

Année 2011

THESE

présentée devant

L'UNIVERSITE DE PROVENCE (AIX-MARSEILLE I)

Par

Nicolas PONS

en vue d'obtenir

LE GRADE DE DOCTEUR

Ecole doctorale : Mécanique, Physique, Micro- et Nano-électronique
Spécialité : Micro- et Nano-électronique

MODELISATION TRIDIMENSIONNELLE MULTIBANDES DU TRANSPORT QUANTIQUE DANS LES TRANSISTORS A NANOFIL

Directeur de thèse : Michel LANNOO
Co-encadrant : Nicolas CAVASSILAS

À soutenir le 8 Juin 2011 devant la commission d'examen :

Rapporteur	M. Arnaud BOURNEL	IEF, Orsay
Rapporteur	Mme. Mireille MOUIS	IMEP-LAHC, Grenoble
Examineur	M. Hervé JAOUEN	STMicroelectronics Crolles
Examineur	M. Marco PALA	IMEP-LAHC, Grenoble
Directeur de thèse	M. Michel LANNOO	IM2NP, Marseille
Co-encadrant	M. Nicolas CAVASSILAS	IM2NP, Marseille
Invité	M. Marc BESCOND	IM2NP, Marseille

Thèse préparée à l'Institut Matériaux Microélectronique Nanosciences de Provence
(IM2NP-CNRS)

Table des matières

Introduction	3
1 Transport électronique dans les nano-composants	7
1.1 Approche classique	8
1.2 Approche quantique	11
1.3 Formule de Landauer et quantum de conductance	15
1.4 Calcul auto-cohérent des équations de Schrödinger et de Poisson . . .	18
2 Formalisme des fonctions de Green	23
2.1 Qu'est ce qu'une fonction de Green ?	24
2.1.1 Point de vue mathématique	24
2.1.2 Sens physique	25
2.1.3 Introduction aux <i>self-energies</i> de contact	26
2.1.4 Conséquences physiques des <i>self-energies</i> de contact	30
2.1.5 Equation de Dyson	32
2.2 Opérateur de Green d'un système fermé	33
2.3 Système ouvert Multibandes en 3 dimensions	36
2.3.1 Hamiltonien multibandes	36
2.3.2 <i>Self-energy</i> de contact d'un Hamiltonien multibandes	39
2.3.3 Décomposition de Schur	43
2.4 Calcul du courant dans le formalisme des fonctions de Green	46
2.4.1 Calcul de la densité de charge et du courant	46
2.4.2 Algorithme de récursion	48
3 Transistor à nanofil <i>notch</i>-MOSFET	53
3.1 Principe de base	54

3.2	Modèle utilisé	56
3.3	Résultats de simulation	58
3.3.1	Influence du <i>notch</i> en régime <i>OFF</i>	59
3.3.2	Influence du <i>notch</i> en régime <i>ON</i>	60
3.3.3	Influence de la hauteur du <i>notch</i>	61
3.3.4	Influence de la longueur du <i>notch</i>	62
3.3.5	Influence du nombre de <i>notches</i>	64
3.3.6	Conclusion	64
4	Transistor à nanofil de type p	67
4.1	Modélisation de la bande de valence en méthode k·p 6 bandes	68
4.1.1	Principe de la théorie k·p	68
4.1.2	Etats de base en bande de valence	69
4.1.3	Modèle à 3 bandes	70
4.1.4	Modèle à 6 bandes : prise en compte du couplage spin-orbite	72
4.1.5	Modèle à 6 bandes pour le nanocomposant	73
4.2	Résultats de simulation	75
4.2.1	Transistor de type p à canal intrinsèque	76
4.2.2	Transistor de type p avec impureté ionisée	77
4.2.3	Conclusion	84
	Conclusion	85
	Bibliographie	88

Introduction

Il était une fois le sable. Constitué principalement d'oxyde de silicium, il est à la base de l'industrie microélectronique. En effet, le silicium, deuxième élément le plus répandu dans la croûte terrestre (environ un quart de son volume), est le principal constituant des transistors. Le silicium est utilisé, évidemment, pour ses propriétés semi-conductrice et a l'avantage d'être peu cher comparé à d'autres matériaux semi-conducteurs comme le germanium. Mais au-delà de son prix, il possède une qualité prépondérante dans l'élaboration de l'empilement métal oxyde semi-conducteur (MOS), qui est une interface entre sa forme pure et son oxyde de grande qualité. L'industrie s'est donc appuyée sur ces deux matériaux pour produire l'élément de base de la microélectronique : le MOSFET.¹

L'augmentation des performances des microprocesseurs passe par l'intégration d'un plus grand nombre de transistors par puce, et donc par la miniaturisation de ceux-ci. Jusque tout récemment, la diminution des tailles était réalisée grâce à des règles de réduction d'échelle, mais l'apparition de phénomènes quantiques a tout remis en cause. Ces phénomènes dégradent les performances du transistor et doivent donc être mieux maîtrisés. Dans cette optique, plusieurs voies sont explorées et même, pour certaines, déjà utilisées. Par exemple, afin de réduire le courant de fuite à travers la grille par effet tunnel, l'oxyde de silicium est remplacé par des matériaux à haute permittivité. Dans le but de réduire le passage tunnel de la source au drain, qui réduit le contrôle du transistor, il est envisagé de changer l'architecture des transistors. Une des voies les plus prometteuses est le transistor à nanofil avec une grille enrobante [1]. C'est cette architecture que nous avons choisi d'explorer théoriquement dans ce travail de thèse.

Ce manuscrit s'articule autour de quatre chapitres, les deux premiers présentant les méthodes de modélisation et simulation utilisées, et les deux suivants regroupant les résultats de simulation.

Le premier chapitre décrit l'approche utilisée pour décrire le transport de charges (électrons ou trous) dans les nano-composants. Nous présenterons les limites des approches classiques comme le modèle de dérive-diffusion [2], l'équation de transport de Boltzmann [3] ou encore le modèle hydrodynamique [4]. Nous verrons ce qu'implique la prise en compte des phénomènes ondulatoires comme l'effet tunnel [6].

¹Metal-Oxide-Semiconductor Field-Effect-Transistor

Nous étudierons ensuite l'approche de Landauer [7–9] qui décrit le système comme étant un conducteur balistique relié à deux réservoirs macroscopiques semi-infinis. Enfin, nous évoquerons l'importance de l'auto-cohérence Schrödinger-Poisson sur le transport [10].

Le deuxième chapitre est consacré au formalisme des fonctions de Green hors équilibre. Nous verrons comment modéliser un composant isolé, puis l'influence des contacts sur celui-ci *via* les *self-energies* [11]. Nous présenterons ensuite le calcul de ces *self-energies* à l'aide de la décomposition de Schur généralisée [12]. Nous verrons aussi que l'approche de Landauer est en accord avec le formalisme des fonctions de Green et qu'elle permet de calculer le courant dans des dispositifs nanométriques [11]. Enfin, nous présenterons un algorithme de récursion basé sur l'équation de Dyson [13] qui permet de réduire considérablement le temps de calcul.

Le troisième chapitre présente une architecture innovante qui a pour but d'améliorer les performances d'un transistor à nano-fils de type n. Avec la réduction des dimensions, les effets quantiques deviennent de plus en plus importants [14], notamment le courant tunnel entre la source et le drain dû à la réduction de la longueur de grille. Nous allons essayer de le réduire en utilisant un autre effet quantique : le confinement. L'impact de cette nouvelle architecture, qui consiste à ajouter un créneau d'oxyde dans le canal, a déjà été étudiée dans des transistors double-grille [15]. Nous détaillerons le principe mettant en jeu des effets quantiques et simulerons, grâce aux outils développés précédemment, un composant dans plusieurs configurations, afin de voir ce que peut apporter cette architecture.

Le quatrième et dernier chapitre est consacré à l'étude d'un transistor à nanofil de type p. Cette étude nécessite une bonne description de la bande de valence. Celle-ci étant fortement non-parabolique, sa description passe par une approche multibandes basée sur un hamiltonien $\mathbf{k} \cdot \mathbf{p}$ [16–18] ou en liaisons fortes [19]. Des simulations en fonctions de Green dans l'espace des modes basé sur un hamiltonien $\mathbf{k} \cdot \mathbf{p}$ ont déjà été effectuées par d'autres groupes [20, 21], mais l'approximation de l'espace des modes néglige certains couplages. D'autres simulations en fonctions de Green basées sur un hamiltonien en liaisons fortes *full-band* ont également déjà été réalisées [22], mais demandent beaucoup de temps et une grande puissance de calcul. De plus, il a été montré que les principaux effets obtenus en *full-band* apparaissent dans une approche $\mathbf{k} \cdot \mathbf{p}$ six bandes [18, 23, 24]. Nous utiliserons donc un modèle en fonctions de Green dans l'espace réel basé sur un hamiltonien $\mathbf{k} \cdot \mathbf{p}$ à six bandes.

La description 3D du modèle que nous avons choisie nous permet de simuler des transistors intrinsèques, mais également de considérer une impureté ionisée dans le canal. Nous verrons que ces impuretés génèrent des couplages supplémentaires qui compliquent encore la physique de ces dispositifs.

Chapitre 1

Transport électronique dans les nano-composants

Les dimensions des composants électroniques n'ont cessé de diminuer de façon spectaculaire depuis plusieurs dizaines d'années. Les dimensions atteignant l'échelle nanométrique, il n'est plus envisageable d'ignorer le comportement quantique des électrons. Nous devons donc modéliser les dispositifs nanométriques avec un formalisme quantique.

Dans ce chapitre, nous évoquerons brièvement l'approche classique utilisée pour la modélisation des dispositifs micrométriques et ses limites pour modéliser des dispositifs nanométriques (chapitre 1.1). Nous étudierons ensuite l'approche ondulatoire et notamment un des phénomènes les plus importants que cette approche permet de décrire : l'effet tunnel (chapitre 1.2). Dans la troisième partie de ce chapitre, nous présenterons la formule de Landauer qui est très utile pour calculer le courant dans des dispositifs nanométriques et qui nous permettra d'obtenir un résultat propre au transport quantique : le quantum de conductance (chapitre 1.3). Enfin, nous verrons l'importance de calculer de façon auto-cohérente les équations de Schrödinger et de Poisson pour la simulation du transport balistique (chapitre 1.4).

1.1 Approche classique

Les électrons, comme toutes les particules, ont la propriété quantique d'être à la fois onde et corpuscule. Il est donc possible d'observer, selon le contexte, des électrons ayant un comportement particulaire (collisions) ou ondulatoire (interférences). Dans un système de grande dimension, les électrons subissent des interactions avec le réseau cristallin, qui détruisent leur phase. Ainsi, les phénomènes d'interférences deviennent marginaux et peuvent être négligés. Dans ce cas l'équation de transport de Boltzmann (ETB) [3] peut être utilisée pour décrire le comportement d'un gaz d'électrons soumis à un champ électrique. Cette équation est obtenue à partir de l'équation de Liouville - Von Neumann, qui régit le transport à partir d'une approche physique, et s'écrit :

$$\frac{\partial f(r, k, t)}{\partial t} + \frac{\hbar k}{m} \frac{\partial f(r, k, t)}{\partial r} + \frac{q}{\hbar k} \frac{\partial U}{\partial r} \frac{\partial f(r, k, t)}{\partial k} = \left[\frac{\partial f(r, k, t)}{\partial t} \right]_{coll} \quad (1.1)$$

Différents modèles classiques reposent sur cette équation.

Nous pouvons également calculer, à partir de l'équation de Liouville - Von Neumann, l'équation de transport de Wigner (ETW) [25] qui est à l'origine de nombreux modèles quantiques. L'ETW s'écrit :

$$\frac{\partial f_w(r, k, t)}{\partial t} + \frac{\hbar k}{m} \frac{\partial f_w(r, k, t)}{\partial r} + \theta(f_w(r, k, t)) = \left[\frac{\partial f_w(r, k, t)}{\partial t} \right]_{coll} \quad (1.2)$$

où

$$\theta(f_w(r, k, t)) = -i \int V_w(r, k - k') f_w(r, k', t) dk' \quad (1.3)$$

est le terme quantique appelé "potentiel quantique de Wigner".

Dans une approche macroscopique, l'ETB est intégrée à différents ordres (ici 0, 1 et 2) afin de relier les paramètres macroscopiques entre eux. Ces paramètres sont :

- Concentration des porteurs : $n = \frac{1}{8\pi^3} \int \varphi d^3k$
- Vitesse moyenne : $v_{moy} = \frac{1}{8\pi^3 n} \int v_g \cdot \varphi d^3k$
- Energie moyenne : $w = \frac{1}{8\pi^3 n} \int \varepsilon \cdot \varphi d^3k$
- Flux d'énergie : $S = \frac{1}{8\pi^3 n} \int v_g \cdot \varepsilon \cdot \varphi d^3k$

avec v_g la vitesse de groupe et ε l'énergie. L'ETB à intégrer dans notre cas est :

$$\frac{\partial}{\partial t} \langle \theta \rangle + \nabla_r \langle v_g \theta \rangle + \frac{q}{\hbar k} \frac{\partial U}{\partial r} \nabla_k \langle \theta \rangle = - \left\langle \frac{\partial}{\partial t} \theta \right\rangle_{coll} \quad (1.4)$$

avec $\langle \theta \rangle = \frac{2}{8\pi^3} \int \varphi \theta d^3k$

Nous obtenons le système d'équation suivant :

$$\frac{\partial}{\partial t} n + \nabla_r (n.v_{moy}) = -R_{net} \quad (\text{ordre 0}) \quad (1.5)$$

(R_{net} est le coefficient de génération recombinaison étant l'intégrale du terme de collision)

$$\frac{\partial}{\partial t} (n.v_{moy}) + \frac{2}{3m} \nabla_r (n.w) + \frac{F}{m} n = -\frac{n(v_{moy} - v_{moy0})}{\tau_m} \quad (\text{ordre 1}) \quad (1.6)$$

(τ_m étant le temps de relaxation en moment)

$$\frac{\partial}{\partial t} (n.w) + \nabla_r (n.S) - F.n.v_{moy} = -n \frac{(w - w_0)}{\tau_\varepsilon} \quad (\text{ordre 2}) \quad (1.7)$$

(τ_ε étant le temps de relaxation en énergie)

Le modèle de dérive-diffusion [2] résulte de l'équation de transport de Boltzmann dans l'hypothèse où les porteurs se déplacent dans des conditions de quasi-équilibre [26]. Dans ce modèle, l'équation de transport de Boltzmann est intégrée à l'ordre 0 et à l'ordre 1. Nous devons ensuite imposer une condition sur l'énergie afin de fermer le système à l'ordre 2. Ici nous considérerons que la température des électrons est égale à celle du réseau. L'énergie est donc constante et vaut :

$$w = \frac{3}{2} k_B T \quad (1.8)$$

En multipliant par la charge élémentaire q , nous obtenons les deux équations régissant le transport du modèle Dérive-Diffusion :

$$q \cdot \frac{\partial}{\partial t} n + \nabla_r (J_n) = -q \cdot R_{net} \quad (1.9)$$

$$J_n = q \cdot D_n \cdot \nabla_r n + q \cdot n \cdot \mu_n \cdot E \quad (1.10)$$

où $\mu_n = \frac{q \cdot \tau_m}{m}$ la mobilité de l'électron, E le champ électrique et $D_n = \mu_n \frac{k_B \cdot T}{q}$ le coefficient de diffusion des électrons. Cette équation couplée à l'équation de Poisson constitue l'approche la plus développée pour décrire un transport classique.

Le modèle HydroDynamique (HD) [4], qui résulte lui aussi de l'équation de Boltzmann, est plus complexe et permet de prendre en compte la variation de la diffusion en fonction de la température. Dans ce modèle, on intègre jusqu'à l'ordre 2 et on impose une condition sur le flux énergétique (ordre 3) qui s'exprime de la manière suivante [27] :

$$n.S = -\frac{5.k_B\mu.T_n}{2.q}J_n + n.Q \quad (1.11)$$

où Q est le flux de chaleur exprimé par la loi de Fourier et la loi de Wiedermann-Frantz [28]. Les trois équations régissant le transport du modèle HD sont :

$$q.\frac{\partial}{\partial t}n + \nabla_r(J_n) = -q.R_{net} \quad (1.12)$$

$$J_n = \mu_n.k_B.[\nabla_r(n.T_n) + \frac{q}{k_B}.E.n] - \tau_m.\frac{\partial J_n}{\partial t} \quad (1.13)$$

$$\frac{3.k_B}{2}\frac{\partial(n.T_n)}{\partial t} + \nabla_r(n.S) - E.J_n = -\frac{3.k_B}{2}\frac{T_n - T_L}{\tau_\varepsilon} \quad (1.14)$$

avec T_n la température des électrons et T_L la température du réseau.

Le modèle *Energie Balance* est une composante du modèle HD considérant une relation entre le temps de relaxation et l'énergie [29]. Les trois équations deviennent :

$$q.\frac{\partial}{\partial t}n + \nabla_r(J_n) = -q.R_{net} \quad (1.15)$$

$$J_n = \mu_n.[k_B.(T_n.\nabla_r(n) + n.\nabla_r(T_n)(1 + \nu_n) + q.E.n)] \quad (1.16)$$

$$\frac{3.k_B}{2}\frac{\partial(n.T_n)}{\partial t} + \nabla_r(n.S) - E.J_n = -\frac{3.k_B}{2}\frac{T_n - T_L}{\tau_\varepsilon} \quad (1.17)$$

où $\nu_n = \frac{T_n.\nabla_r(\mu_n)}{\mu_n.\nabla_r(T_n)}$ est un terme correctif qui traduit la dépendance en température du paramètre mobilité.

