

Propriétés électroniques des semiconducteurs

M. Lannoo

Cette conférence est consacrée aux propriétés particulières des électrons dans les semiconducteurs, dont l'origine réside dans leur structure de bandes spécifique. La première partie décrit qualitativement la formation des bandes dans les solides et se termine par une discussion distinguant isolant, métal et semiconducteur. Suit une description rapide des méthodes permettant de calculer ces structures de bandes, avec application au cas des semiconducteurs de structure zinc-blende. Les semiconducteurs à structure en chaînes ou en couches sont aussi abordés. On montre ensuite que le concept de bandes peut être étendu aux alliages ternaires, aux systèmes amorphes et vitreux. L'effet des impuretés dans les matériaux cristallins est ensuite abordé dans le cadre de l'approximation de la masse effective qui est brièvement discutée. On montre ainsi la possibilité de doper ces matériaux. Au passage, on introduit l'existence des excitons et leurs conséquences sur les propriétés optiques. Ceci se termine par un survol des nanostructures à semiconducteurs de dimensionnalité réduite. Pour terminer, un bref aperçu des applications de ces propriétés est présenté avec les perspectives actuelles de développement.

I. Formation des bandes dans un solide

Nous discutons ici l'origine physique de la structure de bandes des solides. Nous commençons par des arguments qualitatifs, partant de la limite atomique. Nous montrons ensuite qu'il est possible d'établir ces résultats de manière plus rigoureuse pour les solides cristallins. Nous en déduisons les critères permettant de distinguer métal, isolant et semiconducteur.

I.1. Description qualitative de la formation de bandes

Nous allons considérer ici le solide comme une grosse molécule dont la taille devient très grande. Pour simplifier, nous traiterons le cas d'un système formé de N atomes d'hydrogène. Nous commencerons par la molécule H_2 ($N = 2$) et étudierons ce qui se passe quand $N \rightarrow \infty$. Ceci nous permettra d'illustrer très simplement l'origine de la formation des bandes, les arguments qualitatifs pouvant se généraliser facilement à des systèmes plus complexes.

Prenons d'abord le cas de H_2 et étudions l'évolution des niveaux d'énergie en fonction de la distance R entre les atomes H . Quand $R \rightarrow \infty$, les deux atomes ne peuvent pas interagir ; les niveaux d'énergie sont donc ceux des atomes isolés, mais avec une dégénérescence deux fois plus grande puisque la fonction d'onde du système peut correspondre à celle de l'un ou l'autre des atomes. Quand R diminue à partir de cette limite, les atomes commencent à interagir faiblement. L'effet de cette interaction sera de décaler les niveaux atomiques et de lever leur dégénérescence (théorie des perturbations au premier ordre sur un état dégénéré). On aboutit ainsi aux courbes $E(R)$ de la figure 1a permettant de déduire les niveaux de la molécule à partir de ceux des atomes libres et de leur dégénérescence.

L'argument peut se généraliser au cas de N atomes d'hydrogène avec une distance R entre premiers voisins. Pour $R \rightarrow \infty$, on obtient à nouveau les niveaux atomiques mais avec une dégéné-

rescence N fois plus élevée. Lorsque R diminue, l'interaction entre atomes va lever cette dégénérescence provoquant l'apparition d'un nombre de sous-niveaux proportionnel à N , situés pour chaque niveau E_j entre des limites supérieure E_{jmax} et inférieure E_{jmin} (figure 1b) dont la valeur devient indépendante de N lorsque celui-ci est grand. Ces ensembles de sous-niveaux se resserrent donc pour $N \rightarrow \infty$, chacun formant un pseudocontinuum encore appelé bande d'énergies permises. Celles-ci sont séparées par des intervalles d'énergie appelés bandes interdites. Il est évident que ces conclusions qualitatives resteront valables pour des systèmes plus complexes contenant divers types d'atomes. Toutefois, l'existence de limites de bandes bien définies dépend de l'existence d'un certain ordre dans le système.

I.2. Dérivation rigoureuse de l'existence des bandes pour un cristal

Même pour un solide cristallin, le calcul des bandes est compliqué par l'existence de surfaces. Or, dans les conditions usuelles (où le rapport du nombre d'atomes de surface à celui de volume tend vers zéro), l'influence des conditions aux limites de surface sur les propriétés électroniques de volume devient négligeable. Il est donc possible d'utiliser n'importe quel type de conditions aux limites, en particulier celles qui permettront de simplifier le problème au maximum. C'est le cas des conditions cycliques de Born-Von Karman pour lesquelles on impose que le cristal se répète indéfiniment dans toutes les directions de l'espace. Ceci a l'avantage de rendre le système invariant par translation. Cette invariance introduit une simplification mathématique considérable. Ceci s'exprime d'abord par le théorème de Bloch qui permet d'écrire la fonction d'onde d'un électron sous la forme d'une onde plane modulée périodiquement :

$$\psi_{\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} u_{\mathbf{k}}(\mathbf{r}) \quad (1)$$

où \mathbf{r} repère la position de l'électron, \mathbf{k} est son vecteur d'onde et la fonction $u_{\mathbf{k}}(\mathbf{r})$ est périodique dans l'espace, sa période étant la maille élémentaire du cristal. Pour résoudre l'équation de Schrödinger, on développe habituellement la fonction d'onde suivant une base dont le nombre de fonctions est nécessairement

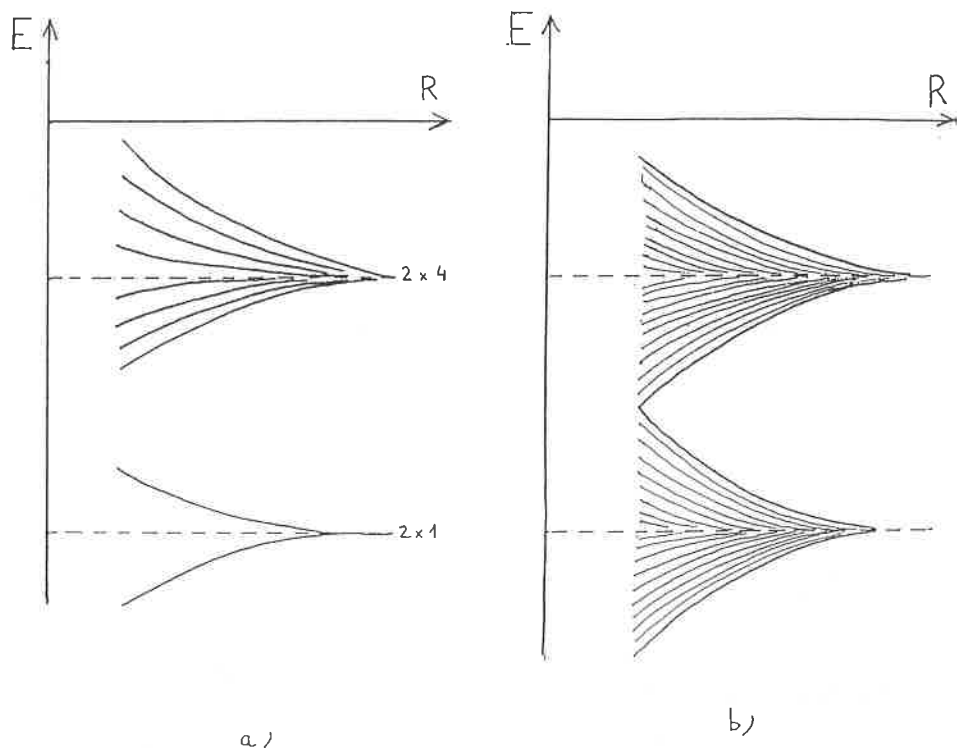


FIGURE 1. - Levée de dégénérescence des niveaux atomiques lors de la formation : a) de la molécule d'hydrogène H_2 , b) d'une molécule fictive de N atomes H_N .

proportionnel à la taille du cristal et est donc égal à $g \times N$, où N est le nombre de mailles élémentaires, ce qui conduit à une matrice $gN \times gN$ à diagonaliser. Le grand avantage d'utiliser des fonctions de Bloch définies par (1) est de réduire la taille des matrices à diagonaliser de $gN \times gN$ à $g \times g$ du fait que les conditions aux limites à appliquer aux fonctions $u_k(r)$ se font en surface de la maille élémentaire au lieu de la surface du cristal.

