e^{+\frac{1}{2} [A_{\vec{k}\lambda}, B_{\vec{k}\lambda}]} . \quad (8.148)$$

Il en résulte :

$$Q = \prod_{\vec{k}, \lambda} \langle e^{A_{\vec{k}\lambda} + B_{\vec{k}\lambda}} \rangle e^{+\frac{1}{2} [A_{\vec{k}\lambda}, B_{\vec{k}\lambda}]} . \quad (8.149)$$

$A_{\vec{k}\lambda} + B_{\vec{k}\lambda}$ est à nouveau une combinaison linéaire des deux opérateurs $b_{\vec{k}\lambda}^\dagger$ et $b_{\vec{k}\lambda}$:

$$A_{\vec{k}\lambda} + B_{\vec{k}\lambda} = \gamma_{\vec{k}\lambda} b_{\vec{k}\lambda} + \mu_{\vec{k}\lambda} b_{\vec{k}\lambda}^\dagger , \quad (8.150)$$

de sorte que l'on doit finalement calculer des valeurs moyennes du genre $q = \langle e^{\gamma b + \mu b^\dagger} \rangle$ où b et b^\dagger sont relatifs à un certain mode. Explicitement :

$$q = Z^{-1} \sum_{n=0}^{+\infty} e^{-\beta \hbar \omega [n + \frac{1}{2}]} \langle n | e^{\gamma b + \mu b^\dagger} | n \rangle, \quad Z = \sum_{n=0}^{+\infty} e^{-\beta \hbar \omega [n + \frac{1}{2}]} . \quad (8.151)$$

Un calcul facile montre que⁴⁷ :

$$q = e^{\frac{1}{2} \langle (\gamma b + \mu b^\dagger)^2 \rangle} . \quad (8.152)$$

Alors ($j \equiv \vec{k}, \lambda$) :

$$Q = \prod_j e^{\frac{1}{2} \langle (\gamma_j b_j + \mu_j b_j^\dagger)^2 \rangle} e^{+\frac{1}{2} [A_j, B_j]} = \prod_j e^{\frac{1}{2} \langle (A_j + B_j)^2 \rangle} e^{+\frac{1}{2} [A_j, B_j]} . \quad (8.153)$$

Maintenant :

$$\prod_j e^{\frac{1}{2} \langle (A_j + B_j)^2 \rangle} = e^{\frac{1}{2} \sum_j \langle (A_j + B_j)^2 \rangle} = e^{\frac{1}{2} \langle [\sum_j A_j + B_j]^2 \rangle} . \quad (8.154)$$

La dernière égalité vient du fait que, si $j \neq j'$, $\langle A_j B_{j'} \rangle = \langle A_j \rangle \langle B_{j'} \rangle$ et que la valeur moyenne d'une fonction linéaire de B et/ou b^\dagger est nulle. Alors, (8.153) donne :

$$Q = e^{\frac{1}{2} \langle [\sum_j A_j + B_j]^2 \rangle} e^{+\frac{1}{2} \sum_j [A_j, B_j]} . \quad (8.155)$$

$[A_j, B_j]$ est un scalaire et peut être reporté dans l'argument de la première exponentielle ; ce dernier devient :

$$\begin{aligned} \frac{1}{2} \sum_j \langle (A_j + B_j)^2 \rangle + [A_j, B_j] &= \frac{1}{2} \sum_j \langle A_j^2 + B_j^2 + A_j B_j + B_j A_j + A_j B_j - B_j A_j \rangle \\ &= \frac{1}{2} \sum_j \langle A_j^2 + B_j^2 + 2A_j B_j \rangle . \end{aligned} \quad (8.156)$$

Au total, Q défini en (8.144) est égal à :

$$Q = e^{\frac{1}{2} \sum_j \langle A_j^2 + B_j^2 + 2A_j B_j \rangle} , \quad (8.157)$$

de sorte que $Q \stackrel{\text{déf}}{=} \langle e^{i\vec{q} \cdot \vec{u}(\vec{0})} e^{-i\vec{q} \cdot \vec{u}(\vec{R})} \rangle$ est donné par :

$$\begin{aligned} Q &= \exp \left\langle -\frac{1}{2} [\vec{q} \cdot \vec{u}(\vec{0})]^2 - \frac{1}{2} [\vec{q} \cdot \vec{u}(\vec{R})]^2 + [\vec{q} \cdot \vec{u}(\vec{0})] [\vec{q} \cdot \vec{u}(\vec{R})] \right\rangle \\ &= \exp \left\langle -[\vec{q} \cdot \vec{u}(\vec{0})]^2 + [\vec{q} \cdot \vec{u}(\vec{0})] [\vec{q} \cdot \vec{u}(\vec{R})] \right\rangle \equiv e^{-2W} \exp \langle [\vec{q} \cdot \vec{u}(\vec{0})] [\vec{q} \cdot \vec{u}(\vec{R})] \rangle . \end{aligned} \quad (8.158)$$