Tant que les dimensions sont suffisamment grandes les interactions inélastiques limitent les interférences. En revanche les phénomènes d'interférences deviennent non-négligeables à partir d'une certaine taille limite. Cette limite est difficile à déterminer. En effet elle dépend autant des électrons eux mêmes (énergie, densité...) que du milieu dans lequel ils évoluent : c'est l'interaction entre les deux qui est déterminante. Des ordres de grandeurs peuvent être trouvés dans la littérature ; par exemple le GaAs à 77 K a un libre parcours moyen de 0,1 μm [30]. Le caractère

ondulatoire peut même devenir prépondérant lorsque le potentiel varie de façon importante sur une dimension de l'ordre de grandeur de la longueur d'onde des électrons. Afin de calculer l'ordre de grandeur de la longueur d'onde d'un électron, considérons un électron thermique, c'est à dire un électron dont l'énergie cinétique est donnée par :

$$E_c = E_{th} = \frac{3}{2}k_B T \quad (1.18)$$

où T est la température et k_B est la constante de Boltzmann. Par ailleurs, l'énergie cinétique d'un électron dans un potentiel périodique (réseau cristallin) peut s'écrire :

$$E_c = \frac{\hbar^2 k^2}{2m^*} \quad (1.19)$$

où k est le module du vecteur d'onde de l'électron et m^* sa masse effective. Cette expression représente l'approximation dite de la masse effective et est vérifiée lorsque E_c est très faible. Dans cette approximation, m^* représente l'influence du cristal sur le comportement dynamique d'un électron de faible énergie. Dans le cas d'un électron thermique, à température ambiante, l'approximation de la masse effective est bien vérifiée. Nous pouvons donc calculer le module du vecteur d'onde de l'électron :

$$k = \frac{\sqrt{3m^*k_B T}}{\hbar} \Rightarrow \lambda = \frac{2\pi}{k} = \frac{h}{\sqrt{3m^*k_B T}} \quad (1.20)$$

où λ est la longueur d'onde de l'électron. A température ambiante et avec $m^* = m_0$ (masse de l'électron libre) nous obtenons $\lambda = 6,2$ nm ; pour $m^* = 0,2 m_0$ on a $\lambda = 13,8$ nm. Il ne fait donc aucun doute que dans les composants en Si (où la masse transverse $m_t^* = 0,2$ et la masse longitudinale $m_l^* = 0,9 m_0$) dont la longueur de grille est inférieure à 20 nm, le comportement ondulatoire des électrons ne peut être ignoré.

1.2 Approche quantique

Comme nous venons de le voir, en deça d'une certaine taille, il n'est plus possible d'ignorer le comportement ondulatoire des particules, le transport devient quantique. Les effets quantiques peuvent être séparés en deux catégories : les effets quantiques électrostatiques (répartition des porteurs et quantification de l'énergie) et le trans-

port quantique (interférences, effet tunnel). Certains modèles se contentent de coupler les effets quantiques électrostatiques avec des équations de transport classique. Le transport n'est pas quantique !

Le modèle *Density Gradient* consiste à ajouter au potentiel classique un potentiel de correction quantique tel que :

$$\phi_n = \phi_{class} + \phi_{quant} \text{ avec } \phi_{quant} = \alpha_n \cdot \left(\frac{\nabla^2 \sqrt{n}}{\sqrt{n}} \right) \quad (1.21)$$

α_n étant une constante dont l'expression varie selon les approches employées. En appliquant la méthode des moments à l'équation de transport de Wigner [31] nous obtenons :

$$\alpha_n = -\frac{\gamma \cdot \hbar^2}{6m_{DOS}} \quad (1.22)$$

En utilisant la formulation de Bohm [32] nous avons :

$$\alpha_n = -\frac{\hbar^2}{2m_{eff}} \quad (1.23)$$

Ancona a quant à lui développé une approche microscopique pour déterminer α [33] :

$$\alpha_n = -\frac{\hbar^2}{12m_{DOS}}. \quad (1.24)$$

Outre le modèle *Density Gradient*, les effets quantiques électrostatiques peuvent être ajoutés en résolvant numériquement le système Poisson-Schrödinger couplé avec l'équation de transport de Dérive-Diffusion [34] [35].

Pour prendre en compte, non seulement les effets quantiques électrostatiques, mais aussi les effets quantiques de transport, il faut coupler l'équation de transport de Wigner avec celle de Schrödinger [36]. Dans ce cas, la simulation du transport est effectuée grâce à la méthode Monte Carlo. Dans un simulateur de type Monte Carlo, on décompose la trajectoire des particules en un grand nombre de séquences aléatoires. Durant chaque séquence, un porteur suit une trajectoire déterministe, puis subit une collision, choisie de façon aléatoire, qui modifie sa trajectoire. La répétition de ces séquences permet d'avoir une solution de plus en plus proche de la réalité et ainsi obtenir le courant, la densité de charges... Cette méthode a l'avantage

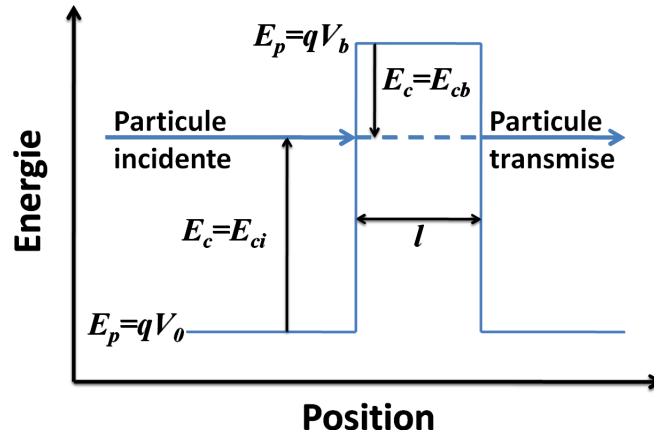


FIG. 1.1: Représentation schématique d'une particule arrivant sur une barrière de potentiel. La largeur l de la barrière est suffisamment faible pour que le passage tunnel soit possible.

de pouvoir appliquer n'importe quelle équation de transport entre deux collisions. Nous pouvons trouver dans la littérature du transport classique [3], du transport classique corrigé [37], les ETB couplées à une résolution de Schrödinger [38], et enfin, la méthode Wigner-Monte Carlo évoquée précédemment.

Nous verrons dans la suite (Chapitre 2) une autre méthode permettant de simuler le transport quantique, le formalisme des fonctions de Green hors-équilibre. C'est cette méthode que j'ai utilisée durant cette thèse car elle présente plusieurs avantages par rapport à la méthode Wigner-Monte Carlo dans le cadre des études qui ont été effectuées, avec notamment une grande importance du courant tunnel. Le courant tunnel est un phénomène qui ne s'explique pas sans une approche ondulatoire et peut devenir prépondérant dans le comportement électrique d'un composant.

L'effet tunnel est un des phénomènes les plus spectaculaires et importants de la mécanique quantique. Avec une approche classique, une particule arrivant sur une barrière de potentiel ne peut franchir cette barrière que si son énergie est supérieure à l'énergie maximale de la barrière. Ce n'est plus le cas dans une approche quantique. Le résultat ne peut pas être déterminé à l'avance : la probabilité de passage n'est jamais tout à fait nulle ou égale à 1. Ceci s'explique à l'aide de la vision ondulatoire de la mécanique quantique.

Supposons le système représenté sur la figure 1.1. L'énergie totale de la particule est donnée par $E = E_c + E_p$ avec E_c l'énergie cinétique, E_p l'énergie potentielle. On

se place dans le cas d'un transport élastique, c'est à dire que l'énergie totale reste invariante. Avant d'atteindre la barrière nous avons :

$$E = E_{ci} + qV_0 = \frac{\hbar^2 k^2}{2m^*} + qV_0 \geq 0 \quad (1.25)$$

avec k le vecteur d'onde, E_{ci} l'énergie cinétique de la particule incidente et V_0 le potentiel à gauche de la barrière. La fonction d'onde de la particule est donnée par $\psi_i = e^{ikx}$, c'est à dire que l'on suppose une particule libre. Lors du passage de la barrière nous avons :

$$E = E_{cb} + qV_b = \frac{\hbar^2 \rho^2}{2m^*} + qV_b \quad (1.26)$$

avec ρ le vecteur d'onde, E_{cb} l'énergie cinétique de la particule au niveau de la barrière et V_b le potentiel de la barrière. Afin de mettre en évidence l'effet tunnel, nous nous plaçons dans le cas où $E_{ci} \leq q(V_b - V_0)$, c'est à dire $\frac{\hbar^2 \rho^2}{2m^*} \leq 0$ et ρ est donc imaginaire. La fonction d'onde associée est une onde évanescente $\psi_b = e^{-|\rho|x}$. Si la largeur de la barrière l n'est pas trop supérieure à la portée de l'onde évanescente $1/|\rho|$, alors une partie de l'onde incidente est transmise à travers la barrière : c'est l'effet tunnel. Nous calculons maintenant des ordres de grandeurs associés à cet effet. Si $l = 1/|\rho|$ il y a approximativement 8 particules sur 10 qui franchissent la barrière par effet tunnel. Calculons $1/|\rho|$ en vérifiant la conservation de l'énergie (égalité des équations 1.25 et 1.26) :

$$\frac{1}{|\rho|} = \sqrt{\frac{\hbar^2}{2m^* (q(V_b - V_0) - E_{ci})}}. \quad (1.27)$$

En considérant la masse effective des électrons dans le Si ($m^* = 0,2 m_0$) et $q(V_b - V_0) - E_{ci} = 0,1$ eV, nous obtenons $1/|\rho| = 1,4$ nm. Une barrière de 1,4 nm et d'une hauteur de 0,1 eV est quasiment transparente dans le Si. Nous constatons que les énergies et les dimensions auxquelles se manifeste cet effet correspondent aux tailles rencontrées en nanoélectronique.

De la même façon qu'une barrière de potentiel n'est jamais tout à fait opaque, une variation de potentiel implique une certaine réflexion des électrons, même si l'énergie cinétique de ces derniers est importante. Nous considérons $E_{cb} > 0$, cela ne signifie pas que la transmission sera parfaite (égale à 1) [39] (contrairement au

cas classique). Autrement dit, la transmission à travers un composant n'est parfaite que si le potentiel ne varie pas fortement le long de la direction de transport. Or pour faire circuler des électrons il faut appliquer une différence de potentiel aux extrémités du composant qui dégrade la transmission. Le caractère quantique des électrons va donc fortement influencer leur transmission et par la même modifier les caractéristiques électriques des transistors.

1.3 Formule de Landauer et quantum de conduction

On considère le système constitué des 3 éléments suivants : un conducteur 1D encadré par deux grands réservoirs d'électrons. Le transport électronique dans le conducteur est balistique [40], c'est-à-dire que les électrons se déplacent sans subir de collisions inélastiques. Les deux réservoirs d'électrons G (gauche) et D (droite) sont macroscopiques (semi-infinis) et à l'équilibre thermodynamique, c'est-à-dire qu'ils fournissent et absorbent des électrons sans modifier la distribution de leurs électrons. On note μ_G et μ_D leur niveau de Fermi respectifs. Le confinement quantique dans les directions transverses au transport transforme le continuum d'états électroniques en sous-bandes appelées modes de conduction. La relation de dispersion des électrons dans les réservoirs est $E(k)$ et leur vitesse est $v_{\pm k} = \frac{1}{\hbar} \left[\frac{dE}{dk} \right]_{\pm k}$. Les électrons se trouvant dans les états $+k$ vont de G vers D et ceux dans les états $-k$ vont de D vers G . Si l'on connecte le conducteur aux deux réservoirs, à chaque sous-bande n correspond une transmission $T_n(E)$. Cette transmission, comprise entre 0 et 1, est la probabilité qu'un électron d'énergie E de passer d'un réservoir à un autre par la sous-bande n . Un électron arrivant dans le conducteur aura donc la probabilité T_n d'être transmis et la probabilité $1 - T_n$ d'être réfléchi. Le courant transmis par le mode de conduction n est donc :

$$I_n = \frac{2e}{L} \sum_{+k} T_n(E(+k)) v_{+k} f_G + \frac{2e}{L} \sum_{-k} T_n(E(-k)) v_{-k} f_D \quad (1.28)$$

L étant la longueur des réservoirs et $f_{G,D}$ les distributions d'électrons dans les réservoirs G et D tels que :

$$f_G = \frac{1}{1 + \exp\left(\frac{E(k) - \mu_G}{k_B T}\right)} \quad (1.29)$$

et

$$f_D = \frac{1}{1 + \exp\left(\frac{E(k) - \mu_D}{k_B T}\right)}. \quad (1.30)$$

Sachant que $v_{+k} = -v_{-k}$ et que la transmission ne dépend pas du sens de déplacement des électrons, on peut réécrire l'équation 1.28 :

$$\begin{aligned} I_n &= \frac{2e}{L} \sum_{+k} T_n(E(k)) v_{+k} (f_G - f_D) \\ &= \frac{2e}{L\hbar} \sum_{+k} T_n(E(k)) \frac{dE}{dk} (f_G - f_D). \end{aligned} \quad (1.31)$$

Dans les réservoirs de longueur L , il y a un état tous les $2\pi/L$ dans l'espace des modes [41]. Les réservoirs étant semi-infinis, L tend vers l'infini. On peut donc transformer la somme discrète par une intégrale dans l'espace des modes tel que :

$$\sum_k \rightarrow \frac{L}{2\pi} \int dk \quad (1.32)$$

On obtient :

$$\begin{aligned} I_n &= \frac{2e}{L\hbar} \frac{L}{2\pi} \int T_n(E(k)) \frac{dE}{dk} (f_G - f_D) dk \\ &= \frac{2e}{h} \int T_n(E) (f_G - f_D) dE. \end{aligned} \quad (1.33)$$

Cette équation est la formule de Landauer qui permet de calculer le courant transmis par un mode de conduction 1D connecté à deux réservoirs semi-infinis [8]. Le courant total transmis par le conducteur s'obtient en sommant les courants de chaque mode [42] :

$$\begin{aligned}
 I &= \sum_n I_n \\
 &= \frac{2e}{h} \int T(E)(f_G - f_D)dE.
 \end{aligned}
 \tag{1.34}$$

On en déduit que la partie essentielle dans le calcul du courant dans le conducteur 1D est le calcul de la transmission entre les deux réservoirs.

Nous allons maintenant nous intéresser à la conductance d'un tel système. On considère un conducteur avec un seul mode de conduction et une transmission parfaite ($T = 1$). Ce système n'a, *a priori*, pas de résistance et donc une conductance infini. On peut calculer le courant dans le conducteur avec la formule de Landauer (1.34) :

$$I = \frac{2e}{h} \int_{-\infty}^{+\infty} (f_G - f_D)dE.
 \tag{1.35}$$

Pour simplifier le calcul, on se place à température nulle. Dans ce cas, tous les états en dessous du niveau de Fermi sont occupés et tous les états en dessus sont vides. On a donc :

$$\begin{aligned}
 I &= \frac{2e}{h} \left(\int_{-\infty}^{\mu_D} 0dE + \int_{\mu_D}^{\mu_G} 1dE + \int_{\mu_G}^{+\infty} 0dE \right) \\
 &= \frac{2e}{h}(\mu_G - \mu_D).
 \end{aligned}
 \tag{1.36}$$

En définissant V la différence de potentiel appliquée aux extrémités du composant, on a $-eV = \mu_G - \mu_D$ donc :

$$I = \frac{2e^2}{h} V_{GD}.
 \tag{1.37}$$

On peut voir que même avec une transmission parfaite, le courant est fini et la résistance est non-nulle. Cette résistance est l'inverse du quantum de conductance :

$$G = \frac{I}{V} = \frac{2e^2}{h} = (12,9 \text{ k}\Omega)^{-1}.
 \tag{1.38}$$

A noter que ce résultat est également valable pour des températures non-nulles.

Pourquoi ce conducteur parfait a-t-il une résistance ? Parce que la quantité de courant pouvant circuler dans le canal de conduction est définie par la fonction de distribution des réservoirs. Un réservoir quasi-infini peut théoriquement fournir une quasi-infinité d'électrons (variation en L), mais le courant généré par chaque électron tend vers 0 car ils sont délocalisés dans le réservoir quasi-infini ce qui offre une densité de charge linéaire quasi-nulle (variation en $1/L$). Ces deux phénomènes se neutralisent à une constante près : le quantum de conductance.

Dans les systèmes macroscopiques, le courant varie linéairement avec la tension en suivant la loi d'Ohm. Dans les nanocomposants, le matériau adopte un comportement qui ne peut être interprété sans quantifier la conductance [43,44] : le courant évolue par paliers successifs.

1.4 Calcul auto-cohérent des équations de Schrödinger et de Poisson

Nous avons vu précédemment que le calcul du courant dans un composant à l'aide de la formule de Landauer nécessitait la connaissance des fonctions de distribution d'électrons dans les réservoirs de gauche et de droite, ainsi que la transmission. Toute la difficulté est de connaître la transmission qui dépend du profil de potentiel dans la région active. Or ce dernier dépend de la différence de potentiel (ddp) et de la densité de charges dans le composant par l'intermédiaire de l'équation de Poisson. Par ailleurs, afin de calculer la densité de charges, il est nécessaire de résoudre l'équation de Schrödinger. Les deux équations de Schrödinger et Poisson étant dépendantes l'une de l'autre (figure 1.2), nous devons les résoudre de façon auto-cohérente [10].

Pour simuler un composant électronique composé d'une partie active (canal de conduction du transistor) et de deux réservoirs, il faut imposer des conditions aux limites. En effet, les réservoirs étant macroscopiques, il n'est pas envisageable de simuler la totalité du composant. Il faut donc modéliser les réservoirs en choisissant correctement les conditions aux limites. Dans notre approche nous considérons les réservoirs comme étant semi-infinis et à l'équilibre thermodynamique. Ils peuvent donc injecter et absorber des électrons sans faire varier leur niveau de Fermi. Ceci implique l'électronneutralité et donc un champ électrique nul, dans le sens du trans-

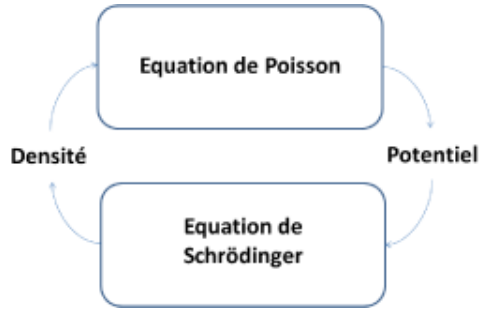
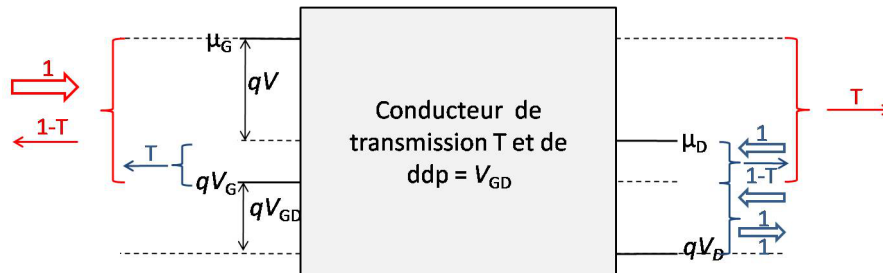


FIG. 1.2: Calcul auto-cohérent de la densité de charges et du potentiel.