Les valeurs propres de l'équation de Schrödinger correspondant aux fonctions de Bloch sont donc des fonctions $E(k)$ du vecteur d'onde de l'électron. On peut montrer que $E(k)$ est périodique dans l'espace réciproque dont les vecteurs de base a_i^* sont définis par rapport à ceux du réseau réel a_i par les relations :

$$a_i^* \cdot a_j = 2\pi \delta_{ij} \quad (2)$$

Comme de plus $E(k)$ est une fonction paire de k , ceci réduit l'étude de $E(k)$ à un volume correspondant à la moitié d'une période symétrique par rapport à l'origine. Cette période est la première zone de Brillouin. La fonction $E(k)$ est une fonction continue du vecteur d'onde k . Cependant l'existence de conditions aux limites en surface du cristal implique une quantification du vecteur k . Avec les conditions aux limites de Born-Von Karman, les valeurs possibles de k sont données par :

$$\psi_k(r + N_i a_i) = \psi_k(r) \quad (3)$$

où N_i est le nombre de mailles élémentaires du cristal dans la direction définie par le vecteur de base a_i : Avec (2) et (3), ceci implique que les valeurs de k sont données par

$$k = \sum_i \frac{n_i}{N_i} a_i^* \quad (4)$$

où les n_i sont des nombres entiers. Ceci montre que, lorsque $N_i \rightarrow \infty$, les vecteurs k possibles forment un pseudo-continuum ainsi que les énergies $E(k)$ correspondantes. Ceci représente une preuve mathématique générale de l'existence de bandes d'énergie dans le cas de cristaux parfaits.

I.3. Distinction métal, isolant, semiconducteur

La situation la plus simple correspond au cas où le problème peut se réduire à une équation de Schrödinger monoélectronique. L'état fondamental du système à 0 K s'obtient alors en remplissant les états les plus bas en énergie, à raison de deux électrons

de spin opposé par niveau, pour tenir compte du principe d'exclusion de Pauli. La ligne de démarcation entre niveaux pleins et niveaux vides définit le niveau de Fermi E_F dont la position ne dépend pas de la taille du cristal (en effet, si le nombre d'électrons est proportionnel à la taille du cristal, il en est de même pour le nombre de niveaux dans un intervalle d'énergie donné).

A 0 K, deux cas peuvent se présenter :

- E_F tombe dans une bande permise (figure 2a). Dans ce cas, tout apport d'énergie externe, d'origine électrique ou optique, même très faible, permet de modifier l'état du cristal. On a un comportement métallique.

- E_F tombe dans une bande interdite (figure 2b) située entre la dernière bande pleine (bande de valence) de maximum E_v et la première bande vide (bande de conduction) de minimum E_c . Toute modification de l'état du cristal demande un apport d'énergie supérieur à la largeur de la bande interdite (ou gap) égale à $E_g = E_c - E_v$. Le système est isolant.

Lorsque $T \neq 0$ K, l'occupation des niveaux est donnée par la fonction de Fermi $(1 + \exp E - E_F/kT)^{-1}$ qui passe de 1 à 0 sur un intervalle de quelques kT . Pour un métal, ceci ne change pratiquement rien puisqu'il est toujours possible d'utiliser des niveaux d'excitation extrêmement faibles. Par contre, dans le deuxième cas, l'apport d'énergie thermique permet de faire passer des électrons de la bande de valence à la bande de conduction, créant ainsi des paires électron-trou. Il devient alors possible d'exciter chacune de ces particules (électrons dans BC, trous dans BV) comme dans le cas d'un métal. Si le gap est grand ($> 2e$ V), la concentration en paires électron-trou est négligeable et le système reste isolant. Dans le cas contraire, le système aura un comportement intermédiaire entre isolant et métal et sera un semiconducteur.

II. Calcul de la structure de bandes des semiconducteurs cristallins

Nous allons d'abord discuter rapidement les méthodes "ab initio" ainsi que leurs limites. Nous présenterons ensuite deux

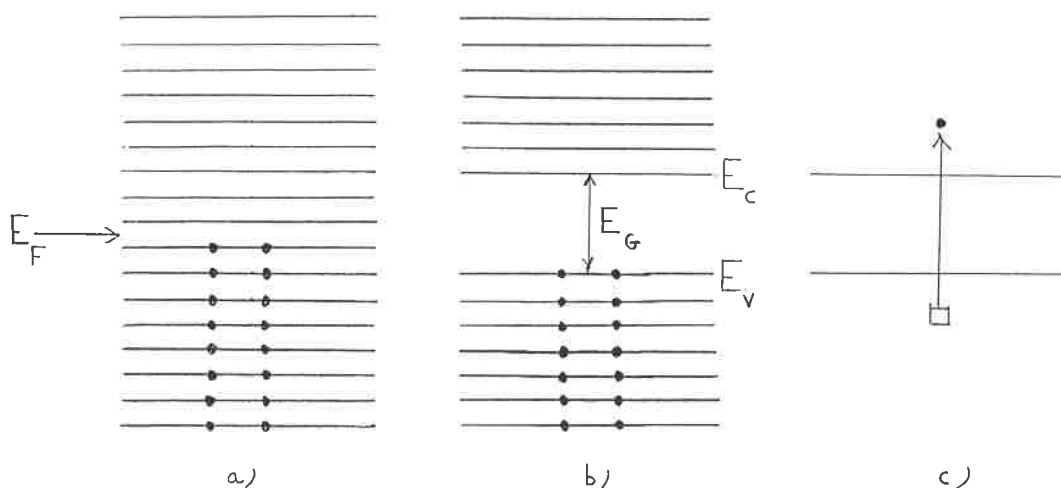


FIGURE 2. - Remplissage des niveaux monoélectroniques : a) pour un métal, b) pour un isolant, c) excitation de paires électron-trou.

méthodes empiriques : pseudopotentiel et liaisons fortes. Nous donnerons ensuite quelques applications aux semiconducteurs de structure zinc-blende, en couches et en chaînes.