Dans la dernière expression, on a introduit la quantité $2W$:

$$2W \stackrel{\text{déf}}{=} \langle [\vec{q} \cdot \vec{u}(\vec{0})]^2 \rangle \quad (8.159)$$

e^{-2W} s'appelle le facteur de Debye - Waller. En définitive, la quantité centrale contenant la diffusion élastique prend la forme (voir (8.142)) :

$$\langle |S|^2 \rangle = N e^{-2W} \sum_{\vec{R}} e^{-i\vec{q} \cdot \vec{R}} \exp \left\langle [\vec{q} \cdot \vec{u}(\vec{0})] [\vec{q} \cdot \vec{u}(\vec{R})] \right\rangle \equiv N e^{-2W} \sum_{\vec{R}} e^{-i\vec{q} \cdot \vec{R}} e^{\Phi(\vec{R})} . \quad (8.160)$$

Ce résultat montre deux effets des vibrations de réseau sur les spectres de diffraction. D'une part, l'apparition du facteur de Debye - Waller révèle une atténuation des intensités de diffraction, d'autant plus grande que la moyenne $\langle [\vec{q} \cdot \vec{u}(\vec{0})]^2 \rangle$ est élevée, c'est-à-dire que la température est haute.

D'autre part, la somme de réseau ne conduit plus à une superposition de fonctions de Dirac, en vertu de la dépendance spatiale de la moyenne $e^{\Phi(\vec{R})}$ définie en (8.160). Si cette fonction – qui est constante et égale

⁴⁷La relation (8.152) est connue sous le nom d'identité de Bloch, voir par exemple [5]. On peut y voir un avatar élaboré du fait que la transformée de Fourier d'une gaussienne est une gaussienne (voir [8], complément L_V).

à 1 pour le réseau immobile – vaut 1 pour la plupart des valeurs de \vec{R} , on retrouve en gros la situation du réseau fixe, les angles spécifiques de diffraction étant légèrement entachés d'un "flou" qui se traduit par un petit élargissement angulaire des taches de diffraction. À l'inverse, si $e^{\Phi(\vec{R})}$ varie significativement à l'échelle du réseau, l'incertitude $\delta\theta$ devient de l'ordre de 2π ; dit autrement, la diffusion coopérative perd tout caractère d'hypersélectivité⁴⁸. Le point crucial est donc l'allure de $\Phi(\vec{R})$.

À l'infini, cette fonction vaut 1 : la moyenne du produit devient le produit des moyennes pour deux points très éloignés l'un de l'autre, chaque moyenne vaut 0, l'exponentielle est donc égale à 1 :

$$e^{\Phi(\infty)} = 1 . \quad (8.161)$$

À l'opposé, en $\vec{R} = \vec{0}$, c'est la valeur quadratique moyenne des déplacements qui apparaît :

$$e^{\Phi(\vec{0})} = \exp\langle [\vec{q} \cdot \vec{u}(\vec{0})]^2 \rangle . \quad (8.162)$$

La question se résume donc à savoir sur quelle distance Φ passe d'une valeur à l'autre. On observe que ceci se produit toujours sur quelques distances atomiques à peine, de sorte que, en première approximation, Φ est presque constante : pour "presque toutes" les valeurs de \vec{R} , $e^{\Phi(\vec{R})} \simeq e^{\Phi(\infty)}$. La somme dans (8.160) se décompose comme suit :

$$\sum_{\vec{R}} e^{-i\vec{q} \cdot \vec{R}} e^{\Phi(\vec{R})} = \sum_{\vec{R}} e^{-i\vec{q} \cdot \vec{R}} e^{\Phi(\infty)} - \sum_{\vec{R}} e^{-i\vec{q} \cdot \vec{R}} [e^{\Phi(\infty)} - e^{\Phi(\vec{R})}] ; \quad (8.163)$$

la première somme est d'ordre N pour un réseau fini, et donne un produit de trois peignes de Dirac si $N \rightarrow +\infty$, reproduisant la condition de von Laue (ou de Bragg) ; au contraire, la deuxième somme est d'ordre quelques unités, et est négligeable devant la première. Il en résulte que l'incertitude angulaire reste très petite et, dans les cas usuels, quasi-inobservable : en pareil cas, l'effet des vibrations se borne à une simple réduction d'intensité sans élargissement perceptible.

Cette conclusion doit cependant être un peu tempérée eu égard à des effets de dimensionnalité : si elle est correcte pour des cristaux à trois dimensions, elle peut être invalidée à $d = 2$ (ou *a fortiori* pour un réseau unidimensionnel), pour les raisons suivantes.