 FIG. 1.3: Représentation schématique du composant comprenant un conducteur (de transmission T et de ddp V_{GD}) et des réservoirs (de ddp V).

port, aux deux extrémités de la partie simulée du composant. Ces conditions aux limites adoptées, nous pouvons maintenant résoudre le système auto-cohérent. Nous noterons V_G et V_D les valeurs respectives du potentiel à gauche et à droite de la partie active. Afin de vérifier l'électroneutralité des réservoirs, la densité électronique doit être identique aux deux extrémités, $n_D = n_G$.

Observons ce qu'implique qualitativement la vérification des conditions aux limites. Afin de calculer facilement les densités de charge, nous allons considérer la température nulle et une tension appliquée V très faible ; cette dernière supposition permet de considérer que la densité d'états est constante sur la gamme d'énergie qui nous intéresse (on la prend égale à 1). On appelle T la transmission du conducteur. A l'extrémité gauche nous avons donc (le système est décrit figure 1.3) :

$$n_G = (\mu_G - qV_G)(1 + (1 - T)) + (\mu_D - qV_G)T. \quad (1.39)$$

Dans cette expression, le premier terme de la somme correspond aux électrons

venant de G (1), dont certains sont réfléchis en G ($1 - T$). Le second terme correspond aux électrons transmis depuis D . Dans les deux cas nous intégrons entre le bas de la bande de conduction qV_G et le niveau de Fermi correspondant au réservoir d'origine. Après simplification on a :

$$n_G = -TqV + 2(\mu_G - qV_G). \quad (1.40)$$

Pour l'extrémité D nous obtenons avec un raisonnement similaire :

$$n_D = (\mu_D - qV_G)(1 + (1 - T)) + (\mu_G - qV_G)T + 2(qV_G - qV_D). \quad (1.41)$$

Dans ce cas, tous les électrons venant de D et ayant une énergie comprise entre qV_D et qV_G sont réfléchis (correspond à la bande interdite en G) : ils sont donc comptés 2 fois. Après simplification nous avons :

$$n_D = TqV + 2(\mu_D - qV_D). \quad (1.42)$$

Comme nous l'avons dit précédemment, les conditions aux limites pour l'équation de Poisson imposent que nous ayons $n_D = n_G$. Un rapide calcul à partir des expressions précédentes permet de trouver :

$$V_{GD} = V_G - V_D = V(1 - T). \quad (1.43)$$

Cette expression, obtenue dans des conditions particulières (température nulle et tension très faible), met néanmoins en évidence un point très important du transport quantique balistique : si $T = 1$ alors $V_{GD} = 0$ V. On ne peut donc pas appliquer de différence de tension aux bornes d'un conducteur dont la transmission est parfaite. Par ailleurs la réciproque est vraie puisque, comme nous l'avons vu, $T = 1$ à condition qu'il n'y ait aucune variation du potentiel dans le composant : on voit dans cet exemple l'importance de l'auto-cohérence dans la simulation. Nous vérifions également que les conditions aux limites adoptées (électronneutralité avec champ nul) sont en parfait accord avec l'approche de Landauer.

Cet exemple nous permet également de mettre en évidence la notion de résistance des réservoirs. Si $T = 1$ la variation du potentiel, due à la différence de tension appliquée aux extrémités du composant (réservoirs et conducteur), n'aura lieu que

dans les réservoirs. Afin de calculer cette résistance on peut calculer le courant : nous sommes dans le cas où $T = 1$ et le courant est donné par l'expression 1.37 :

$$J = \frac{2q^2}{h}V. \quad (1.44)$$

On a donc pour la résistance des réservoirs $R_{res} = \frac{h}{2q^2}$ qui est l'inverse du quantum de conductance. Le quantum de conductance est encaissé dans les réservoirs mais est propre à tous le système réservoirs + conducteur. En effet, les réservoirs isolés, en tant qu'objet macroscopiques, ont une résistance propre qui n'a rien à voir avec la résistance que nous venons de calculer. On peut remarquer au final que cette résistance finie est encaissée sur une longueur infinie (celle des réservoirs) ; cela permet bien de considérer les réservoirs à l'équilibre.

Si on suppose que $T \approx 0$, alors on a $V_{GD} \approx V$. La différence de tension est donc presque entièrement encaissée dans le conducteur quantique et très peu dans les réservoirs. Dans ce cas le conducteur offre une résistance très importante, très supérieure à R_{res} .

Cet exemple simple permet de comprendre l'importance de l'auto-cohérence dans la simulation de composant dont la partie active est un conducteur 1D balistique. Ceci permet également de souligner le comportement d'un composant (même parfait) ne peut pas être compris uniquement au regard de sa partie active ; il faut raisonner en intégrant les réservoirs macroscopiques. Cela sera d'autant plus clair dans le chapitre suivant (chapitre 2) où l'on étudiera la perturbation apportée par les réservoirs sur les modes de transmission du conducteur.

Chapitre 2

Formalisme des fonctions de Green

Ce chapitre sera consacré à l'étude du formalisme des fonctions de Green et à son utilisation pour la modélisation des nano-dispositifs. Nous verrons, tout d'abord, ce qu'est une fonction de Green d'un point de vue mathématique avant de la définir physiquement dans un système quantique (chapitre 2.1). Ce chapitre introduira également la notion de self-energy ainsi que l'équation de Dyson. Nous étudierons ensuite un simple système 1D sans interactions afin de se familiariser avec l'opérateur de Green et les grandeurs physiques qui en découlent (chapitre 2.2). Dans un troisième temps, nous considérerons un système ouvert avec un Hamiltonien multibandes et décrirons l'algorithme de Schur qui permet de calculer les self-energies dans le cas du Hamiltonien multibandes (chapitre 2.3). Enfin, nous présenterons l'algorithme de récursion, basé sur l'équation de Dyson, permettant de réduire de façon significative la taille des matrices à inverser (chapitre 2.4).

2.1 Qu'est ce qu'une fonction de Green ?

2.1.1 Point de vue mathématique

Avant de passer au sens physique, il est important de savoir ce qu'est une fonction de Green d'un point de vue mathématique. Pour cela nous allons tout d'abord poser le problème et introduire la notion de distribution de Dirac. Soit x une position, D un opérateur différentiel linéaire et $j(x)$ une fonction donnée. On cherche $\varphi(x)$ la solution de l'équation aux dérivées partielles linéaire suivante :

$$D\varphi(x) = j(x). \quad (2.1)$$

On appelle distribution de Dirac (que l'on note $\delta(x)$) une fonction nulle pour tout $x \neq 0$ et qui vérifie l'équation suivante :

$$\int \varphi(x)\delta(x_0 - x)dx = \varphi(x_0). \quad (2.2)$$

On appelle fonction de Green $G(x)$ toute solution de l'équation aux dérivées partielles linéaire :

$$DG(x) = \delta(x). \quad (2.3)$$

Pour la même équation aux dérivées partielles linéaire, il existe en général plusieurs solutions $G(x)$. Il est donc indispensable de définir des conditions aux limites pour que la fonction de Green soit unique. Nous venons de voir que la définition d'une fonction de Green est composée de deux éléments :

- l'équation aux dérivées partielles linéaire.
- les conditions aux limites.

Nous connaissons maintenant la fonction de Green, mais quel en est l'intérêt ? Le principal intérêt réside dans le fait que de la solution $\varphi(x)$ de l'équation 2.1 peut s'écrire sous la forme d'un produit de convolution :

$$\varphi(x) = (G * j)(x) = \int G(x - y)j(y)dy. \quad (2.4)$$

En effet, à partir de l'équation 2.1, nous avons :

$$D\varphi(x) = j(x) = \int \delta(x - y)j(y)dy \quad (2.5)$$

et d'après l'équation 2.3 on obtient :

$$D\varphi(x) = \int [DG(x - y)]j(y)dy. \quad (2.6)$$

Or D n'agissant que sur x , nous pouvons écrire :

$$D\varphi(x) = D \int G(x - y)j(y)dy. \quad (2.7)$$

Par simplification de l'équation nous obtenons :

$$\varphi(x) = \int G(x - y)j(y)dy = (G * j)(x) \quad (2.8)$$

qui est bien solution de l'équation 2.1.

2.1.2 Sens physique

Dans le cas d'un système quantique, la fonction de Green est la réponse du système à une impulsion. La fonction de Green est dite retardée si on ne s'intéresse qu'à l'évolution vers le futur. La fonction de Green retardée représente ainsi la probabilité de trouver la particule en position r_2 à l'instant t_2 sachant qu'elle se trouvait en position r_1 à l'instant t_1 . Si on considère un système invariant dans le temps ne dépendant que de l'énergie, l'équation 2.3 devient par transformation de Fourier :

$$\sigma(E)G(E) = 1 \quad (2.9)$$

avec $\sigma(E)$ le symbole de l'opérateur différentiel D , qui est une fonction polynomiale en E . L'équation 2.9 peut se réécrire comme ceci :

$$G(E) = \frac{1}{\sigma(E)}. \quad (2.10)$$

La fonction de Green indépendante du temps $G(E)$ est donc l'inverse de $\sigma(E)$. De façon similaire, nous notons que l'équation de Schrödinger indépendante du temps s'écrit :

$$H|\psi\rangle = E|\psi\rangle \quad (2.11)$$

$|\psi\rangle$ et E étant respectivement les vecteurs propres et les valeurs propres de l'opérateur hamiltonien H . On peut réécrire l'équation 2.11 sous la forme :

$$(E - H)|\psi\rangle = 0. \quad (2.12)$$

Le système étant indépendant du temps, on peut montrer que $\sigma(E) = (E - H)$. D'après l'équation 2.10, la fonction de Green s'écrit donc dans ce cas :

$$G(E) = (E - H)^{-1}. \quad (2.13)$$

Pour que la fonction de Green soit définie pour toutes les valeurs de l'énergie, y compris les valeurs propres du Hamiltonien, il faut ajouter à $E - H$ un terme imaginaire η qui tend vers 0. Nous venons de définir l'opérateur de la fonction de Green retardée [45] du système étudié qui s'écrit :

$$G(E) = \lim_{\eta \rightarrow 0^+} (E - H + i\eta)^{-1}. \quad (2.14)$$

Cet opérateur permet, comme nous le verrons plus tard, de calculer de nombreuses grandeurs (densités de charges, courant, ...).

2.1.3 Introduction aux *self-energies* de contact

Avant d'introduire la notion de *self-energy*, nous allons voir comment traiter numériquement les différentes parties d'un composant. Prenons le cas simple d'un dispositif unidimensionnel constitué de la partie active connectée à deux réservoirs d'électrons. Le Hamiltonien H_0 du composant est :

$$H_0 = E_c - \frac{\hbar^2}{2m^*} \frac{d^2}{dx^2} + U(x) \quad (2.15)$$

où E_c est le bas de la bande de conduction (ou de la sous-bande), m^* est la masse effective des électrons dans le matériau considéré dans la direction x et $U(x)$ l'énergie potentielle.

Afin de traiter numériquement ce problème, il faut définir un maillage le long du composant. Nous décrivons donc le conducteur continue par N points espacés

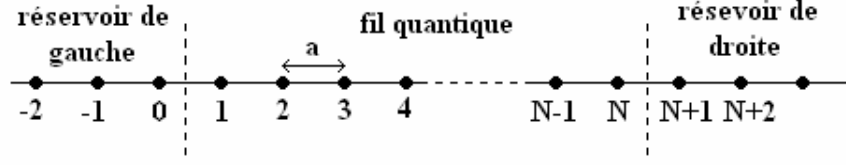


FIG. 2.1: Représentation des sites discrets de la partie active du composant et des réservoirs de gauche et de droite.

d'un pas constant a . Nous adoptons une représentation en kets $|r\rangle$, r indiquant la position dans la partie active. Le site $|1\rangle$ est ensuite relié au réservoir de gauche et le site $|N\rangle$ à celui de droite comme schématisé sur la figure 2.1.

Il faut donc discrétiser le Hamiltonien de l'expression 2.15. Pour cela, appliquons H_0 à une fonction ψ :

$$H_0\psi = (E_c + U(x))\psi - \frac{\hbar^2}{2m^*} \frac{d^2\psi}{dx^2}. \quad (2.16)$$

$E_c + U(x)$ peut facilement être discrétisé et on définit U_r comme le potentiel au point $|r\rangle$. Nous discrétisons ensuite $\frac{d^2}{dx^2}$ en utilisant la méthode des différences finies :

$$\frac{d}{dx} \left(\frac{d\psi}{dx} \right) = \frac{d}{dx} \left(\frac{\psi_{r+1} - \psi_r}{a} \right) \quad (2.17)$$

$$\frac{d^2\psi}{dx^2} = \frac{\psi_{r+1} - 2\psi_r + \psi_{r-1}}{a^2} \quad (2.18)$$

où ψ_r représente ψ sur le site $|r\rangle$. On obtient donc l'expression 2.16 discrétisé :

$$\begin{aligned} H_0\psi_r &= (E_c + U_r)\psi_r - \frac{\hbar^2}{2m^*a^2}(\psi_{r+1} - 2\psi_r + \psi_{r-1}) \\ &= (E_c + U_r + 2t)\psi_r - t\psi_{r+1} - t\psi_{r-1} \end{aligned} \quad (2.19)$$

avec $t = \frac{\hbar^2}{2m^*a^2}$ qui couple deux sites successifs. L'équation 2.19 peut s'écrire sous forme matricielle :

$$H_0 = \begin{bmatrix} E_c + 2t + U_1 & -t & 0 & \cdots & 0 \\ -t & E_c + 2t + U_2 & \ddots & \ddots & \vdots \\ 0 & -t & \ddots & -t & 0 \\ \vdots & \ddots & \ddots & E_c + 2t + U_{N-1} & -t \\ 0 & \cdots & 0 & -t & E_c + 2t + U_N \end{bmatrix}. \quad (2.20)$$

La fonction de Green du composant isolé est alors donnée par (Equ. 2.14) :

$$G_0(E) = \lim_{\eta \rightarrow 0^+} (E - H_0 + i\eta)^{-1}. \quad (2.21)$$

Pour définir le composant complet, c'est-à-dire la partie active et les deux réservoirs (semi-infinis), il faudrait un Hamiltonien de dimension infinie comme celui-ci :

$$H = \left[\begin{array}{ccc|ccc|ccc} \ddots & \ddots & 0 & & & & & & & & & \\ \ddots & 2t + U_1 & -t & 0 & & & & & & & & \\ 0 & -t & 2t + U_1 & -t & 0 & & & & & & & \\ \hline & 0 & -t & A + U_1 & -t & 0 & & & & & & \\ & & 0 & -t & A + U_2 & \ddots & 0 & & & & & \\ & & & 0 & \ddots & \ddots & -t & & 0 & & & \\ & & & & 0 & -t & A + U_N & & -t & 0 & & \\ \hline & & & & & 0 & -t & 2t + U_N & -t & 0 & & \\ & & & & & & 0 & -t & 2t + U_N & \ddots & & \\ & & & & & & & 0 & \ddots & \ddots & & \end{array} \right]. \quad (2.22)$$

Pour des raisons pratiques nous avons remplacé $E_c + 2t$ par A dans l'écriture de la matrice 2.22. Le bloc en haut à gauche représente le Hamiltonien du réservoir de gauche, le bloc du milieu celui du fil et le bloc en bas à droite celui du réservoir de droite. Ce hamiltonien présente, sous cette forme matricielle, des dimensions infinies et ne peut donc pas être implémenté numériquement. La notion de *self-energy* permet de remédier à ce problème. Il s'agit de considérer l'influence du réservoir comme une perturbation appliquée à la partie active du système. Raisonons par exemple sur le premier point $|1\rangle$ de la partie active et appliquons le hamiltonien H :

$$H |1\rangle = -t |0\rangle + (E_c + 2t + U_1) |1\rangle - t |2\rangle. \quad (2.23)$$

$|0\rangle$ appartenant au réservoir de gauche, il ne fait pas partie de la base des $|r\rangle$ (r allant de 1 à N). Nous souhaitons néanmoins exprimer l'influence du réservoir sur les sites de la partie active. Nous allons pour cela exprimer $|0\rangle$ en fonction de ces sites. Soit un électron qui sort du fil (en $|1\rangle$) pour aller dans le réservoir de gauche (en $|0\rangle$). Sa fonction d'onde a la forme suivante :

$$\psi_k(1) = e^{-ikx} \quad (2.24)$$

avec k son vecteur d'onde et x sa position le long du système. On en déduit :

$$\psi_k(0) = e^{-ik(x-a)} = \psi_k(1) e^{ika}. \quad (2.25)$$

Donc d'après 2.23 et 2.25 nous avons :

$$H |1\rangle = \underbrace{(E_c + 2t + U_1) |1\rangle - t |2\rangle}_{H_{fil}} - \underbrace{te^{ika} |1\rangle}_{\Sigma_G}. \quad (2.26)$$

Le réservoir G n'a pas d'influence sur les autres sites du système car on ne considère qu'un couplage au premier voisin. L'influence de G sur la partie active Σ_G s'écrit donc dans la base des $|r\rangle$:

$$\Sigma_G(E) = \begin{bmatrix} -te^{ika} & 0 & \cdots & 0 \\ 0 & 0 & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 \end{bmatrix}. \quad (2.27)$$

De la même manière, nous pouvons calculer Σ_D pour le réservoir de droite :

$$\Sigma_D(E) = \begin{bmatrix} 0 & \cdots & \cdots & 0 \\ \vdots & \ddots & & \vdots \\ \vdots & & 0 & 0 \\ 0 & \cdots & 0 & -te^{ika} \end{bmatrix}. \quad (2.28)$$

On peut alors définir le composant complet par $H_0 + \Sigma_G + \Sigma_D$. Ceci n'est pas un

hamiltonien car il dépend de E . Mais nous verrons que l'on peut calculer la fonction de Green du système :

$$G(E) = \frac{1}{E - H_0 - \Sigma_G - \Sigma_D}. \quad (2.29)$$

Cet opérateur, qui définit tout le système (fil et réservoirs), a la même dimension que l'opérateur de Green du fil seul (région active). L'intérêt des *self-energies* dans cet exemple est de transformer un opérateur de taille infinie en un opérateur de taille finie et de pouvoir ainsi traiter numériquement le problème.

Remarque : Le calcul présenté ci-dessus n'est pas très "rigoureux" (ni général). Un calcul de $\Sigma_{G,D}$ général sera présenté dans le chapitre 2.3.2.