II.1. Méthodes "ab initio"

En ce qui concerne les structures de bandes, l'énorme majorité de ces méthodes se fondent sur la séparation des variables électroniques permettant de ramener le problème à des équations de Schrödinger individuelles, du type suivant :

$$\left(-\frac{\hbar^2}{2m} \Delta + V_N + V_H + \Sigma \right) \psi_i = \epsilon_i \psi_i \quad (5)$$

où V_N est le potentiel dû aux noyaux, V_H le potentiel répulsif exercé par l'ensemble des électrons sur l'électron considéré (potentiel de Hartree) et où Σ est un opérateur contenant les corrections à apporter aux termes précédents. Bien sûr, une telle séparation des variables ne peut pas se faire de manière exacte et il existe diverses approximations correspondant à des formes différentes de Σ :

- Hartree, avec $\Sigma = 0$,
- Hartree-Fock où Σ correspond aux interactions d'échange,
- Densité locale où Σ est remplacé par un potentiel local V_{XC} correspondant à celui du gaz d'électrons libres de même densité électronique locale,
- GW, où Σ prend une forme plus complexe intermédiaire entre les deux précédentes [1].

La méthode actuellement la plus populaire est celle de la densité locale, plus performante que Hartree-Fock, tout en étant beaucoup plus simple à mettre en œuvre. Son avantage est de produire des résultats très précis pour les propriétés structurales, un peu moins pour les énergies de liaison. Par contre, elle pose plus de problèmes en ce qui concerne la prédiction des bandes interdites (0,6 au lieu de 1,2 eV pour Si, 0 pour Ge). Il a été montré récemment [2, 3], mais au prix de calculs très lourds, que l'approximation GW apporte une amélioration très substantielle de ce dernier point de vue.

En ce qui concerne les propriétés des électrons de valence, il n'est pas nécessaire d'utiliser le potentiel V_N réel, mais il est possible de le remplacer par un pseudopotentiel à variations plus douces, éliminant les effets de cœur, ce qui rend la convergence des calculs beaucoup plus rapide. Il existe même maintenant des tables de pseudopotentiels "premiers principes" [4] permettant des calculs "ab initio".

II.2. Méthodes semi-empiriques

De manière générale, l'interprétation de la structure électronique d'un solide s'effectue d'abord sur la base de modèles simpli-

fiés de nature semi-empirique. Il existe essentiellement deux approches de ce type : la méthode du pseudopotentiel empirique et l'approximation des liaisons fortes.

La méthode du pseudopotentiel empirique consiste à remplacer le potentiel $V_N + V_H + \Sigma$ par une somme de pseudopotentiels atomiques (dont les effets du cœur sont éliminés). Comme ceux-ci varient assez doucement dans l'espace, la méthode se prête naturellement à un développement de la fonction d'onde dans une base d'ondes planes. Chaque pseudopotentiel atomique intervient alors par ses composantes de Fourier ou facteurs de forme $v_{at}(\mathbf{G})$ où les \mathbf{G} sont les vecteurs du réseau réciproque. La fonction $v_{at}(\mathbf{G})$ décroît rapidement en fonction de $|\mathbf{G}|$ et la méthode empirique est de tronquer aussi rapidement que possible cette fonction, considérant les quelques $v_{at}(\mathbf{G})$ retenus comme autant de paramètres ajustables. En pratique, la structure de bandes du silicium est très bien décrite à partir de trois facteurs de forme.

La seconde méthode empirique est fournie par l'approximation des liaisons fortes. Le principe est de développer la fonction d'onde suivant une base formée des fonctions des atomes libres. Toutefois, on est obligé de limiter le nombre de fonctions de base et on se restreint habituellement à celles de la couche externe des atomes considérés (base minimale). Les solutions sont donc données par l'équation séculaire :

$$\det(H - ES) = 0 \quad (6)$$

où H est la matrice de l'hamiltonien et S celles des recouvrements $\langle \Phi_i | \Phi_j \rangle$ entre orbitales atomiques Φ_i et Φ_j . L'approximation dite des "liaisons fortes" consiste à négliger ces recouvrements et donc à remplacer S dans (6) par la matrice unité I . Le problème se réduit donc à une diagonalisation de la matrice H . L'aspect empirique de la méthode intervient dans la détermination des éléments de cette matrice. Ceux-ci décroissent assez rapidement en fonction de la distance interatomique et une approximation supplémentaire consiste à tronquer la matrice dans l'espace en se limitant aux interactions aux premiers, seconds ou troisièmes voisins. Les éléments de matrice retenus sont ensuite ajustés de manière à reproduire le mieux possible la structure de bandes expérimentale. Une telle méthode n'a évidemment d'intérêt que si elle permet une interprétation simple (ce qui est le cas) et si on arrive à trouver des règles générales pour les paramètres de manière à pouvoir effectuer des prédictions. C'est le cas des règles de Harrison [5] qui permettent l'application à de nombreux matériaux.

Les deux méthodes ont leurs mérites respectifs. Celle du pseudopotentiel permet d'obtenir assez facilement une bonne description de la structure de bandes des semiconducteurs. L'approximation des liaisons fortes décrit facilement la bande de valence mais beaucoup moins les bandes de conduction même les plus basses. Son avantage essentiel est de fournir une description très simple de la formation des bandes à partir de la structure des atomes libres, éventuellement par l'intermédiaire de modèles moléculaires comme nous le verrons plus loin. C'est pourquoi nos dis-

cussions physiques seront dans la suite essentiellement axées sur cette méthode.

II.3. Semiconducteurs de type zinc-blende

Il s'agit de systèmes élémentaires de la colonne IV et de composés III-V ou II-VI où les atomes ont une coordinance tétraédrique. La base minimale de liaisons fortes est composée d'une orbitale s et de trois orbitales p. On pourrait diagonaliser directement la matrice de l'hamiltonien dans cette base et obtenir ainsi la structure de bandes. Il est cependant plus judicieux du point de vue de la compréhension physique d'effectuer un changement de base tel que, dans la nouvelle base, certains éléments de matrice vont devenir beaucoup plus grands que les autres. Ceci permettra de procéder par étapes en traitant d'abord les éléments dominants et en introduisant ensuite les autres comme des corrections. C'est la base générale des modèles moléculaires.

Le changement de base naturel est ici de bâtir des orbitales sp^3 de la forme :

$$\varphi_{ij} = \frac{\varphi_{s,i} + \sqrt{3}\varphi_{p,ij}}{2} \quad (7)$$

où $\varphi_{s,i}$ est la fonction s de l'atome i et $\varphi_{p,ij}$ l'une de ses orbitales p pointant vers le premier voisin j. Si l'on procède ainsi pour tous les atomes, chaque liaison du système est caractérisée par une paire d'orbitales hybridées $sp^3\varphi_{ij}$ et φ_{ij} se recouvrant fortement (figure 3). L'élément de matrice dominant de l'hamiltonien sera donné par

$$\langle \varphi_{ij} | H | \varphi_{ij} \rangle = -\beta \quad \beta > 0 \quad (8)$$

alors que les éléments diagonaux pourront s'écrire

$$\langle \varphi_{ij} | H | \varphi_{ij} \rangle = E_A \quad (9)$$

E_A et E_B étant les énergies moyennes des atomes A ou B dans l'état sp^3 . La première étape consiste à ne retenir que les éléments de matrice définis ci-dessus. Le problème est alors équivalent à un ensemble de molécules biatomiques, chacune conduisant à un niveau liant et antiliant donné par :

$$E_{AL} = \frac{E_A + E_B}{2} \pm \sqrt{\left(\frac{E_A - E_B}{2}\right)^2 + \beta^2} \quad (10)$$

Ces états ont une dégénérescence égale au nombre N de liaisons dans le solide. Pour une liaison entre atomes i et j, la fonction d'onde de ces états est donnée par :

$$\psi_{L,ij} = \frac{\varphi_{ij} + \lambda \varphi_{ij}}{\sqrt{1 + \lambda^2}} \quad (11)$$

$$\psi_{AL,ij} = \frac{\lambda \varphi_{ij} - \varphi_{ij}}{\sqrt{1 + \lambda^2}}$$

Comme il y a deux électrons par liaison, l'état fondamental du système correspond à des états liants pleins et des états antiliants vides. Cette description définit le modèle moléculaire de ces matériaux, qui a permis l'interprétation de nombreuses propriétés physiques [5].