L'équation (8.162) implique la valeur moyenne du carré d'une quantité qui est à nouveau une combinaison linéaire d'opérateurs de création et d'annihilation. Pour un seul oscillateur de fréquence ω , on a :

$$u = \sqrt{\frac{\hbar}{2M\omega}} (b + b^\dagger) . \quad (8.164)$$

La moyenne de u^2 est donc :

$$\langle u^2 \rangle = \frac{\hbar}{2M\omega} \frac{1}{Z} \sum_{n=0}^{+\infty} e^{-\beta[n+\frac{1}{2}]\hbar\omega} \langle b^2 + b^{\dagger 2} + bb^\dagger + b^\dagger b \rangle . \quad (8.165)$$

Un calcul simple donne le résultat⁴⁹ :

$$\langle u^2 \rangle = \frac{\hbar}{2M\omega} \coth \frac{\beta\hbar\omega}{2} . \quad (8.166)$$

Pour les combinaisons linéaires que constituent les $\vec{q} \cdot \vec{u}(\vec{0})$, on va donc voir apparaître des arguments en exponentielle qui sont des sommes analogues sur l'ensemble des modes :

$$q^2 \sum_{\vec{k}\lambda} \frac{\hbar}{2M\omega_\lambda(\vec{k})} \coth \frac{\beta\hbar\omega_\lambda(\vec{k})}{2} . \quad (8.167)$$

Fondamentalement, pour chaque branche, la somme fait apparaître une intégrale du genre :

$$\int \frac{k^{d-1} dk}{\omega_\lambda(\vec{k})} \coth \frac{\beta\hbar\omega_\lambda(\vec{k})}{2} , \quad (8.168)$$

⁴⁸On raconte que von Laue, le pionnier de la diffraction X, ne croyait pas trop à la possibilité d'observer la diffusion sélective, en raison des effets de vibration de réseau. Si l'expérience s'est révélée positive, c'est *justement* parce que, en pratique, la fonction $\Phi(\vec{R})$ est presque partout constante.

⁴⁹À la limite classique, on retrouve bien le théorème d'équipartition : $\frac{1}{2}M\langle u^2 \rangle = \frac{1}{2}k_B T$.

où d est la dimension spatiale. Cette intégrale présente une singularité en $\vec{k} = 0$ en dimension inférieure à 3. En effet, près de l'origine, l'intégrand est comme $k^{d-1}\omega_\lambda(\vec{k})^{-2} \sim k^{d-3}$; l'intégrale ne converge donc que si $d - 3 > -1$ soit $d > 2$. À $d = 2$, la divergence est "douce" (logarithmique), à $d = 1$, elle est relativement dure. On sait évidemment traiter ces divergences⁵⁰ en introduisant des coupures appropriées faisant intervenir (par exemple) la taille de l'échantillon. Quoiqu'il en soit, elles constituent un signal d'alarme montrant que, dans ces systèmes de basse dimensionnalité, les corrections dues aux vibrations jouent un rôle important et donnent, ici, une véritable largeur angulaire aux taches de diffraction⁵¹.

⁵⁰dites infra-rouges car survenant à basse fréquence.

⁵¹Finalement, ce sont les phonons acoustiques qui donnent une amplification spectaculaire de l'effet vibrationnel. Ceci n'est pas surprenant : la diffusion sélective étant un phénomène coopératif impliquant l'ordre à longue distance, les phonons de grande longueur d'onde sont les plus capables de modifier sérieusement le comportement par rapport au réseau fixe.

Bibliographie

- [1] I. Z. Fischer, *Statistical Theory of Liquids* (The University of Chicago Press, Chicago and London, 1964)
- [2] D. A. McQuarrie, *Statistical Mechanics* (Harper & Row, New York, 1973)
- [3] F. Reif, *Fundamentals of Statistical and Thermal Physics* (McGraw-Hill, 1965)
- [4] S. A. Rice et P. Gray, *The Statistical Mechanics of Simple Liquids* (Interscience Publishers, New York, 1965)
- [5] W. Marshall et S. W. Lovesey, *Theory of Thermal Neutron Scattering* (Holt, Rinehart et Winston, Oxford, 1971)
- [6] N. W. Ashcroft et N. D. Mermin, *Solid State Physics* (Clarendon Press, Oxford, New York 1976)
- [7] Albert Messiah, *Mécanique Quantique* tome 1 (Dunod, Paris 1965)
- [8] Cl. Cohen-Tannoudji, B. Diu et F. Laloë, *Mécanique Quantique* (Hermann, Paris, 1973)
- [9] R.D. Mattuck, *A Guide to Feynman Diagrams in the Many-Body Problem* (McGraw-Hill, New York, 1967)