2.1.4 Conséquences physiques des *self-energies* de contact

Nous venons de voir que les *self-energies* ont un grand intérêt pour le calcul numérique. Elles nous permettent en effet de modéliser un système ouvert (*i.e.* infini) avec un opérateur de taille finie. Mais les *self-energies* que nous présentons dans la partie qui suit ne sont pas seulement un outil mathématique, elles ont aussi un sens physique. Soit H_0 le hamiltonien de la partie active du composant sans *self-energy*. En le diagonalisant, nous obtenons des valeurs propres E_k parfaitement définies. Avec les *self-energies*, nous avons :

$$H_0 + \Sigma_G + \Sigma_D = H_{rel} - \frac{i}{2}(\Gamma_G + \Gamma_D) \quad (2.30)$$

où H_{rel} est la partie réelle donnée par :

$$H_{rel} = H_0 + \frac{\Sigma_G + \Sigma_G^+}{2} + \frac{\Sigma_D + \Sigma_D^+}{2}. \quad (2.31)$$

La partie imaginaire est constituée de $\Gamma_G = i(\Sigma_G - \Sigma_G^+)$ et de $\Gamma_D = i(\Sigma_D - \Sigma_D^+)$. Nous remarquons que la partie réelle est modifiée par les *self-energies*, ce qui modifie les énergies propres E_k du système. Les *self-energies* font aussi apparaître une partie imaginaire. Cette partie imaginaire crée un élargissement des niveaux d'énergie qui caractérise la perturbation induite par les réservoirs. Les niveaux ne sont donc plus parfaitement définis, le temps de vie d'une particule sur ces niveaux n'est donc plus infini et le transport devient possible. En effet, des niveaux d'énergie parfaitement définis (*i.e.* discrets), correspondent à des états non-perturbés du système isolé dans

lequel le transport électronique n'est pas possible. Nous verrons plus tard que la densité d'états est liée à la partie imaginaire de la fonction de Green :

$$G(E) = (E - H_{rel} + \frac{i}{2}(\Gamma_G + \Gamma_D))^{-1}. \quad (2.32)$$

Or G n'a de partie imaginaire que pour Γ_G et Γ_D non-nuls. Γ_G et Γ_D sont donc responsables de l'élargissement des niveaux, c'est pourquoi on les appelle fonctions d'élargissement.

Nous avons vu au chapitre 2.1.3, que la *self-energy* de gauche (équation 2.26) s'écrit dans le cas simple à une bande :

$$\Sigma_G(1, 1) = -te^{ika}. \quad (2.33)$$

Nous pouvons en déduire la fonction d'élargissement suivante

$$\begin{aligned} \Gamma_G(1, 1) &= i(\Sigma_G(1, 1) - \Sigma_G^+(1, 1)) \\ &= i(-te^{ika} + te^{-ika}) \\ &= 2t\sin(ka). \end{aligned} \quad (2.34)$$

Or la relation de dispersion dans le fil est donnée par $E = E_c + 2t(1 - \cos(ka))$ [46]. La vitesse de groupe v_g est donc :

$$v_g = \frac{1}{\hbar} \frac{dE}{dk} = \frac{2ta}{\hbar} \sin(ka). \quad (2.35)$$

Il y a ainsi une relation directe entre la vitesse de groupe et la fonction d'élargissement :

$$\Gamma_G(1, 1) = v_g \frac{\hbar}{a}. \quad (2.36)$$

Après avoir vu leur intérêt numérique, nous venons de voir que les *self-energies* avaient aussi un sens physique. Elles génèrent l'élargissement des niveaux d'énergie, et donc la limitation du temps de vie des électrons sur ces niveaux, indispensable à la notion de transport.

2.1.5 Equation de Dyson

Les fonctions de Green peuvent également se présenter comme des propagateurs, chaque propagateur caractérisant un chemin possible entre deux états. Considérons un électron sur l'état $|k_1\rangle$ pouvant aller sur l'état $|k_2\rangle$ par deux chemins différents. Si on appelle G_1 et G_2 les propagateurs associés à ces deux chemins, la probabilité de présence de l'électron en $|k_2\rangle$ est donnée par :

$$P = |G_1|^2 + |G_2|^2 + G_1^+ G_2 + G_2^+ G_1. \quad (2.37)$$

La notion de propagateur permet aussi de formaliser les interactions dues à une perturbation extérieure. Soit un système non perturbé dont le propagateur est donné par :

$$G_0(E) = \lim_{\eta \rightarrow 0^+} (E - H_0 + i\eta)^{-1} \quad (2.38)$$

η représentant l'élargissement d'un niveau d'énergie, la durée de vie d'un électron sur ce niveau est $1/\eta$. η tendant vers 0, la durée de vie des électrons dans ce système est infinie. Ce système est donc parfaitement à l'équilibre avec un système extérieur.

Si le système est maintenant perturbé, la perturbation peut, en général, se décrire à l'aide d'une *self-energy* $\Sigma(E)$. Si l'électron n'interagit pas avec cette perturbation son propagateur est donné par $G = G_0$. S'il interagit une seule fois il est donné par $G = G_0 \Sigma G_0$. Afin de considérer tous les chemins possibles, nous pouvons montrer que le propagateur d'un tel système perturbé est donné par [47] :

$$G = G_0 + G_0 \Sigma G_0 + G_0 \Sigma G_0 \Sigma G_0 + G_0 \Sigma G_0 \Sigma G_0 \Sigma G_0 + \dots \quad (2.39)$$

Cette expression peut se factoriser par G_0 :

$$\begin{aligned} G &= G_0 (1 + G_0 \Sigma + G_0^2 \Sigma^2 + G_0^3 \Sigma^3 + \dots) \\ &= \frac{G_0}{1 - G_0 \Sigma} \\ &= \frac{1}{G_0^{-1} - \Sigma}. \end{aligned} \quad (2.40)$$

On peut en effet calculer la suite géométrique $Suite = 1 + c + c^2 + c^3 + \dots$ puisque

Suite $-c \times$ *Suite* = 1. L'opérateur de Green du système perturbé s'écrit donc d'après l'équation 2.38 :

$$G(E) = (E - H_0 - \Sigma(E))^{-1} \quad (2.41)$$

où $\eta = 0$ puisque $\Sigma(E)$ a une partie imaginaire non nulle. $\Sigma(E)$ peut bien sûr représenter la perturbation apportée par un réservoir sur une région active comme nous l'avons vu précédemment, mais également définir n'importe quel type d'interaction (électron-électron, électron-phonon, électron-photon...). L'expression 2.41 est appelée l'équation de Dyson. On peut l'écrire différemment en repartant de l'équation 2.39 :

$$\begin{aligned} G &= G_0 + G_0 \Sigma G_0 + G_0 \Sigma G_0 \Sigma G_0 + G_0 \Sigma G_0 \Sigma G_0 \Sigma G_0 + \dots \\ &= G_0 + G_0^2 \Sigma + G_0^3 \Sigma^2 + G_0^4 \Sigma^3 + \dots \\ &= G_0 + G_0 \Sigma (G_0 + G_0^2 \Sigma + G_0^3 \Sigma^2 + \dots) \\ &= G_0 + G_0 \Sigma G. \end{aligned} \quad (2.42)$$

L'équation de Dyson est une équation générale permettant de formaliser n'importe quel système perturbé.

2.2 Opérateur de Green d'un système fermé

Nous avons vu précédemment (chapitre 1.4) que pour modéliser correctement un composant, il est nécessaire de résoudre de manière auto-cohérente les équations de Schrödinger et Poisson. En effet, le calcul du courant requiert la connaissance du potentiel dans la région active, qui est lui-même dépendant de la densité de charges.

Nous allons maintenant voir comment calculer la densité de charges du composant dans le formalisme des fonctions de Green. Soit un composant unidimensionnel orienté suivant la direction x , sans interaction (système fermé). La partie active du composant est définie par un Hamiltonien H_0 :

$$H_0 = E_c - \frac{\hbar^2}{2m^*} \frac{d^2}{dx^2} + U(x) \quad (2.43)$$

où E_c est le bas de la bande de conduction (ou de la sous-bande), m^* est la masse effective des électrons dans le matériau considéré dans la direction x et $U(x)$ l'énergie potentielle.

Soit $|k\rangle$ la base propre de l'opérateur H_0 et E_k les valeurs propres correspondantes :

$$H_0 |k\rangle = E_k |k\rangle. \quad (2.44)$$

La fonction d'onde pour l'état $|k\rangle$ en $|r\rangle$ est alors $\psi_k(r) = \langle k | r \rangle$. Pour un tel système on définit la densité d'états comme une somme de pics de Dirac :

$$\rho(E) = \sum_k \delta(E - E_k). \quad (2.45)$$

L'opérateur de Green correspondant à ce système défini dans l'équation 2.14 s'écrit :

$$G_0(E) = \lim_{\eta \rightarrow 0^+} \frac{1}{E - H_0 + i\eta}. \quad (2.46)$$

Exprimé dans la base des $|k\rangle$ la trace de cet opérateur vérifie l'expression suivante :

$$\begin{aligned} \text{Tr}(G_0(E)) &= \sum_k \lim_{\eta \rightarrow 0^+} \langle k | G_0(E) | k \rangle \\ &= \sum_k \lim_{\eta \rightarrow 0^+} \frac{1}{E - E_k + i\eta}. \end{aligned} \quad (2.47)$$

Notons que la trace de cet opérateur est indépendante de la base sur laquelle l'opérateur est développé. On peut maintenant calculer la partie imaginaire de cette trace :

$$\text{Im}(\text{Tr}(G_0(E))) = - \sum_k \lim_{\eta \rightarrow 0^+} \frac{\eta}{(E - E_k)^2 + \eta^2}. \quad (2.48)$$

La valeur de la limite tend vers l'infini lorsque E tend vers une valeur propre E_k et elle vaut 0 sinon. On a donc une succession de pics (Dirac) dont la largeur η tend vers 0. Ceci est assimilable à la densité d'états définie par les pics de Dirac

dans l'expression 2.45 :

$$\rho(E) = \sum_k \delta(E - E_k) = -\frac{1}{\pi} \text{Im}(\text{Tr}(G_0(E))). \quad (2.49)$$

Remarque : Le facteur $1/\pi$ résulte du fait que $\int dE \lim_{\eta \rightarrow 0^+} \frac{\eta}{(E - E_k)^2 + \eta^2} = \pi$. On peut également profiter de cette remarque pour voir l'importance de η dans l'écriture de $G_0(E)$ sans lequel cette intégrale ne serait pas définie.

Nous allons maintenant voir comment aboutir à la densité de porteurs en fonction de la position, c'est à dire sur chaque état $|r\rangle$. L'opérateur $G_0(E)$ dans la base des $|r\rangle$ est donné par l'expression :

$$\begin{aligned} \langle r | G_0(E) | r' \rangle &= \sum_k \lim_{\eta \rightarrow 0^+} \frac{\langle r | k \rangle \langle k | r' \rangle}{E - E_k + i\eta} \\ &= \sum_k \lim_{\eta \rightarrow 0^+} \frac{\psi_k(r)^* \psi_k(r')}{E - E_k + i\eta}. \end{aligned} \quad (2.50)$$

De la même façon que nous sommes arrivé à l'expression 2.49, nous pouvons vérifier que :

$$\text{Im}(\langle r | G_0(E) | r \rangle) = -\pi \sum_k |\psi_k(r)|^2 \delta(E - E_k). \quad (2.51)$$

Or si l'on calcule la densité de porteurs sur chaque état $|r\rangle$ avec une fonction de distribution $f_0(k)$ on a :

$$\begin{aligned} n(r) &= \sum_k f_0(k) |\psi_k(r)|^2 \\ &= \int dE f_0(E) \sum_k |\psi_k(r)|^2 \delta(E - E_k) \\ &= \int dE f_0(E) \left[-\frac{1}{\pi} \text{Im}(\langle r | G_0(E) | r \rangle) \right] \\ &= \int dE f_0(E) \rho_r(E) \end{aligned} \quad (2.52)$$

avec

$$\rho_r(E) = \left[-\frac{1}{\pi} \text{Im} (\langle r | G_0(E) | r \rangle) \right] \quad (2.53)$$

qui définit la densité locale d'états. La partie imaginaire de la diagonale de l'opérateur de Green dans l'espace $|r\rangle$ représente donc la densité locale d'états. Dans le cas où nous connecterions des *self-energies* au système, la connaissance de l'opérateur de Green nous permettrait de calculer la densité de porteurs dans tout le système et donc de faire le calcul auto-cohérent dans un système ouvert.

2.3 Système ouvert Multibandes en 3 dimensions

Dans le chapitre précédent nous avons calculé la densité d'états d'un système physique unidimensionnel sans interaction dans l'approximation de la masse effective. Nous allons maintenant le faire dans le cas le plus général possible. Une partie du travail de cette thèse a été de passer d'un hamiltonien simple bande à un hamiltonien plus général afin de pouvoir traiter une gamme plus large de problèmes. Nous allons considérer un conducteur en 3 dimensions (3D) connecté à deux réservoirs semi-infini. Nous utiliserons un hamiltonien multibandes afin de traiter le cas où les sous-bandes ne sont pas paraboliques et couplées entre elles. Nous verrons notamment qu'une approche multibandes est indispensable pour décrire le transport des trous dans la bande de valence. Nous utiliserons ensuite les *self-energies* pour simuler l'influence des réservoirs comme nous l'avons vu au chapitre 2.1.3.

2.3.1 Hamiltonien multibandes

Soit le hamiltonien 3D à une bande du matériau massif :

$$H_{\text{massif}(1 \text{ bande})} = \frac{\hbar^2 k_x^2}{2m_0} + \frac{\hbar^2 k_y^2}{2m_0} + \frac{\hbar^2 k_z^2}{2m_0}. \quad (2.54)$$

Les dimensions sont infinies et le potentiel périodique dans les trois directions de l'espace. En revanche, en présence de confinement suivant une ou plusieurs directions, le vecteur d'onde n'est plus défini et doit être remplacé dans le cadre de la

théorie $\mathbf{k} \cdot \mathbf{p}$ (que nous étudierons au chapitre 4) par : $k_z \rightarrow -i \frac{\partial}{\partial z}$ si le confinement a lieu selon z . L'équation 2.54 devient :

$$H_{1D(1 \text{ bande})} = \frac{\hbar^2 k_x^2}{2m_0} + \frac{\hbar^2 k_y^2}{2m_0} - \frac{\hbar^2 \partial^2}{2m_0 \partial z^2}. \quad (2.55)$$

Comme nous l'avons vu dans le chapitre 2.1.3, il faut projeter cet Hamiltonien sur la base de sites (positions de 1 à N_z selon la direction z) pour pouvoir le traiter numériquement. On obtient :

$$H_{1D(1 \text{ bande})}(k_x, k_y) = \frac{\hbar^2 k_x^2}{2m_0} + \frac{\hbar^2 k_y^2}{2m_0} + \underbrace{\begin{bmatrix} 2t & -t & 0 & \cdots & 0 \\ -t & 2t & \ddots & \ddots & \vdots \\ 0 & -t & \ddots & -t & 0 \\ \vdots & \ddots & \ddots & 2t & -t \\ 0 & \cdots & 0 & -t & 2t \end{bmatrix}}_{N_z \times N_z} \quad (2.56)$$

avec $t = \frac{\hbar^2}{2m_0 a^2}$ et a le pas entre les sites. Passons maintenant au cas où deux bandes sont couplées entre elles. Dans la direction z , le Hamiltonien du matériau massif s'écrit :

$$H_{\text{massif}(2 \text{ bandes})}(k_x = 0, k_y = 0, k_z) = \begin{bmatrix} \frac{\hbar^2 k_z^2}{2m_0} + E_g & Pk_z \\ Pk_z & \frac{\hbar^2 k_z^2}{2m_0} \end{bmatrix} \quad (2.57)$$

avec P un paramètre d'ajustement formalisant le couplage entre les deux bandes et E_g l'énergie séparant les deux bandes en $k_z = 0$.

En considérant un confinement selon z nous obtenons :

$$H_{1D(2 \text{ bandes})}(k_x = 0, k_y = 0) = \underbrace{\begin{bmatrix} \begin{bmatrix} \frac{\hbar^2}{l^2 m_0} + E_g & 0 \\ 0 & \frac{\hbar^2}{l^2 m_0} \end{bmatrix} & \begin{bmatrix} \frac{-\hbar^2}{2l^2 m_0} & \frac{iP}{l} \\ \frac{iP}{l} & \frac{-\hbar^2}{2l^2 m_0} \end{bmatrix} & 0 & \cdots \\ \begin{bmatrix} \frac{-\hbar^2}{2l^2 m_0} & \frac{-iP}{l} \\ \frac{-iP}{l} & \frac{-\hbar^2}{2l^2 m_0} \end{bmatrix} & \ddots & \ddots & \ddots \\ 0 & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{bmatrix}}_{2N_z \times 2N_z} \quad (2.58)$$

On passe ici d'une matrice de dimension N_z (dans une description à une bande) à une matrice de dimension $2N_z$. De la même manière, si un hamiltonien à N bandes est considéré :

$$H_{1D(N \text{ bandes})}(k_x = 0, k_y = 0) = \underbrace{\begin{bmatrix} [\alpha_1] & [\beta] & 0 & \cdots & 0 \\ [\beta^+] & [\alpha_2] & [\beta] & \ddots & \vdots \\ 0 & [\beta^+] & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & [\alpha_{N_z-1}] & [\beta] \\ 0 & \cdots & 0 & [\beta^+] & [\alpha_{N_z}] \end{bmatrix}}_{N.N_z \times N.N_z} \quad (2.59)$$

avec α et β des matrices de dimensions N . Les α représentent les plans perpendiculaires au transport et les β les couplages entre ces plans. Nous pouvons écrire un hamiltonien général à N bandes pour n'importe quelle configuration de la manière suivante :

$$H = \begin{bmatrix} [\alpha_1] & [\beta] & 0 & \cdots & 0 \\ [\beta^+] & [\alpha_2] & [\beta] & \ddots & \vdots \\ 0 & [\beta^+] & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & [\alpha_{M-1}] & [\beta] \\ 0 & \cdots & 0 & [\beta^+] & [\alpha_M] \end{bmatrix} \quad (2.60)$$

Dans le cas très simple d'un composant unidimensionnel à une bande (chapitre 2.1.3), $[\alpha]$ et $[\beta]$ sont des scalaires tels que $[\alpha_i] = E_c + 2t + U_i$ et $[\beta] = -t$ comme

nous avons pu le voir dans l'équation 2.20. H est de dimension M .