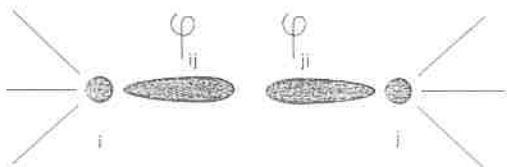


FIGURE 3. - Paire d'orbitales hybridées sp^3 impliquées dans une liaison.

L'incorporation d'éléments de matrice de l'hamiltonien négligés jusqu'ici peut se faire de manière simple. Pour les raisons discutées dans la section I, les interactions entre les N états liants dégénérés vont lever la dégénérescence et provoquer la formation d'une bande à caractère liant. Il en est de même pour les états antiliants. Le cas intéressant est celui où ces deux bandes ne se recouvrent pas et sont donc séparées par une bande interdite. Les interactions entre états liants et antiliants provoquent alors une répulsion mutuelle de ces deux bandes ayant pour effet de conserver sinon d'augmenter la bande interdite. Le schéma final est celui de la figure 4, avec une bande de valence à caractère liant remplie et une bande de conduction à caractère antiliant vide. L'élargissement en bandes sera analysé un peu plus en détail pour les semiconducteurs amorphes (section III.1).

Les conclusions qualitatives, développées ci-dessus, sont en accord complet avec les résultats de calculs détaillés. Un exemple de structure de bandes est celle du silicium, représentée figure 5.

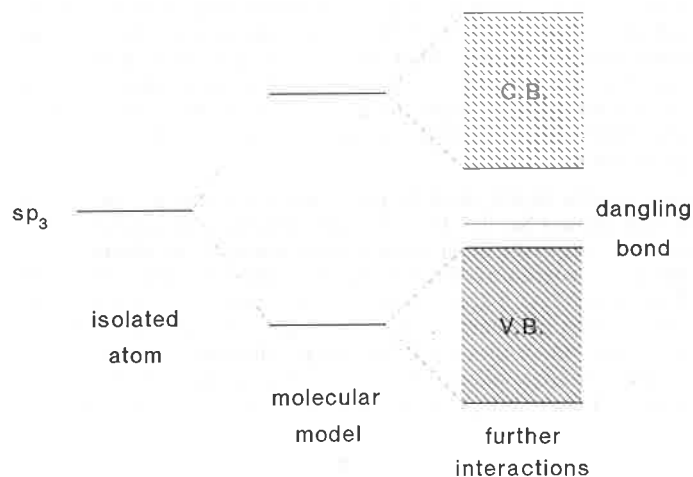


FIGURE 4. - Elargissement des niveaux liants et antiliants en bandes de valence et de conduction.

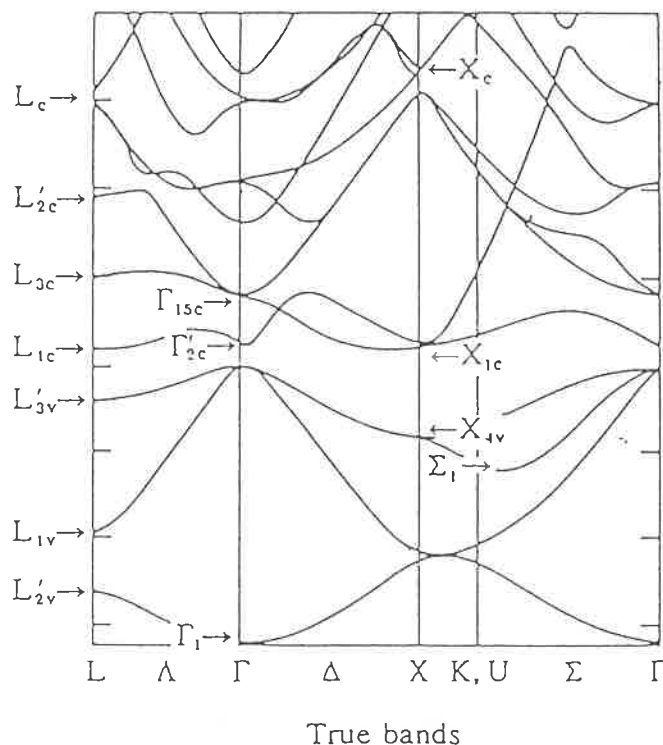


FIGURE 5. - Exemple de structure de bandes du silicium.

II.4. Semiconducteurs en chaînes ou en couches

Nous considérons ici des systèmes où chaque atome forme N liaisons équivalentes avec ses premiers voisins. Comme pour les cristaux tétraédriques, il est intéressant de bâtir un modèle moléculaire qui fournit une base pour la compréhension de leur structure de bandes. Sur chaque atome, on bâtit N orbitales équivalentes qui pointent exactement ou de manière approchée vers les premiers voisins. Comme nous considérons une base minimale sp de quatre orbitales, ceci est toujours possible, puisque les systèmes considérés correspondent à $N < 4$. Les états atomiques restants sont alors choisis en tenant compte de la symétrie locale (des exemples détaillés sont donnés dans la suite).

Une fois ceci effectué, on obtient directement la structure électronique. A nouveau, les orbitales dirigées se couplent fortement par paires comme dans les molécules biatomiques et donnent lieu à des états liants σ et antiliants σ^* . C'est ce couplage qui domine la matrice de l'hamiltonien et les propriétés de cohésion. A ce premier stade d'approximation, tous les autres états restent non couplés à leur valeur atomique. La structure de niveaux est formée de $N/2$ états liants σ et $N/2$ états antiliants σ^* plus $4-N$ états non liants par atome. Bien sûr, ces niveaux moléculaires sont tous fortement dégénérés et, si le modèle moléculaire est valable, les interactions supplémentaires auront pour effet de lever cette dégénérescence pour donner naissance à des bandes d'énergie séparées.

Un exemple simple de cette procédure est constitué par les solides en chaînes Se, Te. Pour discuter ces cas, nous considérons la structure idéalisée de la figure 6 qui conduit à une situation proche de la structure réelle, l'angle interliaison étant pris égal à 90° . Avec le système d'axes locaux définis figure 6, les orbitales dirigées vers les premiers voisins seront des états "p" purs, notés p_x et p_y . Les états atomiques p_z , perpendiculaires au plan de la chaîne, resteront découplés des autres et donneront un état atomique à l'énergie E_p . Finalement, les états "s", d'énergie beau-

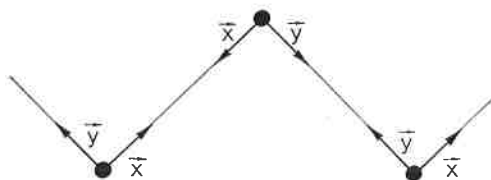


FIGURE 6. - Modèle idéalisé de structure en chaînes pour Se et Te.

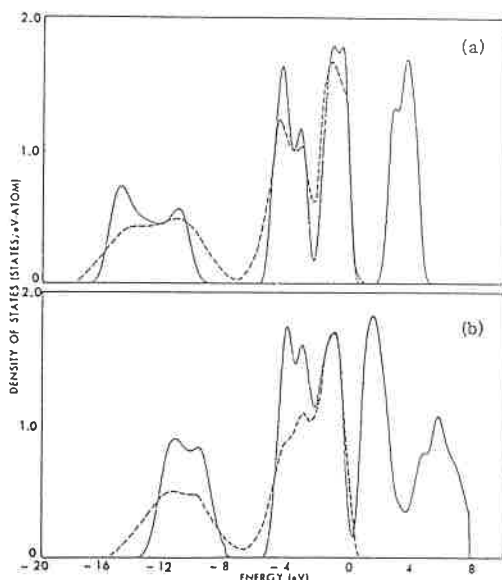


FIGURE 7. - Comparaison théorie-expérience pour les densités d'états de Se et Te, chaque sous-bande dérivant d'un état du modèle moléculaire discuté dans le texte.

coup plus basse, donneront un état atomique à l'énergie E_s . Le calcul détaillé montre que ce modèle moléculaire trivial correspond directement à la formation des bandes et décrit bien les résultats expérimentaux de la figure 7.

Des considérations similaires s'appliquent aux composés cristallins GeSe et GeTe. Ici $N = 3$, les angles entre liaisons étant aussi proches de 90° . Les orbitales dirigées seront donc les trois orbitales p qui donnent lieu aux états liants et antiliants. Il reste donc un état non liant s par atome donnant naissance à une bande correspondante d'énergie plus basse. A nouveau, ce modèle simple trouve une confirmation expérimentale et théorique.

III. Semiconducteurs non cristallins et désordonnés

De notre discussion, section I, il ressort qu'il est vraisemblable que ces matériaux donnent lieu à des bandes d'énergie. Cependant, il n'est pas évident que les limites de bandes soient aussi bien définies que dans les cristaux. En particulier, on s'attend à ce que le désordre fasse apparaître des queues de bandes.

III.1. Semiconducteurs amorphes

Suivant une méthode analogue à celle de la réf. [6], nous considérons un modèle idéalisé d'un amorphe covalent tel que a-Si pour montrer l'existence de bandes bien définies. La structure atomique est celle d'un réseau continu aléatoire qui préserve l'environnement tétraédrique de chaque atome et minimise les distorsions radiales et angulaires. La limite idéale, que nous considérons dans un premier temps, consiste à supposer que ces distorsions sont nulles. La structure électronique peut alors se discuter aisément en liaisons fortes, à partir des mêmes arguments que dans la section II.3. On commence par former des états liants et antiliants (définis par les équations (10) et (11) avec $\lambda = 1$ pour les covalents). Ce qui nous intéresse maintenant, c'est d'étudier la formation des bandes à partir de ce modèle moléculaire pour montrer qu'il existe encore une bande interdite. Considérons d'abord l'élargissement en bandes des niveaux moléculaires liants d'énergie E_L . L'élément de matrice le plus important est alors l'interaction $\langle \psi_{Lij} | H | \psi_{Llk} \rangle = \Delta$ entre deux états antiliants adjacents. L'élargissement de la bande liante s'obtient alors en écrivant la fonction d'onde :

$$\psi = \sum a_{L,ij} \psi_{L,ij} \quad (12)$$

et en projetant l'équation de Schrödinger sur chacun des états de base. Ceci donne :

$$(E - E_L) a_{L,ij} = \Delta \left\{ \sum_{j' \neq j} a_{L,ij'} + \sum_{i' \neq i} a_{L,i'j} \right\} \quad (13)$$

où les sommes se font sur les liaisons adjacentes à la liaison ij . Il est intéressant d'introduire les inconnues auxiliaires :

$$S_i = \sum_j a_{L,ij} \quad (14)$$

ce qui permet d'écrire (13) sous la forme :

$$(E - E_L + 2\Delta) a_{ij} = \Delta (S_i + S_j) \quad (15)$$

qui, sommée sur j , donne

$$(E - E_L - 2\Delta) S_i = \Delta \sum_j S_j \quad (16)$$

Il est intéressant d'écrire ces équations sous la forme :

$$\epsilon S_i = \sum_j S_j \quad (17)$$

Les valeurs propres des deux matrices sont donc reliées par :

$$E = E_L + 2\Delta + \Delta\varepsilon \quad (18)$$

La matrice définie par l'équation (17), avec des éléments égaux à 1 entre premiers voisins et 0 partout ailleurs, est appelée matrice connectivité du système. Des théorèmes topologiques généraux permettent de montrer que ses valeurs propres sont situées entre -4 et $+4$. Avec (18), ceci donne :

$$E_L + 2\Delta - 4|\Delta| \leq E \leq E_L + 2\Delta + 4|\Delta| \quad (19)$$

L'équation (18) ne donne pas toutes les solutions du problème. Il existe en particulier des solutions telles que tous les S_i soient nuls. D'après (15), ceci est possible ($a_{ij} \neq 0$) si :

$$E = E_L - 2\Delta \quad (20)$$

ce qui donne une bande plate (sans dispersion) à l'une des extrémités de l'intervalle défini par (19).

Le résultat essentiel que nous obtenons est donc que tous les états permis se trouvent dans l'intervalle donné par (19). Ce raisonnement ne dépend pas du tout de l'ordre à longue distance et est valable pour notre modèle idéalisé de a-Si comme pour c-Si. Le même raisonnement s'applique à la bande dérivée des états antiliants. Donc, dans ce modèle, la bande interdite entre les bandes liante et antiliante est au moins aussi large dans l'amorphe que dans le cristal. L'incorporation des interactions entre états liants et antiliants ne peut qu'élargir cette bande interdite par répulsion mutuelle des deux bandes. Bien sûr, cette conclusion ne vaut que pour le modèle idéalisé de l'amorphe. Les distortions de longueurs et d'angles de liaisons produiront un certain désordre qui doit donner lieu à des queues de bandes s'étendant dans la bande interdite.

III.2. Alliages désordonnés

Nous ne discuterons ce cas que brièvement, une analyse beaucoup plus détaillée étant donnée dans la réf [7]. Les systèmes les plus simples de ce genre correspondent à des composés où le réseau est bien défini, mais où il existe un désordre de substitution sur les sites de ce réseau. C'est le cas des alliages semiconducteurs pseudobinaires $Ga_xAl_{1-x}As$, $In_xGa_{1-x}As$,... Le problème que pose leur structure électronique est assez compliqué et requiert l'utilisation de diverses approximations.

La plus simple d'entre elles est celle du "cristal virtuel". On définit un hamiltonien moyen où le potentiel est celui d'atomes fictifs résultant de la pondération statistique des atomes constituants. Par exemple, dans $Ga_xAl_{1-x}As$, le cation moyen a un potentiel égal à x fois celui de Ga plus $(1-x)$ fois celui de Al. L'hamiltonien résultant a la symétrie de translation, ce qui permet d'employer les méthodes usuelles de calcul de structures de bandes cristallines, habituellement celle du pseudopotential empirique ou des liaisons fortes.