Dans le cas d'un composant tridimensionnel à six bandes, que nous étudierons au chapitre 4 avec un confinement selon deux directions (y et z) et un potentiel qui varie selon l'autre direction (x), $[\alpha]$ et $[\beta]$ sont des matrices de dimension $6 \times N_y \times N_z$. H est de dimension $6 \times N_y \times N_z \times N_x$.

2.3.2 *Self-energy* de contact d'un Hamiltonien multibandes

Le composant que nous voulons modéliser est connecté à deux réservoirs G et D . Il nous faut donc calculer les *Self-energies* simulant les réservoirs semi-infinis connectés au composant qui a pour hamiltonien $H_{3D(N \text{ bandes})}$. Le hamiltonien total s'écrit :

$$H_{\text{composant}} = \begin{bmatrix} [H_G] & [\tau] & [0] \\ [\tau^+] & [H_{3D(N \text{ bandes})}] & [\tau] \\ [0] & [\tau^+] & [H_D] \end{bmatrix} \quad (2.61)$$

avec

$$\tau = \begin{bmatrix} \vdots & & & \\ [0] & [0] & & \\ [\beta] & [0] & \dots & \end{bmatrix} \quad (2.62)$$

le couplage entre les différentes parties,

$$H_G = \begin{bmatrix} \ddots & \ddots & \ddots & \vdots \\ \ddots & \ddots & [\beta] & [0] \\ \ddots & [\beta^+] & [\alpha_1] & [\beta] \\ \dots & [0] & [\beta^+] & [\alpha_1] \end{bmatrix} \quad (2.63)$$

le hamiltonien du réservoir semi-infini de gauche et

$$H_D = \begin{bmatrix} [\alpha_N] & [\beta] & [0] & \cdots \\ [\beta^+] & [\alpha_N] & [\beta] & \ddots \\ [0] & [\beta^+] & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{bmatrix} \quad (2.64)$$

le hamiltonien du r eservoir semi-infini de droite. Nous pr esentons ci-dessous le calcul de la *Self-energy* Σ_G du r eservoir de droite. Un raisonnement identique permet d'aboutir  a l'expression de son homologue du contact de gauche. Pour cela, nous consid erons un conducteur de hamiltonien H_C connect e, via la matrice de couplage τ ,  a un r eservoir de hamiltonien H_D . La matrice du hamiltonien pr esente la forme suivante :

$$\begin{bmatrix} [H_C] & [\tau] \\ [\tau^+] & [H_D] \end{bmatrix}. \quad (2.65)$$

La matrice de la fonction de Green G du syst eme total s' ecrit donc :

$$G = \begin{bmatrix} [G_C] & [G_{CD}] \\ [G_{DC}] & [G_D] \end{bmatrix} = \begin{bmatrix} [(E + i\eta)I - H_C] & -[\tau] \\ -[\tau^+] & [(E + i\eta)I - H_D] \end{bmatrix}^{-1} \quad (2.66)$$

o u $[G_C]$ et $[G_D]$ sont respectivement les fonctions de Green du conducteur et du r eservoir D . $[G_{CD}]$ et $[G_{DC}]$ sont les fonctions de Green qui couplent le conducteur au r eservoir.

Nous d eduisons de l' egalit e (2.66) le syst eme  a deux  equations suivant :

$$[(E + i\eta)I - H_C][G_C] - [\tau][G_{DC}] = I \quad (2.67)$$

$$-[\tau^+][G_C] + [(E + i\eta)I - H_D][G_{DC}] = 0. \quad (2.68)$$

A partir de l' equation (2.68) nous pouvons  ecrire :

$$[G_{DC}] = [G_D][\tau^+][G_C] \quad (2.69)$$

o u

$$[G_D] = [(E + i\eta)I - H_D]^{-1}. \quad (2.70)$$

En utilisant 2.69 dans 2.67, nous obtenons :

$$[G_C] = [[(E + i\eta)I - H_C] - [\tau][G_D][\tau^+]]^{-1}. \quad (2.71)$$

Nous pouvons donc en déduire que $\Sigma_D = \tau G_D \tau^+$. H_D étant de taille infinie, il ne semble pas évident de calculer G_D et encore moins Σ_D . Supposons néanmoins que G_D soit connue et que son écriture matricielle s'écrive :

$$G_D = \begin{bmatrix} [G_{00}] & [G_{10}] & \cdots \\ [G_{01}] & [G_{11}] & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \quad (2.72)$$

où les $[G_{ij}]$ sont des blocs $N.N_y.N_z \times N.N_y.N_z$ de même dimension que $[\beta]$. On peut calculer Σ_D tel que :

$$\Sigma_D = \tau G_D \tau^+ = \begin{bmatrix} [0] & \cdots & \cdots & [0] \\ \vdots & \ddots & & \vdots \\ \vdots & & [0] & [0] \\ [0] & \cdots & [0] & [\beta G_{00} \beta^+] \end{bmatrix}. \quad (2.73)$$

Σ_D est donc de dimension $N \times N_x \times N_y \times N_z$ et ne dépend que du premier bloc de dimension $N \times N_y \times N_z$ de l'opérateur de Green G_D . Il nous reste maintenant à calculer le premier bloc $[G_{00}]$. Pour cela nous allons considérer deux hamiltoniens pour le réservoir de droite :

Un Hamiltonien complet tel que nous l'avons déjà défini :

$$[H_D] = \begin{bmatrix} [\alpha_N] & [\beta] & [0] & \cdots \\ [\beta^+] & [\alpha_N] & [\beta] & \ddots \\ [0] & [\beta^+] & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{bmatrix}, \quad (2.74)$$

ayant comme opérateur de Green (également déjà défini) :

$$G_D = \begin{bmatrix} [G_{00}] & [G_{10}] & \cdots \\ [G_{01}] & [G_{11}] & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}. \quad (2.75)$$

Nous considérons également un système similaire dans laquelle le premier plan n'est pas connecté. Son hamiltonien est donc donné par :

$$[h_D] = \begin{bmatrix} [\alpha_N] & [0] & [0] & \cdots \\ [0] & [\alpha_N] & [\beta] & \ddots \\ [0] & [\beta^+] & [\alpha_N] & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{bmatrix} \quad (2.76)$$

et son opérateur de Green a la forme :

$$g_D = \begin{bmatrix} [g_{00}] & [0] & \cdots & \cdots \\ [0] & [g_{11}] & [g_{21}] & \ddots \\ \vdots & [g_{21}] & [g_{22}] & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{bmatrix}. \quad (2.77)$$

La seule différence entre ces deux réservoirs tient à la présence ou pas du couplage $[\beta]$ entre le site $|0\rangle$ et le reste du système. G_D représente donc le même système que g_D à une "perturbation près". Il est ainsi possible de relier ces deux opérateurs de Green par la formule de Dyson (Equ. 2.42) :

$$G_D = g_D + g_D \delta G_D \quad (2.78)$$

où δ représente la perturbation qui est donnée par :

$$\delta = \begin{bmatrix} [0] & [\beta] & [0] & \cdots \\ [\beta^+] & [0] & [0] & \ddots \\ [0] & [0] & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{bmatrix}. \quad (2.79)$$

Calculons les premiers blocs de la matrice $g_D \delta G_D$:

$$g_D \delta G_D = \begin{bmatrix} [g_{00}\beta G_{10}] & [g_{00}\beta^+ G_{01}] & [?] & \cdots \\ [g_{11}\beta^+ G_{00}] & [g_{11}\beta^+ G_{01}] & [?] & \ddots \\ [?] & [?] & [?] & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{bmatrix}. \quad (2.80)$$

Parmis ces blocs, seul le premier terme $[G_{00}]$ intervient dans le calcul de Σ_D . D'après les deux équations précédentes nous avons :

$$\begin{aligned} G_{00} &= g_{00} + g_{00}\beta G_{10} \\ G_{10} &= 0 + g_{11}\beta^+ G_{00}. \end{aligned} \quad (2.81)$$

On a donc :

$$G_{00} = g_{00} + g_{00}\beta g_{11}\beta^+ G_{00}. \quad (2.82)$$

Or $g_{11} = G_{00}$ puisqu'ils représentent tous les deux le premier bloc de l'opérateur de Green d'un réservoir infini. On a finalement :

$$G_{00} = g_{00} + g_{00}\beta G_{00}\beta^+ G_{00} \quad (2.83)$$

avec $g_{00} = [EI - \alpha_N]^{-1}$ l'opérateur de Green d'un site isolé. Il faut donc résoudre l'équation matricielle 2.83 afin de trouver G_{00} pour chaque énergie E considérée. A cette fin, nous utilisons une technique mathématique qui repose sur la décomposition de Schur [12]. Cette approche a l'avantage d'être non itérative et rapide.

2.3.3 Décomposition de Schur

La décomposition de Schur est une méthode de décomposition de matrices. Elle permet de décomposer une matrice carré complexe A en trois matrices, une étant triangulaire supérieure et les deux autres unitaires. Comme nous venons de le voir, le calcul des *self-energies* nécessite de connaître le premier terme G_{00} de l'opérateur de Green des réservoirs. Pour cela, nous devons résoudre l'équation matricielle 2.83. Nous allons voir dans ce chapitre comment résoudre cette équation quadratique en utilisant une décomposition de Schur. Nous allons d'abord modifier l'équation 2.83 :

$$G_{00} = g_{00} + g_{00} \underbrace{\beta G_{00} \beta^+}_{\Sigma} G_{00}. \quad (2.84)$$

D'après Dyson l'équation 2.84 peut se réécrire sous la forme :

$$\begin{aligned} G_{00} &= (E - \alpha - \Sigma)^{-1} \\ &= (E - \alpha - \beta G_{00} \beta^+)^{-1}. \end{aligned} \quad (2.85)$$

Pour un système orienté selon la direction x et confiné selon les deux autres directions, α représente le Hamiltonien transverse et sa matrice est de dimension $N \times N_y \times N_z$.

En notant $X = G_{00} \beta^+$ l'équation 2.85 devient :

$$X = (E - \alpha - \beta X)^{-1} \beta^+ \quad (2.86)$$

$$\beta X^2 + (\alpha - E)X + \beta^+ = 0 \quad (2.87)$$

$$X^2 + \underbrace{\beta^{-1}(\alpha - E)}_B X + \underbrace{\beta^{-1}\beta^+}_C = 0. \quad (2.88)$$

Nous obtenons donc une équation matricielle quadratique de la forme $X^2 + BX + C = 0$ à résoudre. Pour cela, nous allons utiliser la décomposition de Schur.

Définissons tout d'abord la matrice suivante :

$$F = \begin{bmatrix} 0 & I \\ -C & -B \end{bmatrix} \quad (2.89)$$

de dimension $2 \times N \times N_y \times N_z$ avec I la matrice identité de dimension $N \times N_y \times N_z$. Nous multiplions ensuite F par une matrice telle que :

$$F \begin{bmatrix} I \\ X \end{bmatrix} = \begin{bmatrix} X \\ -C - BX \end{bmatrix}. \quad (2.90)$$

X est solution de 2.88 si et seulement si [12] :

$$F \begin{bmatrix} I \\ X \end{bmatrix} = \begin{bmatrix} I \\ X \end{bmatrix} X \quad (2.91)$$

c'est à dire si $X^2 + BX + C = 0$.

F peut se réécrire à l'aide d'une décomposition de Schur :

$$F = QTQ^+ \quad (2.92)$$

où Q est une matrice unitaire ($QQ^+ = 1$) et T une matrice triangulaire supérieure. Ces matrices sont de dimension $2 \times N_y \times N_z$ et s'écrivent sous la forme :

$$Q = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \quad (2.93)$$

et

$$T = \begin{bmatrix} T_{11} & T_{12} \\ O & T_{22} \end{bmatrix}. \quad (2.94)$$

Toutes les solutions de l'équation quadratique sont de la forme :

$$X = Q_{21}Q_{11}^{-1}. \quad (2.95)$$

En effet, admettons que $X = Q_{21}Q_{11}^{-1}$. Nous avons $FQ = QT$:

$$\begin{bmatrix} 0 & I \\ -C & -B \end{bmatrix} \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \begin{bmatrix} T_{11} & T_{12} \\ O & T_{22} \end{bmatrix}. \quad (2.96)$$

Et d'après l'équation 2.96 :

$$Q_{21} = Q_{11}T_{11} \quad (2.97)$$

et

$$-CQ_{11} - BQ_{21} = Q_{21}T_{11}. \quad (2.98)$$

La multiplication de 2.98 par Q_{11}^{-1} conduit à :

$$-C - BQ_{21}Q_{11}^{-1} = Q_{21}T_{11}Q_{11}^{-1}. \quad (2.99)$$

D'après 2.97 nous avons :

$$T_{11} = Q_{11}^{-1} Q_{21}. \quad (2.100)$$

On remplace 2.100 dans 2.99 et nous obtenons :

$$-C - \underbrace{BQ_{21}Q_{11}^{-1}}_X = \underbrace{Q_{21}Q_{11}^{-1}}_X \underbrace{Q_{21}Q_{11}^{-1}}_X. \quad (2.101)$$

On a bien $X^2 + BX + C = 0$ donc $X = Q_{21}Q_{11}^{-1}$ est bien solution de cette équation. Cette méthode nous permet donc d'obtenir $X = G_{00}\beta^+$ et on peut en déduire la *self-energy* Σ_D qui ne dépend que du bloc $[\beta G_{00}\beta^+]$.

2.4 Calcul du courant dans le formalisme des fonctions de Green

2.4.1 Calcul de la densité de charge et du courant

Connaissant les *self-energies* des réservoirs, nous pouvons désormais calculer la densité de porteurs et le courant dans le dispositif.

La densité de porteurs dans le dispositif, comme nous l'avons vu dans l'équation 2.52 dans le cas d'un système fermé, se calcule à partir de la densité locale d'états. Cette dernière est directement liée à la partie imaginaire de G . Nous allons donc calculer cette partie imaginaire :

$$\text{Im}(G) = \frac{i}{2}(G - G^+). \quad (2.102)$$

Si nous nous plaçons dans le cas d'un modèle à deux réservoirs, l'équation 2.32 donne :

$$G = (E - H_{rel} + \frac{i}{2}(\Gamma_G + \Gamma_D))^{-1} \quad (2.103)$$

et son conjugué s'écrit

$$G^+ = (E - H_{rel} - \frac{i}{2}(\Gamma_G + \Gamma_D))^{-1}. \quad (2.104)$$

Nous avons donc :

$$\text{Im}(G) = \frac{i}{2}(G - G^+) = \frac{1}{2} \frac{\Gamma_G + \Gamma_D}{(E - H_{rel})^2 + \frac{1}{4}(\Gamma_G + \Gamma_D)^2}. \quad (2.105)$$

La trace de la partie imaginaire est, à un facteur près, la densité d'états (DOS) dans le composant. La densité de porteurs s'obtient ensuite en multipliant la DOS par la fonction de Fermi-Dirac du système. Le système étant hors-équilibre, deux fonctions de Fermi différentes (relatives à chacun des réservoirs) interviennent. Il est donc nécessaire de calculer séparément les contributions associées à chaque réservoir. Pour cela, nous allons faire le calcul suivant :

$$\begin{aligned} G(\Gamma_G + \Gamma_D)G^+ &= \frac{\Gamma_G + \Gamma_D}{(E - H_{rel} + \frac{i}{2}(\Gamma_G + \Gamma_D))(E - H_{rel} - \frac{i}{2}(\Gamma_G + \Gamma_D))} \\ &= \frac{\Gamma_G + \Gamma_D}{(E - H_{rel})^2 + \frac{1}{4}(\Gamma_G + \Gamma_D)^2} \\ &= 2\text{Im}(G). \end{aligned} \quad (2.106)$$

Nous avons donc séparé en deux contributions la partie imaginaire de G qui peut donc s'écrire :

$$\text{Im}(G) = \frac{1}{2}(G\Gamma_G G^+ + G\Gamma_D G^+), \quad (2.107)$$

où $G\Gamma_G G^+$ est associé à la densité d'états des électrons injectés depuis le réservoir de gauche et $G\Gamma_D G^+$ de ceux injectés depuis le réservoir de droite [11]. Nous définissons les fonctions spectrales A_G et A_D telles que :

$$A_G = G\Gamma_G G^+ \quad (2.108)$$

et

$$A_D = G\Gamma_D G^+. \quad (2.109)$$

Nous pouvons écrire les densités locales d'états $\frac{1}{2\pi} \langle r|A_G|r \rangle$ et $\frac{1}{2\pi} \langle r|A_D|r \rangle$ respectivement pour les charges injectées depuis le réservoir G et pour celles injectées depuis le réservoir D . La densité totale de charges dans le composant s'exprime alors de la manière suivante :

$$n(r) = \frac{1}{2\pi} \int [f_0(E, \mu_G) \langle r|A_G|r \rangle + f_0(E, \mu_D) \langle r|A_D|r \rangle] dE \quad (2.110)$$

avec $f_0(E, \mu_G)$ et $f_0(E, \mu_D)$ respectivement les fonctions de Fermi des réservoirs G et D .

Nous pouvons également calculer la transmission du composant et en déduire le courant le transversant grâce à la formule de Landauer. En effet, la transmission peut s'écrire [11] :

$$\begin{aligned} T(E) &= \text{trace}(A_D \Gamma_G) \\ &= \text{trace}(A_G \Gamma_D), \end{aligned} \quad (2.111)$$

et la formule de Landauer (1.34) permet de calculer le courant à partir de la transmission. Le courant vérifie donc l'expression :

$$\begin{aligned} I &= -\frac{2e}{h} \int [f_0(E, \mu_G) - f_0(E, \mu_D)] T(E) dE \\ &= -\frac{2e}{h} \int [f_0(E, \mu_G) - f_0(E, \mu_D)] \text{trace}(A_D \Gamma_G) dE \\ &= -\frac{2e}{h} \int [f_0(E, \mu_G) - f_0(E, \mu_D)] \text{trace}(A_G \Gamma_D) dE. \end{aligned} \quad (2.112)$$

Ainsi la connaissance des *self-energies*, du potentiel et du hamiltonien de la partie active, permet de calculer la fonction de Green du système à partir de laquelle seront déduits le courant et les charges. Le cas des systèmes 3D, dont la fonction de Green nécessite l'inversion d'une matrice de très grande taille, constitue un défi numérique. Afin de réduire le temps de calcul, nous avons utilisé un algorithme de récursion basé sur l'équation de Dyson [48].