L'approximation du cristal virtuel néglige les fluctuations locales par rapport à l'hamiltonien moyen. On peut traiter ces effets par différentes méthodes, les plus courantes étant l'approximation de la matrice de transfert moyenne (ATA), et celle du potentiel cohérent (CPA [8]). Toutes deux sont des approximations de type champ moyen ne différant que par le fait que la seconde est autocohérente. La technique est de passer par les densités d'états locales qui représentent, par exemple, la contribution d'un atome (ou d'une cellule) à la densité d'états totale. Considérons, par exemple, un système $A_xB_{1-x}C$. L'ATA procède de la manière suivante : on calcule la densité d'états locale sur l'atome A, $n_A(E)$, en supposant cet atome connecté à un milieu moyen défini par le cristal virtuel ; on détermine de la même façon $n_B(E)$ pour l'atome B. La densité d'états moyenne s'obtient alors comme :

$$\langle n(E) \rangle = x n_A(E) + (1-x) n_B(E) \quad (21)$$

Pour obtenir la CPA, on procède de la même manière, mais en modifiant la définition du milieu moyen auquel seront connectés les cations A ou B. Celui-ci n'est plus le cristal virtuel mais est un milieu effectif dans lequel les cations ont la même densité d'états $\langle n(E) \rangle$ que celle que l'on cherche à calculer. Donc lorsqu'on applique (21) au calcul de $\langle n(E) \rangle$ sur un site de cation, les

quantités $n_A(E)$ et $n_B(E)$ dépendent du milieu moyen auquel A et B sont connectés et donc aussi de la quantité $\langle n(E) \rangle$ que l'on cherche à déterminer. Les équations de la CPA doivent donc être résolues de manière autocohérente.

La figure 8 illustre les différences entre une approximation de type CPA et de cristal virtuel pour les pseudobinaires de structure zinc-blende [7]. Il est évident que certaines caractéristiques spectrales nécessitent l'emploi de la CPA.

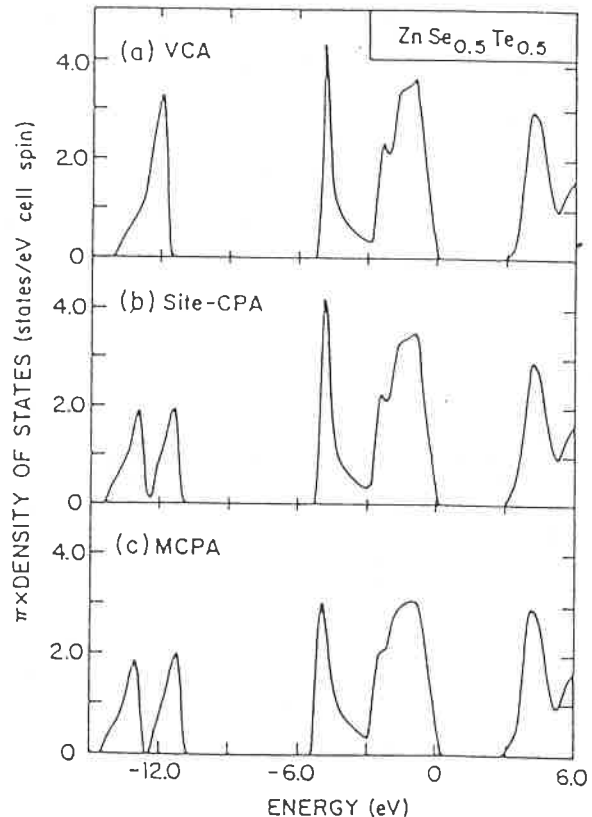


FIGURE 8. - Comparaison de l'approximation du cristal virtuel (VCA) et de deux approximations de type CPA pour l'alliage $ZnSe_{0.5}Te_{0.5}$, tirée de la réf. [7].

IV. Utilisation de la masse effective

Une classe de problèmes extrêmement importante dans les semi-conducteurs correspond à l'application d'un potentiel lentement variable dans l'espace. C'est le cas des donneurs et accepteurs dans ces matériaux, des excitons ainsi que des potentiels macroscopiques comme ceux existant dans de nombreux composants pour la microélectronique. Dans la plupart de ces cas, il est possible d'utiliser l'approximation de la masse effective ou encore de la fonction enveloppe qui simplifie notablement le problème.

IV.1. Donneurs et accepteurs

Il s'agit d'impuretés de substitution ayant un électron de valence en plus (donneur) ou en moins (accepteur) que l'atome remplacé. Prenons le cas du donneur simple, par exemple As dans du germanium. L'atome As a cinq électrons sur sa couche externe au lieu de quatre pour Ge. Sur ces cinq électrons, quatre vont se trouver dans des états liants tels que nous les avons discutés plus haut, c'est-à-dire dans la bande de valence. Le problème est donc de connaître l'énergie de l'électron supplémentaire. Le potentiel vu par celui-ci est celui du germanium cristallin plus une charge

+ e sur le noyau As. Il est donc égal à $-e^2/\epsilon r$, où ϵ est la permittivité diélectrique de Ge. Il s'agit donc d'un potentiel coulombien attractif à variations relativement lentes. Son effet sera de faire sortir des états du bas de la bande de conduction. Une très bonne approximation de la fonction d'onde de ces états est alors de l'écrire sous la forme [9].

$$\psi(\mathbf{r}) = \Phi_{k_m}(\mathbf{r}) \cdot F(\mathbf{r}) \quad (22)$$

où Φ_{k_m} est la fonction du bas de bande de conduction du cristal parfait et $F(\mathbf{r})$ la fonction "enveloppe", à variations lentes dans l'espace. On peut montrer que l'équation obéie par $F(\mathbf{r})$ est une équation de Schrödinger effective où il suffit de remplacer la masse de l'électron par des masses effectives caractéristiques de la courbure des bandes $E(\mathbf{k})$ au voisinage du bas de la bande de conduction. Avec un potentiel en $-e^2/\epsilon r$, on obtient alors un problème hydrogénoïde donnant pour l'état fondamental 1s :

$$E = E_c - \frac{m^* e^4}{2 \hbar^2 \epsilon^2} \quad (23)$$

$$F_{1s}(\mathbf{r}) = A \exp - \frac{r}{a_0^*} \quad a_0^* = \frac{\hbar^2 \epsilon}{m^* e^2}$$

Avec les valeurs typiques $m^* = 0,1$ et $\epsilon \sim 10$, le niveau fondamental est à 13,6 meV, très proche du bas de la bande de conduction E_c et le rayon de Bohr $a^* = 53 \text{ \AA}$ est grand devant la distance interatomique. A basse température, l'électron supplémentaire se trouve donc dans l'état hydrogénoïde défini par (23). A température ambiante, par contre, où $kT \sim 25 \text{ meV}$, l'énergie thermique est suffisante pour que cet électron soit ionisé dans la bande de conduction. Celle-ci, au lieu d'être vide, comporte alors autant d'électrons qu'il y a d'atomes donneurs. Le comportement se rapproche d'un métal : le semiconducteur est dopé de type n.

Le raisonnement symétrique peut s'effectuer pour des atomes accepteurs ayant un électron de moins (ex : Ga dans Ge). On peut alors considérer que l'on a une bande de valence pleine plus un manque d'électron ou "trou" de charge + e attiré par la charge en excès - e du noyau accepteur. Ceci introduit des niveaux hydrogénoïdes dans le gap, proches du sommet de la bande de valence. A T = 0 K les trous sont sur le niveau hydrogénoïde le plus stable. A l'ambiante, ils sont ionisés dans la bande de valence. La situation est symétrique et le semiconducteur est dopé de type p.

IV.2. Excitons

L'absorption optique dans un semiconducteur se produit au-dessus d'un certain seuil tel que $h\nu > E_g$, la largeur de bande interdite. Lors de l'absorption d'un photon, un électron est excité de la bande de valence vers la bande de conduction. Il y a création d'une paire électron-trou, le trou étant une particule fictive, c'est-à-dire un manque d'électron de charge + e. Cependant, dans l'espace, l'électron et le trou ainsi créés ne sont pas indépendants mais s'attirent par un potentiel coulombien $- e^2/\epsilon r$. Ce potentiel coulombien doit donner lieu à des états liés de la paire électron-trou. Ces états liés peuvent s'obtenir par une simple extension de l'approximation de la masse effective mais appliquée maintenant à des paires de particules.

Le principe est d'écrire la fonction d'onde totale du système sous la forme

$$\psi = \psi_0 \cdot F(\mathbf{r}_e, \mathbf{r}_h) \quad (24)$$

où ψ_0 est l'état excité du semiconducteur (le plus bas) et $F(\mathbf{r}_e, \mathbf{r}_h)$, la fonction enveloppe décrivant le mouvement combiné de l'électron et du trou. La fonction enveloppe satisfait alors une équation de Schrödinger effective (voir réf. [10] pour une dérivation plus détaillée)

$$\left(-\frac{\hbar^2}{2m_e} \Delta_e - \frac{\hbar^2}{2m_h} \Delta_h - \frac{e^2}{\epsilon r_{eh}} \right) F(\mathbf{r}_e, \mathbf{r}_h) = (E - E_g) F(\mathbf{r}_e, \mathbf{r}_h) \quad (25)$$

où m_e, m_h sont les masses effectives de l'électron et du trou. On peut séparer mouvement du centre de masse et mouvement relatif. L'état fondamental correspond à une énergie de translation nulle et à l'état 1s du mouvement relatif pour lequel la masse est la masse réduite. On obtient donc des pics d'absorption optique à des énergies égales à E_g moins les énergies de liaison des états hydrogénoïdes du mouvement relatif.

V. Systèmes à dimensionalité réduite

Grâce aux progrès des techniques de croissance, il est maintenant possible de fabriquer des structures à semiconducteurs ayant un caractère bi- ou unidimensionnel et même, plus récemment, des boîtes quantiques (OD). Ces structures présentent un intérêt à la fois sur le plan de la physique et des composants. Nous développons ici rapidement quelques aspects de ces systèmes.

V.1. Aspects qualitatifs

Partons d'un système simple 2D formé de couches de GaAs et GaAlAs alternées. Le gap de GaAlAs est plus grand que celui de GaAs, la différence se reportant à la fois sur la bande de conduction et de valence. On a donc la situation schématique de la figure 9, c'est-à-dire la formation de puits carrés de potentiel le long de l'axe de croissance. En supposant que l'on puisse traiter les électrons et trous dans une approximation de masse effective, on aboutit à un problème trivial de mécanique quantique à une dimension (sauf pour le mouvement libre dans les directions parallèles aux couches). Il en est de même pour les fils ou boîtes quantiques.

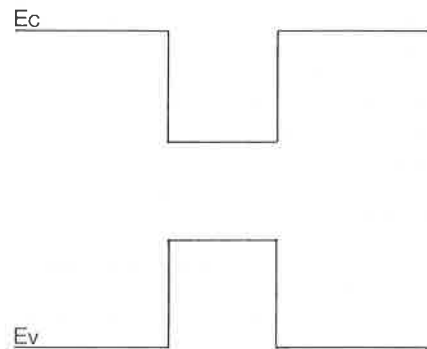


FIGURE 9. - Discontinuités des bandes de valence et de conduction pour une couche de GaAs entourée de GaAlAs.

Considérons donc les électrons de la bande de conduction, par exemple, et discutons les caractéristiques générales de la densité d'états suivant le nombre de dimensions. Pour un système confiné dans 3-d directions et libre dans les d restantes, les énergies permises sont la somme de niveaux E_n d'un puits à 3-d dimensions et de E_k énergie d'une particule libre de vecteur d'onde k à d dimensions. La densité d'états $n(E)$ d'énergie E est, par définition :

$$n(E) = \sum_{n,k} \delta(E - E_n - E_k) \quad (26)$$

qu'on peut réécrire :

$$n(E) = \sum_n \int dE' \delta(E - E_n - E') \sum_k \delta(E' - E_k) \quad (27)$$

La quantité $\sum_k \delta(E' - E_k)$ définit la densité d'états $n_d(E')$ du gaz

d'électrons libres à d dimensions (définie pour $E' > 0$). Ceci permet d'écrire :

$$n(E) = \sum_n n_d (E - E_n) \quad (28)$$

Pour un gaz d'électrons libres à d dimensions, il est bien connu que $n_d(E) \propto E^{(d/2-1)}$ ce qui donne :

$$n(E) \propto \sum_n (E - E_n)^{d/2-1} \quad (29)$$

L'allure de $n(E)$ est représentée *figure 10* pour le volume, les puits, fils et boîtes quantiques et montre des comportements nettement différents.

On peut passer graduellement d'un comportement 2D à 3D par exemple. Il suffit d'augmenter l'épaisseur des couches de telle façon que les E_n forment un pseudo-continuum. On peut aussi fabriquer des superréseaux avec des puits équidistants et suffisamment proches pour les E_n s'élargissent en bandes 1D. Les échelons de la densité d'états de la *figure 10* s'adoucissent alors et $n(E)$ acquiert un caractère 3D. Tous ces cas représentent des illustrations particulièrement nettes des considérations générales sur la formation des bandes.

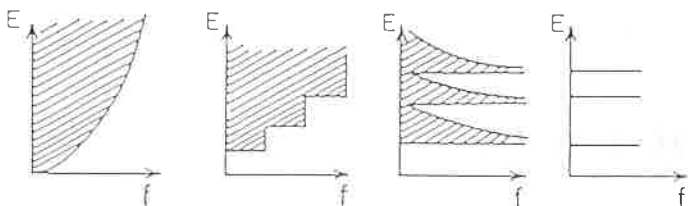


FIGURE 10. - *Forme des densités d'états en fonction de la dimensionnalité du système.*

V.2. Utilisation de la fonction enveloppe

La structure électronique de telles structures est très complexe, en particulier quand on inclut des effets de charge d'espace auquel cas le potentiel présente non seulement des discontinuités mais aussi une courbure macroscopique (ex. : hétérojonctions p-n). Heureusement, dans les régions où le potentiel est lentement variable et où l'énergie des porteurs est proche d'un extremum de bandes, il est possible de simplifier le traitement par l'approximation de la masse effective discutée auparavant. La fonction d'onde dans le matériau considéré s'exprimera donc sous la forme générale (22). Dans le cas simple d'une interface entre deux matériaux A et B, par exemple, la fonction enveloppe prendra la forme simple

$$F(\mathbf{r}) = f(z) e^{ik_{\parallel} \cdot \mathbf{r}_{\parallel}} \quad (30)$$

le symbole \parallel correspondant aux quantités parallèles à l'interface. La fonction $f(z)$ prend des formes différentes dans le matériau A ($f_A(z)$) et B ($f_B(z)$). Chacune de ces fonctions est solution d'une équation de Schrödinger à une dimension différente avec des masses effectives m_A^* et m_B^* . Le problème se pose alors de raccorder correctement les deux fonctions de part et d'autre de l'interface. Habituellement, on assure la continuité de $f(z)$ et de $1/m^* df/dz$, ce qui conduit automatiquement à la continuité du courant. Toutefois la validité de ces conditions aux limites n'est pas garantie et tout ce qu'on peut montrer c'est qu'il doit exister une relation linéaire pour $f(z)$ et df/dz , de part et d'autre de l'interface, par l'intermédiaire d'une matrice de transfert. Celle-ci doit alors être déterminée directement par un calcul microscopique détaillé (un exemple et des références sont données en [10]).