2.4.2 Algorithme de récursion

Considérons le cas d'un fil discrétisé en N_x points selon la direction du transport, et en N_y et N_z points selon les directions de confinement. L'algorithme de récursion permet de calculer la fonction de Green en inversant N_x fois des matrices de taille

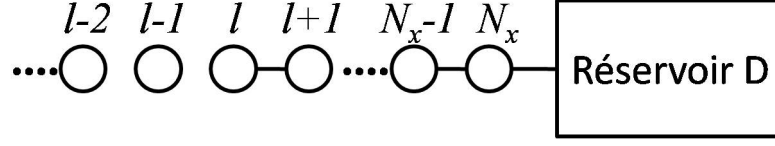


FIG. 2.2: Représentation schématique du plan l connecté aux plans de droite et au réservoir D, et libre à gauche.

$N_y N_z \times N_y N_z$ au lieu de matrices de taille $N_x N_y N_z \times N_x N_y N_z$. Dans le calcul qui suit, la notation $G(l, n)$ représente la matrice de couplage de dimension $N_y N_z$ entre les plans l et n . Comme nous l'avons vu précédemment (2.73), la *self-energy* des contacts s'exprime comme :

$$\Sigma_D = \tau G_D \tau^+ = \begin{bmatrix} [0] & \cdots & \cdots & [0] \\ \vdots & \ddots & & \vdots \\ \vdots & & [0] & [0] \\ [0] & \cdots & [0] & [\beta G_{D00} \beta^+] \end{bmatrix}. \quad (2.113)$$

De la même manière, on peut montrer que :

$$\Sigma_G = \tau^+ G_G \tau = \begin{bmatrix} [\beta^+ G_{G00} \beta] & [0] & \cdots & [0] \\ [0] & [0] & & \vdots \\ \vdots & & \ddots & \vdots \\ [0] & \cdots & \cdots & [0] \end{bmatrix}. \quad (2.114)$$

Or $\Gamma_D = i(\Sigma_D - \Sigma_D^+)$ et $\Gamma_G = i(\Sigma_G - \Sigma_G^+)$ donc seuls $\Gamma_D(N_x, N_x)$ et $\Gamma_G(1, 1)$ sont non nuls. Donc d'après les équations 2.108 et 2.109 :

$$A_G(l, l) = G(l, 1) \Gamma_G(1, 1) G^+(1, l) \quad (2.115)$$

et

$$A_D(l, l) = G(l, N_x) \Gamma_D(N_x, N_x) G^+(N_x, l) \quad (2.116)$$

pour l allant de 1 à N_x .

Nous allons maintenant décrire l'algorithme de récursion. Nous commençons par

définir $G_{right}(l, l)$ la fonction de Green du plan l connectée à droite mais pas à gauche comme décrit sur la figure 2.2. Afin de calculer cette fonction, nous calculons dans un premier temps celle du plan N_x . Pour cela nous avons besoin de la *self-energy* du contact de drain $\Sigma_D(N_x, N_x) = \beta G_{D00} \beta^+$ calculée à l'aide de la décomposition de Schur. Nous rappelons que β est la matrice de couplage entre deux plans successifs.

Nous pouvons alors définir $G_{right}(N_x, N_x)$ la fonction de Green du dernier plan de droite N_x libre à gauche et connectée au réservoir D :

$$G_{right}(N_x, N_x) = (E - H - U(N_x) - \Sigma_D(N_x, N_x))^{-1} \quad (2.117)$$

avec H le hamiltonien d'une section transverse et $U(N_x)$ l'énergie potentielle transverse du plan N_x . Nous avons maintenant tous les éléments pour commencer la récursion. Partons du plan le plus à droite (plan N_x) et calculons de plan en plan G_{right} jusqu'au plan 1 de la manière suivante :

$$G_{right}(l, l) = (E - H - U(l) - \beta G_{right}(l + 1, l + 1) \beta^+)^{-1} \quad (2.118)$$

l allant de $N_x - 1$ à 1.

A la fin de la boucle, nous avons calculé $G_{right}(1, 1)$, la fonction de Green du premier plan de la zone active (plan 1), prenant en compte l'influence du reste du dispositif en dehors de celle du contact de source. Comme nous l'avons fait pour le drain, nous calculons avec l'algorithme de Schur, la fonction de Green du contact de source et en déduisons la *self-energy* de la source :

$$\Sigma_G(1, 1) = \beta^+ G_{G00} \beta. \quad (2.119)$$

Nous calculons également la *self-energy* associée à la partie droite du dispositif :

$$\Sigma_{right} = \beta G_{right}(1, 1) \beta^+ \quad (2.120)$$

afin de calculer la fonction de Green $G(1, 1)$ au premier plan en considérant l'influence de la partie droite et du contact de source :

$$G(1, 1) = (E - H - U(1) - \Sigma_G - \Sigma_{right})^{-1}. \quad (2.121)$$

L'obtention de A_G nécessite le calcul des $G(l, 1)$ fonctions de Green de couplage entre les plans l et le plan 1. L'équation de Dyson donne :

$$G = G_{right} + G_{right}V^{1,2}G \quad (2.122)$$

avec

$$V^{1,2} = \begin{bmatrix} [0] & [\beta] & [0] & \cdots \\ [\beta^+] & [0] & [0] & \ddots \\ [0] & [0] & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{bmatrix} \quad (2.123)$$

la matrice couplant le plan 1 au plan 2. Nous montrons que :

$$G(2, 1) = G_{right}(2, 1) + G_{right}(2, 2)\beta^+G(1, 1). \quad (2.124)$$

G_{right} n'étant pas connectée à gauche, $G_{right}(2, 1)$ est nulle et finalement :

$$G(2, 1) = G_{right}(2, 2)\beta^+G(1, 1). \quad (2.125)$$

Nous pouvons généraliser cette équation à tous les plans l jusqu'à $l = N_x$:

$$G(l, 1) = G_{right}(l, l)\beta^+G(l - 1, 1). \quad (2.126)$$

De manière similaire nous pouvons calculer les $G(l, N_x)$ et ainsi obtenir la densité totale (2.110) et le courant (2.112) traversant le dispositif.

Chapitre 3

Transistor à nanofil *notch*-MOSFET

Dans ce chapitre nous allons décrire une architecture innovante pour les transistors à nanofil. Cette architecture appelée notch-MOSFET (notch pour encoche) a pour but d'améliorer le contrôle du courant qui se dégrade avec la diminution des dimensions du transistor.

Nous verrons que cette architecture nécessite une simulation du transport en trois dimensions dans l'espace réel. En effet l'approche dite de "l'espace des modes" ne permet pas de simuler correctement un transistor à nanofil n'ayant pas un canal de conduction invariant le long de l'axe de transport.

Dans une première partie, nous verrons le principe de base de l'architecture notch-MOSFET. Dans une deuxième partie, nous présenterons le modèle utilisé pour simuler cette architecture. Dans une troisième partie, nous étudierons l'influence de diverses caractéristiques du notch.

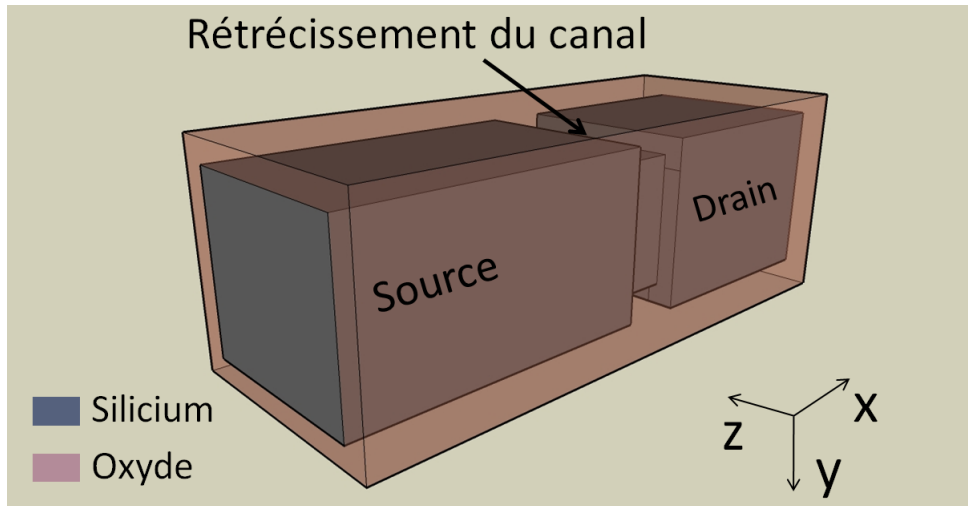


FIG. 3.1: Représentation schématique d'un nanofil en silicium entouré d'oxyde. Le canal de conduction présente un rétrécissement proche du drain.

3.1 Principe de base

Avec la diminution effrénée des dimensions des transistors, les effets quantiques deviennent de plus en plus importants. Le courant tunnel augmente notamment avec la diminution de la longueur de grille. Ce dernier augmente de façon importante le courant *OFF* (courant à l'état bloquant) et légèrement le courant *ON* (courant à l'état passant). Il s'en suit une dégradation du rapport I_{ON}/I_{OFF} qui représente un facteur mérite essentiel dans les transistors. Il est donc important de trouver des solutions pour diminuer ce phénomène. Le passage par effet tunnel est inévitable car il est intrinsèque au caractère quantique des particules. En revanche il est possible de le réduire en tirant partie de ce même caractère quantique. En effet, un autre effet quantique bien connu est le confinement qui apparaît lorsque les dimensions transverses sont de l'ordre du rayon de Bohr du matériau (4.9 nm pour le silicium) ce qui est le cas dans les transistors à nanofil. Nous allons donc essayer de proposer une architecture tirant partie du confinement quantique.

Pour réduire le courant tunnel nous pouvons envisager d'élargir la barrière de potentiel dans le canal. Pour cela nous proposons d'augmenter localement le confinement en rétrécissant le canal (Fig. 3.1) [15]. Ceci aura comme effet d'augmenter localement le minimum des sous-bandes et donc de mimer une augmentation de

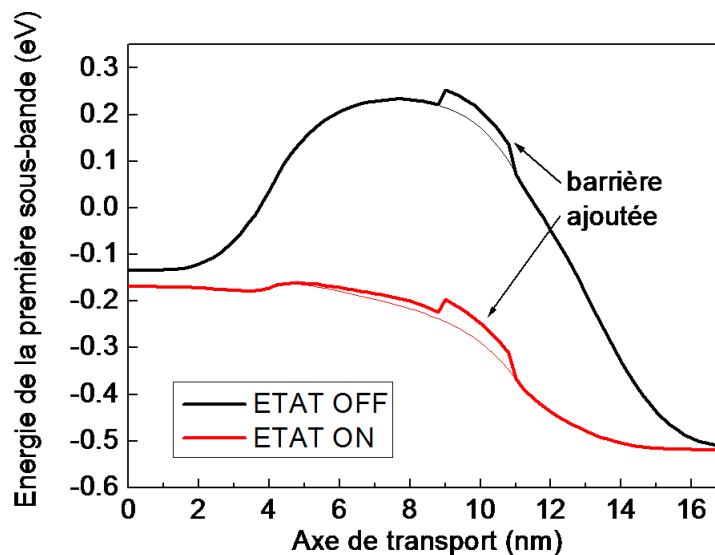


FIG. 3.2: Représentation schématique de la première sous-bande le long d'un nanofil avec un élargissement de barrière de potentiel en régime *ON* ($V_G = 0.6$ V) et en régime *OFF* ($V_G = 0$ V).

l'énergie potentielle. Ce rétrécissement ne doit pas être effectué à n'importe quel endroit du canal. Il doit se trouver côté drain afin d'augmenter la largeur de la barrière tout en évitant de la sur-élever (ce qui réduirait le courant en régime *ON*). Afin d'illustrer cet aspect, nous représentons schématiquement sur la figure 3.2, l'influence sur la première sous-bande d'un rétrécissement du canal (appelé *notch*). Dans ce chapitre et le chapitre 4 nous parlons de représentation schématique de la première sous-bande. Cette pseudo sous-bande est obtenue en calculant le minimum des énergies propres pour chaque plan le long de la direction du transport. Elle n'est pas exacte mais permet d'expliquer les phénomènes qui se produisent dans le nanofil. La présence du *notch* doit générer une augmentation de la largeur de barrière sans altérer le courant thermoïonique en *ON*. Pour cela, la barrière ajoutée ne doit pas être trop haute en énergie. Idéalement, elle doit être "cachée" derrière la barrière principale. En revanche, en régime *OFF* la présence du *notch* doit élargir au maximum la barrière afin de réduire le courant.

Dans ce chapitre nous nous proposons de simuler l'effet du *notch* sur les caractéristiques d'un transistor à nanofil (Figure 3.3). Nous verrons en particulier l'impact de la géométrie du *notch*.

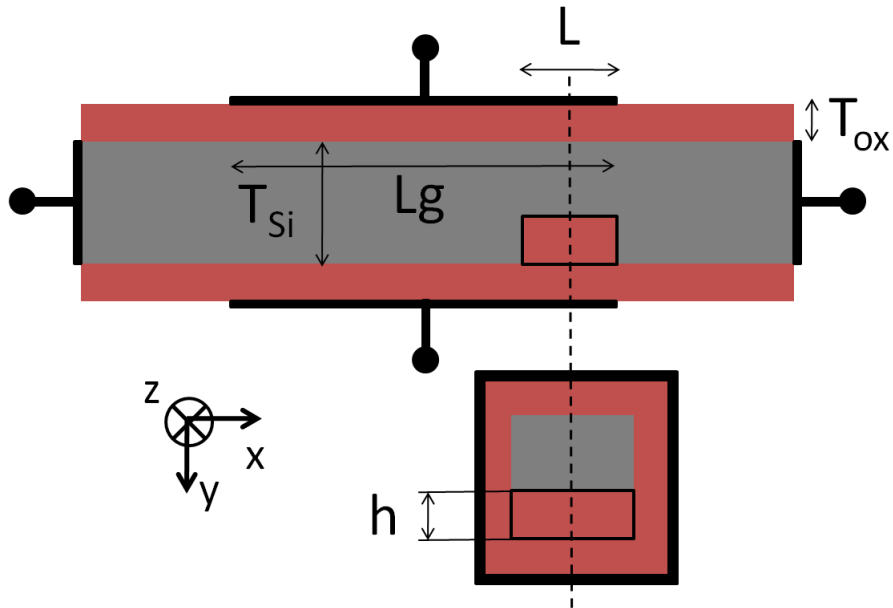


FIG. 3.3: Représentation schématique d'un transistor à nanofil GAA avec un *notch* de hauteur h et de longueur L . L'oxyde qui entoure le fil a une épaisseur $T_{ox} = 1$ nm et le fil a une épaisseur $T_{Si} = 2$ nm. La source, le drain et la grille ont des longueurs respectives de 4, 6 et $L_g = 7$ nm.

3.2 Modèle utilisé

Nous allons considérer un transistor à nanofil de silicium de type n orienté dans la direction $[100]$. Nous nous plaçons dans l'approximation de la masse effective (EMA pour Effective Mass Approximation). Selon le matériau utilisé et l'orientation du nanofil, l'EMA permet une description plus ou moins bonne de la structure de bandes. En l'occurrence, il a été démontré que le bas des bandes de conduction d'un nanofil de silicium orienté suivant $[100]$ reste parabolique jusqu'à un diamètre de 1.5 nm [49–51]. Étant donné que nous ne descendons pas en dessous de 2 nm de diamètre, l'EMA est donc valable. L'équation de Schrödinger s'écrit donc selon l'EMA avec les masses effectives longitudinales m_l et transverses m_t extraites, pour différentes sections de fil, à partir d'un calcul de structure de bandes en liaisons fortes [52].

Notre modèle est basé sur le formalisme des fonctions de Green hors équilibre couplé de façon auto-cohérente à l'équation de Poisson 3D. L'approximation de

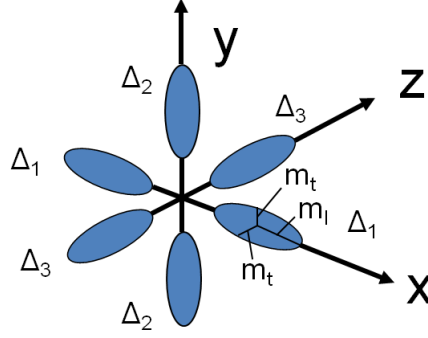


FIG. 3.4: Représentation des six ellipsoïdes équivalents du silicium massif. En cas de confinement, les vallées Δ_1 , Δ_2 et Δ_3 prennent des énergies différentes.

l'espace des modes [53] est très intéressante pour gagner en temps de calcul car toutes les sous-bandes sont considérées comme étant indépendantes les unes des autres. Le problème 3D se transforme en N (nombre de sous-bandes présent en compte) problèmes 1D. Malheureusement cette approximation n'est valable que pour des nanofils dont le potentiel varie lentement à l'échelle nanométrique (ce qui n'est pas du tout le cas pour un nanofil avec *notch*). Nous devons donc effectuer une résolution 3D dans l'espace réel de l'équation de Schrödinger. L'approche de l'espace des modes est malgré tout utilisée pour initialiser le potentiel électrostatique. Toutefois afin de réduire le temps de notre calcul 3D, nous utilisons l'algorithme de récursion décrit au chapitre 2.4.

Nous avons pour chaque vallée Δ du silicium (Fig. 3.4) un hamiltonien 3D que nous devons discrétiser afin de le traiter numériquement. Nous le discrétisons en utilisant la méthode des différences finies (avec un pas de 0.2 nm) dans les directions [100], [010] et [001], respectivement, en N_x , N_y et N_z points. Nous obtenons une matrice de dimension $N_x \times N_y \times N_z$ de la forme

$$H = \begin{bmatrix} [\alpha_1] & [\beta] & 0 & \cdots & 0 \\ [\beta^+] & [\alpha_2] & [\beta] & \ddots & \vdots \\ 0 & [\beta^+] & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & [\alpha_{N_x-1}] & [\beta] \\ 0 & \cdots & 0 & [\beta^+] & [\alpha_{N_x}] \end{bmatrix} \quad (3.1)$$

avec $[\alpha_i]$ la matrice du i ème plan transverse et $[\beta]$ les matrices de couplage entre deux plans successifs. Ces matrices sont de dimension $N_y \times N_z$.

Dans le cas étudié, c'est-à-dire un nanofil fortement confiné et de section transverse carrée, les vallées $\Delta 2$ et $\Delta 3$ sont dégénérées. De plus les sous-bandes issues de la vallée $\Delta 1$ ont une influence négligeable sur le courant car elles sont à très haute énergie [54]. Ainsi, nous ne représenterons que les coefficients de transmission et les spectres de courant issus des vallées $\Delta 2$ et $\Delta 3$.

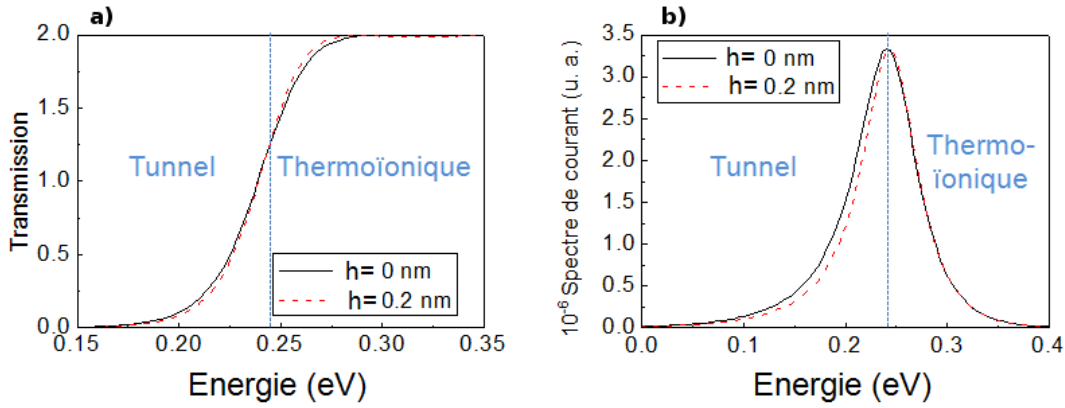


FIG. 3.5: a) Coefficient de transmission des vallées $\Delta 2$ et $\Delta 3$, et b) Spectre de courant en régime *OFF* ($V_G = 0$ V) pour un transistor sans ($h = 0$ nm) et avec un *notch* ($h = 0.2$ nm et $L = 1$ nm). $V_{DS} = 0.4$ V.

3.3 Résultats de simulation

Dans cette partie, nous analysons l'impact du *notch* sur les régimes *OFF* et *ON* avec différents paramètres comme la hauteur h ou la longueur L du *notch*. Le dispositif étudié est un transistor *Gate-All-Around* (GAA) de section transverse 2×2 nm² (hormis au niveau du *notch*). La source et le drain ont des longueurs respectives de 4 et 6 nm. La longueur de grille L_g est de 7 nm et l'oxyde qui entoure le fil a une épaisseur $T_{ox} = 1$ nm. La tension de drain V_{DS} est constante et égale à 0.4 V. La tension de grille V_G , quant à elle, varie de 0 V (régime *OFF*) à 0.6 V (régime *ON*). Le dopage de la source et du drain est de 10^{20} cm⁻³. Nous définissons aussi deux paramètres R et I . R est le rapport I_{ON}/I_{OFF} comparé au rapport I_{ON}/I_{OFF} d'un

MOSFET conventionnel (sans *notch*).

$$R = \frac{(I_{ON}/I_{OFF})}{(I_{ON}/I_{OFF})_{conventionnel}} \quad (3.2)$$

I est le rapport entre le courant ON et le courant ON d'un MOSFET conventionnel.

$$I = \frac{I_{ON}}{(I_{ON})_{conventionnel}} \quad (3.3)$$

3.3.1 Influence du *notch* en régime OFF

Nous étudions dans un premier temps l'influence du notch en régime OFF en comparant le coefficient de transmission (Fig. 3.5 a) d'un MOSFET conventionnel à celui d'un MOSFET avec un *notch*. Comme attendu la présence du *notch* réduit la transmission tunnel. La conséquence sur le spectre de courant est clairement visible sur la figure 3.5 b). On note un courant OFF réduit par le *notch* ($I_{OFF} = 5.64 \times 10^{-11}$ A contre $I_{OFF} = 6.15 \times 10^{-11}$ A). On voit donc un effet bénéfique de la présence du *notch* dès $h = 0.2$ nm.

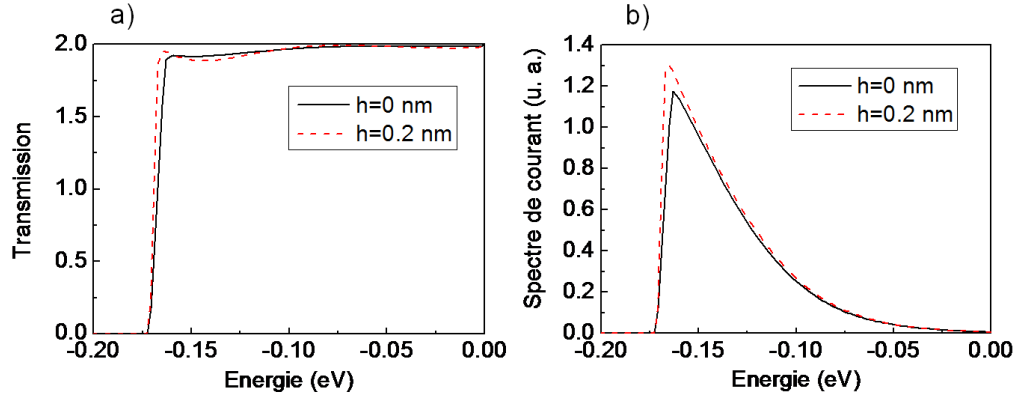


FIG. 3.6: a) Coefficient de transmission des vallées $\Delta 2$ et $\Delta 3$, et b) Spectre de courant en régime ON ($V_G = 0.6$ V) pour un transistor sans ($h = 0$ nm) et avec un *notch* ($h = 0.2$ nm et $L = 1$ nm). $V_{DS} = 0.4$ V.

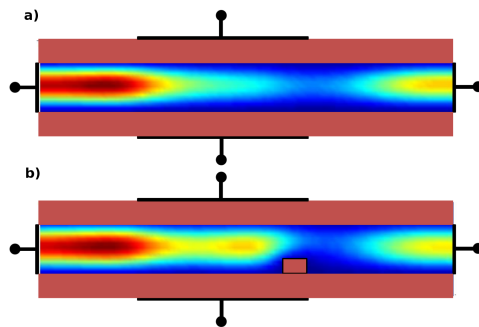


FIG. 3.7: Densité d'électrons dans le canal de conduction pour : a) un transistor conventionnel et b) un transistor avec *notch* de hauteur $h = 0.4$ nm et de longueur $L = 1$ nm. $V_G = 0.6$ V et $V_{DS} = 0.4$ V.

3.3.2 Influence du *notch* en régime *ON*

Dans un MOSFET à nanofil GAA conventionnel, le courant I_{ON} est principalement thermoïonique. Le but premier du *notch* étant de réduire le courant tunnel, il ne devrait donc pas avoir autant d'impact en régime *ON* qu'en régime *OFF*. Nous considérons dans cette partie le même dispositif que précédemment avec une hauteur $h = 0.2$ nm. La figure 3.6 a) qui compare les coefficients de transmission de transistors avec et sans *notch* nous montre une augmentation de la transmission en présence du *notch*. Afin de comprendre ce résultat surprenant nous pouvons observer la densité d'électrons correspondante (Fig. 3.7). En présence de *notch*, cette dernière affiche un maximum dans la région de canal. Cet effet semble traduire l'existence d'états localisés qui induisent des phénomènes de quasi-résonances sur la transmission. Au final la figure 3.6 b) montre que le courant *ON* augmente avec la présence du *notch* ($I_{ON} = 8.13 \times 10^{-5}$ A contre $I_{ON} = 7.86 \times 10^{-5}$ A). Cette augmentation n'était pas attendue mais s'avère être très positive.

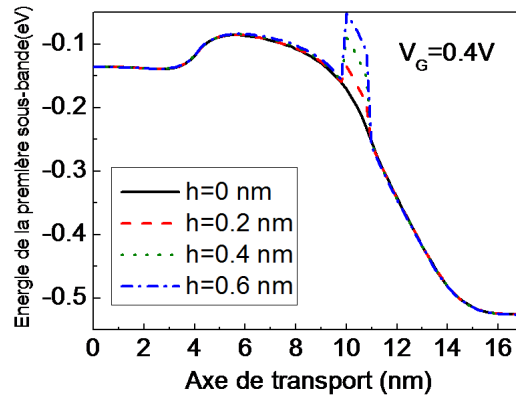


FIG. 3.8: Représentation schématique de la première sous-bande le long du nanofil avec un *notch* de hauteur h variable.

3.3.3 Influence de la hauteur du *notch*

La figure 3.8 donne une représentation schématique de la première sous-bande le long du canal pour différentes hauteurs de *notch* (de 0 à 0.6 nm). La longueur reste égale à 1 nm. Nous avons calculé le courant I_D pour ces différentes hauteurs de notch et pour des tensions de grille V_G allant de 0 à 0.6 V. Nous avons obtenu les quatre caractéristiques $I_D(V_G)$ de la figure 3.9. Nous observons que sous le seuil, le courant diminue lorsque la hauteur h augmente.

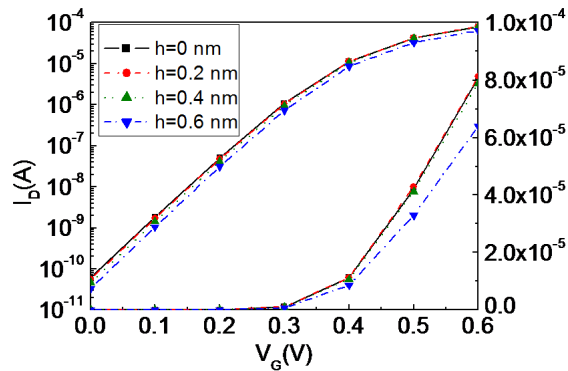


FIG. 3.9: Caractéristiques $I_D(V_G)$ d'un transistor à nanofil avec une hauteur h de *notch* variant de 0 à 0.6 nm. $V_{DS} = 0.4$ V.

Le régime ON est un peu plus compliqué. En effet, pour $h = 0.2$ et $h = 0.4$ nm I_{ON} est supérieur à celui d'un transistor traditionnel (pour les raisons évoquées précédemment). Il diminue à partir de $h = 0.6$ nm comme nous pouvons le voir sur la figure 3.10 (I devient inférieur à 100%). Cette dimension de *notch* génère une trop forte augmentation de la hauteur de barrière de potentiel et dégrade les propriétés du régime ON . La figure 3.10 nous montre aussi que le rapport I_{ON}/I_{OFF} augmente avec la hauteur du *notch* (R augmente avec h). Un bon compromis est obtenu pour $h = 0.4$ nm car le rapport I_{ON}/I_{OFF} augmente de 32% ($R = 132\%$) sans dégradé le courant ON ($I = 100.4\%$).

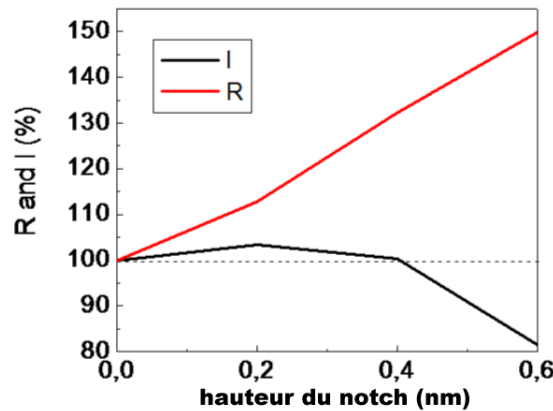


FIG. 3.10: Evolution des paramètres I et R en fonction de la hauteur du *notch*. $V_{DS} = 0.4$ V.

3.3.4 Influence de la longueur du *notch*

Dans le chapitre 3.3.3 nous avons observé l'effet de la hauteur du *notch* sur le courant en gardant la même longueur $L = 1$ nm. Nous allons maintenant étudier l'influence de la longueur du *notch* en gardant la même hauteur $h = 0.2$ nm. La figure 3.11 nous montre une représentation schématique de la première sous-bande le long du canal pour différentes longueurs de *notch* en régime ON . Le *notch* modifie très fortement le potentiel localement (au niveau du *notch*) et a un léger impact sur le potentiel dans le reste du canal. En revanche, à partir d'une certaine longueur, le *notch* peut avoir un fort impact sur le potentiel dans le canal. Sur la figure 3.11 nous pouvons observer une élévation de la barrière lorsque $L \geq 1.6$ nm. Cette

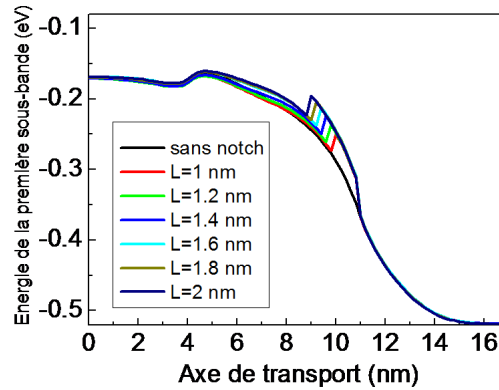


FIG. 3.11: Evolution de la première sous-bande le long du nanofil avec un *notch* de longueur L variable. $V_G = 0.6$ V et $V_{DS} = 0.4$ V.

augmentation locale de l'énergie potentielle est due à l'accumulation de charges dans le "puits" formé par le *notch*. Ceci entraîne une chute du courant ON comme illustré figure 3.12 où I est représenté en fonction de L . Cette figure nous montre aussi que le rapport I_{ON}/I_{OFF} augmente avec la longueur du *notch* (R augmente avec L). Nous avons vu dans le chapitre 3.3.3 que R augmente également avec h . Donc plus les dimensions du *notch* sont importantes, plus le rapport I_{ON}/I_{OFF} sera grand. Mais au delà de certaines dimensions ($h \geq 0.6$ nm et $L \geq 1.6$ nm) l'augmentation de ce rapport s'accompagne d'une forte chute du courant ON , ce qui n'est pas le but de cette architecture.

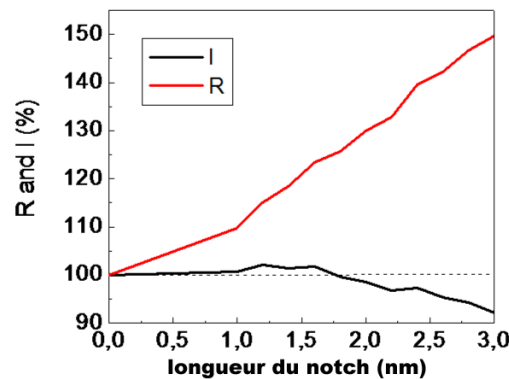


FIG. 3.12: Evolution des paramètres I et R en fonction de la longueur du *notch*. $V_{DS} = 0.4$ V.

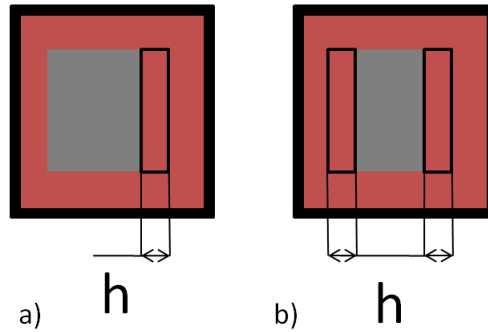


FIG. 3.13: Représentation schématique de sections transverses de transistors à *notch* avec deux configurations différentes. a) configuration “*notch simple*” et b) configuration “*notch double*”.

3.3.5 Influence du nombre de *notches*

Nous allons maintenant évaluer l’impact du nombre de *notches*. Plus précisément, nous souhaitons comparer l’effet d’un “*notch double*”, c’est-à-dire un *notch* situé de chaque côté du canal (Fig. 3.13 b)), à l’effet d’un “*notch simple*” (Fig. 3.13 a)) mais deux fois plus haut. La figure 3.14 a) montre les coefficients de transmission en régime *ON* et *OFF* pour les configurations suivantes : un “*notch simple*” de hauteur $h = 0.4$ nm et un “*notch double*” ayant chacun une hauteur $h = 0.2$ nm. Nous observons que la transmission est légèrement supérieure pour le “*notch simple*” en régime *ON* et que la transmission est plus faible pour cette même configuration en régime *OFF*. La configuration “*notch simple*” est donc avantageuse dans les deux régimes de fonctionnement. Ceci est confirmé par les spectres de courant de la figure 3.14 b) où nous pouvons remarquer que le “*notch simple*” offre un courant *ON* très légèrement supérieur et un courant *OFF* bien inférieur au “*notch double*” ($I_{ON}/I_{OFF} = 1.689 \times 10^6$ contre $I_{ON}/I_{OFF} = 1.55 \times 10^6$).

3.3.6 Conclusion

Nous avons montré que l’ajout d’un *notch* d’oxyde correctement positionné et de taille adéquate peut améliorer les caractéristiques de courant de MOSFETs à nanofil. Cette amélioration est due en partie à l’apparition d’états localisés quasi-

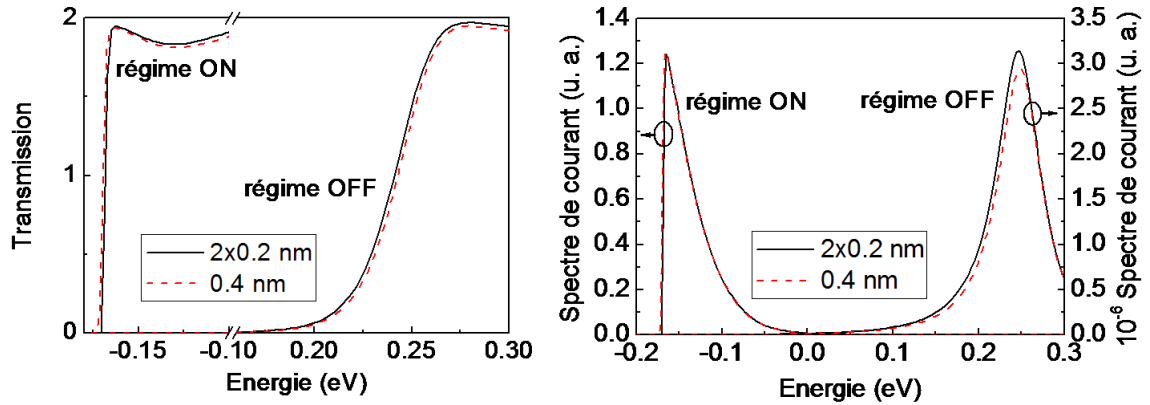


FIG. 3.14: a) Coefficient de transmission et b) spectre de courant en régime *ON* et *OFF* ($V_G = 0.6$ V et $V_G = 0$ V respectivement) pour un transistor avec un “notch simple” ($h = 0.4$ nm) et avec un “notch double” ($h = 0.2$ nm). $V_{DS} = 0.4$ V.

résonnants qui peuvent augmenter le courant *ON*, mais surtout à l’élargissement de la barrière de potentiel qui diminue le courant tunnel en régime *OFF*. Les meilleures performances ont été obtenues pour un “notch simple” de longueur $L = 1$ nm et de hauteur $h = 0.4$ nm. Cette configuration permet d’améliorer de 32% le rapport I_{ON}/I_{OFF} sans réduction du courant *ON*.