Les applications de la fonction enveloppe à ces systèmes sont nombreuses. En premier lieu, on peut citer le calcul de la structure électronique des puits, superréseaux, fils et boîtes quantiques. Ceci s'avère relativement évident pour la bande de conduction (une fois précisées les conditions de raccordement), moins

pour la bande de valence à cause de la dégénérescence au maximum de celle-ci (voir [11] pour plus de détails).

Un autre problème intéressant est celui du dopage de telles structures. En effet, on sait que l'énergie de liaison d'un système hydrogénoïde est quatre fois plus grande à deux dimensions qu'à trois. Il y aura donc une évolution notable des propriétés des donneurs et accepteurs dans un puits quantique suivant la largeur de celui-ci et la position de l'impureté dans le puits. Des études systématiques ont été entreprises pour calculer l'évolution de ces propriétés, à partir du formalisme de la fonction enveloppe (pour une revue voir [11]). Ceci a permis de calculer l'évolution de l'énergie de liaison de l'impureté en fonction des dimensions du puits quantique.

La situation des excitons est similaire mais, ici, le problème est compliqué par la dégénérescence du haut de la bande de valence, ce qui est aussi le cas des accepteurs. Le traitement est fondé sur les mêmes principes que précédemment, avec utilisation d'une méthode variationnelle [11]. On peut calculer, au moins dans des modèles simples, l'absorption optique liée à l'exciton et montrer qu'elle augmente lorsqu'on réduit le nombre de dimensions [10].

Toutes ces structures peuvent être placées dans des champs électrique, magnétique, etc. Les effets de dopage peuvent induire des régions de charge d'espace qui provoquent des variations macroscopiques de potentiel. On peut malgré tout encore calculer la structure électronique de ces systèmes, à condition d'appliquer la méthode de la fonction enveloppe et de résoudre numériquement les équations de Schrödinger ainsi obtenues.

VI. Applications

Comme nous l'avons vu, la propriété essentielle des semiconducteurs est d'avoir un comportement intermédiaire entre isolant et métal, que l'on peut moduler à volonté par la concentration de dopants (donneurs ou accepteurs). L'adjonction de donneurs, par exemple, provoque une conduction par les électrons ionisés dans la bande de conduction. Pour les accepteurs, on aura un effet symétrique avec les trous de la bande de valence. Dans les deux cas, si la concentration augmente, on observe une transition isolant-métal.

Il s'est avéré possible de tirer un parti extraordinaire de ces propriétés modulables. Le cas le plus simple est la jonction p-n où une partie d'un semiconducteur est dopée n (avec des donneurs) et l'autre p (accepteurs), l'interface, étant aussi abrupte que possible. Lorsqu'on applique une tension aux bornes d'une telle jonction, son comportement est évidemment asymétrique lorsqu'on inverse la polarisation. On peut montrer que l'asymétrie est telle que l'on aboutit à un effet redresseur, c'est-à-dire une diode. C'est aussi le cas d'un certain nombre de systèmes métal-semiconducteur dopé, lorsque certaines conditions sont réunies.

A ces composants simples se sont ajoutées une multitude d'autres possibilités. La première est le transistor bipolaire, correspondant à un sandwich p-n-p ou n-p-n. Une autre, non moins importante, est le transistor MOS (métal-oxyde-semiconducteur) dont le fonctionnement est fondé sur l'effet de champ. C'est ce dernier type de composant qui a permis l'intégration à haute échelle, les dimensions des composants élémentaires atteignant le micron ou moins. Dans le futur, ce sont d'ailleurs les effets physiques associés à la réduction des dimensions qui poseront des problèmes pour une intégration plus poussée. Notons que la majorité des circuits intégrés actuels sont à base de silicium, pour des raisons pratiques évidentes. D'autres composants encore plus sophistiqués sont apparus plus récemment, nécessitant des méthodes de croissance élaborées telles que l'EJM (épitaxie par jets moléculaires). Celles-ci permettent une croissance plan par plan et l'élaboration de composants multicouches faisant appel à des matériaux divers (essentiellement des combinaisons de com-

posés III-V). Ces composants sont utilisés pour des applications spécifiques diverses : hautes fréquences, propriétés optiques (lasers)...

Note : Une discussion plus détaillée des effets abordés ici est donnée en [10] avec de nombreuses références.

Bibliographie

[1] L. Hedin, S. Lundquist, *Solid State Physics*, **1969**, 23, 1.

[2] M.S. Hybertsen, S.G. Louie, *Phys. Rev. Lett.*, **1985**, 55, 1418, and *Phys. Rev. B*, **1986**, 34, 5390.

[3] R.W. Godby, M. Schlüter, L.J. Sham, *Phys. Rev. Lett.*, **1986**, 56, 2415 and *Phys. Rev. B*, **1987**, 35, 4170.

[4] D.R. Hamann, M. Schlüter, C. Chiang, *Phys. Rev. Lett.*, **1979**, 43, 1494 and G.B. Bachelet, D.R. Hamann, M. Schlüter, *Phys. Rev. B*, **1982**, 26, 4199.

[5] W.A. Harrison, *Electronic Structure and The Properties of Solids, The Physics of the Chemical Bond*, Freeman, New York, **1980**.

[6] D. Weaire, M.F. Thorpe, *Phys. Rev. B*, **1971**, 4, 2508.

[7] R.J. Lempert, K.C. Hass, H. Ehrenreich, *Phys. Rev. B*, **1987**, 36, 1111.

[8] J.M. Ziman, *Models of disorder*, Cambridge University Press, **1979**.

[9] Une revue est donnée dans : F. Bassani, G. Iadonisi, B. Preziosi, *Rep. Prog. Phys.*, **1974**, 37, 1099.

[10] M. Lannoo, "Band theory applied to semiconductors", to be published.

[11] G. Bastard, *Wave Mechanics applied to Semiconductors Heterostructures*, Les Editions de Physique, Paris, **1988**.



SCHWEIZERISCHE
CHEMISCHE GESELLSCHAFT
Verlag Helvetica Chimica Acta
Postfach, CH-4002 Basel

HELVETICA CHIMICA ACTA

Subscription

Vol. 74, 1991

sFr. 515.- + postage

32.- Europe

sFr. 515.- + postage

50.- Oversea

Still available

Reprinted editions

Vols 1-27 (1918-1944)

Vol. 28 Out of print

Original editions

Vols 29-73 (1946-1990)

Please request our
price list