Chapitre 4

Transistor à nanofil de type p

Dans ce chapitre, nous nous intéressons à la simulation des transistors à nanofil de type p. Pour cela, nous avons développé un code de transport quantique basé sur le formalisme des fonctions de Green avec un hamiltonien $\mathbf{k} \cdot \mathbf{p}$ à six bandes. Ce code nous permet de considérer le caractère multibandes de la bande de valence et d'étudier l'influence d'une impureté dans le canal.

Dans un premier temps, nous faisons une brève introduction à la théorie $\mathbf{k} \cdot \mathbf{p}$ en commençant par le hamiltonien massif que nous discrétisons afin de modéliser un composant. Nous analyserons ensuite les résultats des simulations dans le cas d'un transistor à canal intrinsèque. Enfin, nous verrons comment simuler la présence d'une impureté ionisée dans le canal et quel en est l'impact de sur les propriétés de transport.

4.1 Modélisation de la bande de valence en méthode $\mathbf{k}\cdot\mathbf{p}$ 6 bandes

4.1.1 Principe de la théorie $\mathbf{k}\cdot\mathbf{p}$

Dans cette partie, nous allons faire une brève introduction à la théorie $\mathbf{k}\cdot\mathbf{p}$ dont nous nous servirons pour décrire la bande de valence dans les nanofils. Plaçons nous dans l'approximation des électrons indépendants. Nous souhaitons résoudre l'équation de Schrödinger pour un électron dans un cristal périodique de potentiel cristallin U_c . Nous avons donc :

$$H\phi = \left(\frac{\mathbf{p}^2}{2m_0} + U_c \right) \phi = E\phi \quad (4.1)$$

avec E l'énergie et ϕ la fonction d'onde de l'électron, et m_0 la masse de l'électron libre. Le potentiel étant périodique, nous savons que les fonctions d'onde sont des fonctions de Bloch l'expression est :

$$\phi_{n\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\mathbf{r}} u_{n\mathbf{k}}(\mathbf{r}) \quad (4.2)$$

avec $u_{n\mathbf{k}}$ une fonction périodique de même période que U_c . \mathbf{k} est le vecteur d'onde, \mathbf{r} la position et n l'indice de bande.

Si nous appliquons H sur les fonctions de Bloch, alors nous avons :

$$H\phi_{n\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\mathbf{r}} \left(\frac{\mathbf{p}^2}{2m_0} + U_c + \frac{\hbar^2\mathbf{k}^2}{2m_0} + \frac{\hbar}{m_0}\mathbf{k}\cdot\mathbf{p} \right) u_{n\mathbf{k}}(\mathbf{r}) = E_{n\mathbf{k}}u_{n\mathbf{k}}(\mathbf{r}) e^{i\mathbf{k}\mathbf{r}}. \quad (4.3)$$

En simplifiant ensuite par $e^{i\mathbf{k}\mathbf{r}}$ et on obtient :

$$H_{\mathbf{k}\cdot\mathbf{p}}u_{n\mathbf{k}}(\mathbf{r}) = \left(H + \frac{\hbar^2\mathbf{k}^2}{2m_0} + \frac{\hbar}{m_0}\mathbf{k}\cdot\mathbf{p} \right) u_{n\mathbf{k}}(\mathbf{r}) = E_{n\mathbf{k}}u_{n\mathbf{k}}(\mathbf{r}). \quad (4.4)$$

Dans un cristal périodique, la recherche des états électroniques revient donc à résoudre l'équation aux valeurs propres :

$$H_{\mathbf{k}\cdot\mathbf{p}}u_{n\mathbf{k}}(\mathbf{r}) = E_{n\mathbf{k}}u_{n\mathbf{k}}(\mathbf{r}). \quad (4.5)$$

$H_{\mathbf{k}\cdot\mathbf{p}}$ est la somme de trois termes : le hamiltonien H qui définit $H_{\mathbf{k}\cdot\mathbf{p}}$ à $\mathbf{k} = \mathbf{0}$,

un terme d'énergie cinétique $\frac{\hbar^2 \mathbf{k}^2}{2m_0}$ et un terme de couplage $\frac{\hbar}{m_0} \mathbf{k} \cdot \mathbf{p}$.

La méthode $\mathbf{k} \cdot \mathbf{p}$ consiste à résoudre cette dernière équation en supposant connues les fonctions propres $u_{n\mathbf{0}}$ et les valeurs propres $E_{n\mathbf{0}}$ associées à l'équation 4.5 à $\mathbf{k} = \mathbf{0}$:

$$Hu_{n\mathbf{0}} = E_{n\mathbf{0}}u_{n\mathbf{0}}. \quad (4.6)$$

$\{u_{n\mathbf{0}}\}$ est une base complète de l'espace engendré par les parties périodiques $u_{n\mathbf{k}}$ des fonctions de Bloch [55]. Afin de traiter $H_{\mathbf{k},\mathbf{p}}$ nous pouvons donc le projeter sur cette base :

$$\langle u_{n'\mathbf{0}} | H_{\mathbf{k},\mathbf{p}} | u_{n\mathbf{0}} \rangle = \left(E_{n\mathbf{0}} + \frac{\hbar^2 \mathbf{k}^2}{2m_0} \right) \delta_{nn'} + \langle u_{n'\mathbf{0}} | \frac{\hbar}{m_0} \mathbf{k} \cdot \mathbf{p} | u_{n\mathbf{0}} \rangle. \quad (4.7)$$

La base des $u_{n\mathbf{0}}$ étant infinie, il faut choisir une restriction de cette base. Dans le cas où cette restriction ne serait pas satisfaisante, on peut toujours inclure de façon perturbative l'influence des bandes en dehors de la restriction. Si maintenant nous prenons en compte le couplage avec les bandes éloignées en énergie (indicées ν), des termes supplémentaires vont alors apparaître dans la matrice du hamiltonien. Luttinger et Kohn ont montré que les perturbations d'ordre 2 provenant des états $u_{\nu\mathbf{0}}$ sont introduites dans la matrice par la transformation [55] :

$$H_{ij} \rightarrow H_{ij} + \sum_{\nu \neq i,j} \frac{1}{2} \left(\frac{1}{E_i - E_\nu} + \frac{1}{E_j - E_\nu} \right) H_{i\nu} H_{\nu j}. \quad (4.8)$$

En premier lieu, la restriction est choisie en fonction du système physique à traiter.

4.1.2 Etats de base en bande de valence

Etant donné que nous voulons étudier le comportement de transistors de type p, nous présenterons les cas d'une restriction à trois bandes, puis à six bandes en incluant le spin de l'électron et en prenant en compte l'interaction spin-orbite pour décrire les états de trous. En effet, la bande de valence des semiconducteurs regroupe trois bandes deux fois dégénérées et appelées bandes de trous lourds, trous légers et split-off. Dans le cas de l'étude du transistor de type n, un hamiltonien de masse effective est utilisé à cause de l'allure "parabolique" de la bande de conduction. Mais dans le cas du transistor de type p, le couplage entre ces trois bandes entraîne une

forte non-parabolicité et anisotropie de la structure de bandes : il faut un hamiltonien $\mathbf{k} \cdot \mathbf{p}$ à six bandes pour décrire correctement la bande de valence [55]. Dans ce cas, la bande de valence correspond à un seul n et celui-ci sera donc omis dans les notations. Nous fixons également $E_0 = 0$. Les six états de Bloch à $\mathbf{k} = \mathbf{0}$ s'écrivent communément $\{iX \uparrow, iY \uparrow, iZ \uparrow, iX \downarrow, iY \downarrow, iZ \downarrow\}$ [56] avec \uparrow et \downarrow qui représente le spin. Enfin, La théorie des groupes donne [57] :

$$\langle u'_0 | \mathbf{p} | u_0 \rangle = 0 \quad (4.9)$$

au sein de cette restriction.

4.1.3 Modèle à 3 bandes

Dans un premier temps on ignore le spin de l'électron. Dans ce cas la restriction de la base est de dimension 3. Le hamiltonien $H'_{\mathbf{k},\mathbf{p}}$ est donc une matrice 3×3 sur la base $\{iX, iY, iZ\}$ en $\mathbf{k} = (k_x, k_y, k_z)$:

$$[H'_{\mathbf{k},\mathbf{p}}] = \frac{\hbar^2}{2m_0} \begin{bmatrix} |iX\rangle & |iY\rangle & |iZ\rangle \\ k^2 & 0 & 0 \\ 0 & k^2 & 0 \\ 0 & 0 & k^2 \end{bmatrix}. \quad (4.10)$$

Les énergies propres du hamiltonien s'écrivent alors : $E_X(\mathbf{k}) = E_Y(\mathbf{k}) = E_Z(\mathbf{k}) = \frac{\hbar^2 k^2}{2m_0}$, de fonctions propres $e^{i\mathbf{k}\mathbf{r}} |iX\rangle$, $e^{i\mathbf{k}\mathbf{r}} |iY\rangle$ et $e^{i\mathbf{k}\mathbf{r}} |iZ\rangle$ en $\mathbf{k} = \mathbf{0}$. Dans cette méthode les $|u_0\rangle$ jouent le rôle de spineurs. Dans ce cas, les trois bandes de valence associées à ces fonctions propres sont dégénérées et ont une courbure positive. Or on sait qu'en réalité ce n'est pas le cas (en convention électron). La restriction choisie n'est donc pas satisfaisante et il faut tenir compte de l'influence des autres bandes en perturbation. Dans ce cas, la matrice du hamiltonien s'écrit alors [55] :

$$[H'_{\mathbf{k},\mathbf{p}}] = \frac{\hbar^2}{2m_0} \begin{bmatrix} |iX\rangle & |iY\rangle & |iZ\rangle \\ \{k^2 + Lk_x^2 + M(k_y^2 + k_z^2)\} & Nk_xk_y & Nk_xk_z \\ Nk_xk_y & \{k^2 + Lk_y^2 + M(k_x^2 + k_z^2)\} & Nk_yk_z \\ Nk_xk_z & Nk_yk_z & \{k^2 + Lk_z^2 + M(k_x^2 + k_y^2)\} \end{bmatrix}. \quad (4.11)$$

Avec les coefficients non nuls L , M et N tels que :

$$L = \frac{2}{m_0} \sum_{\nu} \frac{\langle X|p_x|\nu\rangle\langle\nu|p_x|X\rangle}{E_V - E_{\nu}}, \quad (4.12)$$

$$M = \frac{2}{m_0} \sum_{\nu} \frac{\langle X|p_y|\nu\rangle\langle\nu|p_y|X\rangle}{E_V - E_{\nu}}, \quad (4.13)$$

$$N = \frac{2}{m_0} \sum_{\nu} \frac{\langle X|p_x|\nu\rangle\langle\nu|p_y|Y\rangle + \langle X|p_y|\nu\rangle\langle\nu|p_x|Y\rangle}{E_V - E_{\nu}}. \quad (4.14)$$

Traisons le cas particulier $\mathbf{k} = (0, 0, k_z)$. La matrice du hamiltonien s'écrit alors :

$$[H'_{\mathbf{k},\mathbf{p}}] = \frac{\hbar^2}{2m_0} \begin{bmatrix} |iX\rangle & |iY\rangle & |iZ\rangle \\ k_z^2(1 + M) & 0 & 0 \\ 0 & k_z^2(1 + M) & 0 \\ 0 & 0 & k_z^2(1 + L) \end{bmatrix}. \quad (4.15)$$

Les valeurs propres associées aux fonctions propres $e^{i\mathbf{k}\mathbf{r}} |iX\rangle$, $e^{i\mathbf{k}\mathbf{r}} |iY\rangle$ et $e^{i\mathbf{k}\mathbf{r}} |iZ\rangle$ sont respectivement :

$$E_X(\mathbf{k}) = E_Y(\mathbf{k}) = \frac{\hbar^2 k_z^2}{2 \left(\frac{m_0}{1+M} \right)} \quad (4.16)$$

et

$$E_Z(\mathbf{k}) = \frac{\hbar^2 k_z^2}{2 \left(\frac{m_0}{1+L} \right)}. \quad (4.17)$$

Pour la plupart, les bandes éloignées en énergie que nous prenons en compte

sont issues de la bande de conduction, donc $E_V - E_\nu < 0$ et par conséquent L et M sont négatifs. De plus, les termes $\frac{2}{m_0} \langle X | p_x | \nu \rangle \langle \nu | p_x | X \rangle$ et $\frac{2}{m_0} \langle X | p_y | \nu \rangle \langle \nu | p_y | X \rangle$ entrant respectivement dans les expressions de L et M sont plus grand que la bande interdite, donc $|L|$ et $|M|$ sont supérieurs à 1. Il en résulte que les coefficients $1 + L$ et $1 + M$ sont négatifs. Les trois bandes associées aux fonctions X , Y et Z en $\mathbf{k} = \mathbf{0}$ ont donc désormais une courbure négative : le traitement perturbatif au deuxième ordre est donc satisfaisant. A noter que les paramètres L , M et N sont équivalents aux paramètres de Luttinger (γ_1 , γ_2 et γ_3) qui sont plus communément utilisés. Les deux jeux de paramètres sont liés par les relations suivantes :

$$L = -(\gamma_1 + 4\gamma_2 + 1), M = 2\gamma_2 - \gamma_1 - 1 \text{ et } N = -6\gamma_3. \quad (4.18)$$

4.1.4 Modèle à 6 bandes : prise en compte du couplage spin-orbite

En incluant le spin de l'électron, la restriction de la base prend la dimension 6. La matrice du hamiltonien $H_{\mathbf{k},\mathbf{p}}$ devient donc une matrice 6×6 diagonale par blocs 3×3 sur la base $\{|\uparrow\rangle, |\downarrow\rangle\} \otimes \{|iX\rangle, |iY\rangle, |iZ\rangle\}$:

$$[H_{\mathbf{k},\mathbf{p}}] = \begin{bmatrix} |\uparrow\rangle & |\downarrow\rangle \\ [H'_{\mathbf{k},\mathbf{p}}] & 0 \\ 0 & [H'_{\mathbf{k},\mathbf{p}}] \end{bmatrix} \quad (4.19)$$

avec $H'_{\mathbf{k},\mathbf{p}}$ définit précédemment (équation 4.11).

Le hamiltonien six bandes total H devient $H = H_{\mathbf{k},\mathbf{p}} + H_{so}$, avec H_{so} le hamiltonien de l'interaction spin-orbite dont la matrice s'écrit [58] :

$$[H_{so}] = \frac{\Delta}{3} \begin{bmatrix} |iX \uparrow\rangle & |iY \uparrow\rangle & |iZ \uparrow\rangle & |iX \downarrow\rangle & |iY \downarrow\rangle & |iZ \downarrow\rangle \\ 0 & -i & 0 & 0 & 0 & 1 \\ i & 0 & 0 & 0 & 0 & -i \\ 0 & 0 & 0 & -1 & i & 0 \\ 0 & 0 & -1 & 0 & i & 0 \\ 0 & 0 & -i & -i & 0 & 0 \\ 1 & i & 0 & 0 & 0 & 0 \end{bmatrix} \quad (4.20)$$

avec Δ l'énergie de décalage spin-orbite. La bande de valence est maintenant composée de six bandes (trois bandes dégénérées deux fois). En $\mathbf{k} = \mathbf{0}$ la bande la plus haute en énergie est dégénérée quatre fois et elle est séparée de l'autre bande dégénérée deux fois, que l'on appelle bande *split-off*, de l'énergie Δ . En $\mathbf{k} \neq \mathbf{0}$ la bande la plus haute en énergie est dégénérée deux fois et est appelée bande de trous lourds pour $\mathbf{k} = (0, 0, k_z)$, alors que la suivante (également dégénérée deux fois) est appelée bande de trous légers. Ces bandes sont représentées pour $\mathbf{k} = (0, 0, k_z)$ sur la figure 4.1.

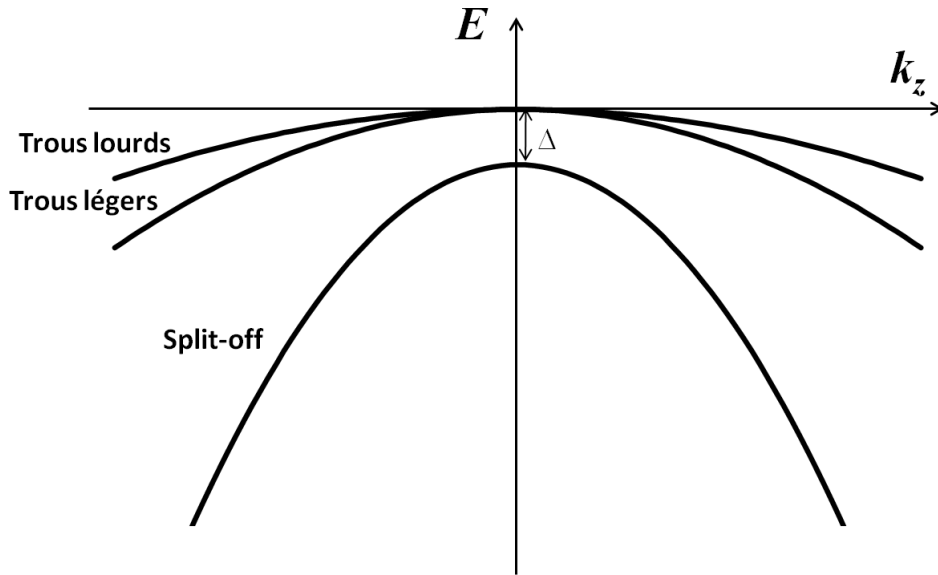


FIG. 4.1: Représentation schématique de la bande de valence.

4.1.5 Modèle à 6 bandes pour le nanocomposant

Maintenant que nous avons $H_{\mathbf{k},\mathbf{p}}$ le hamiltonien $\mathbf{k} \cdot \mathbf{p}$ six bandes du matériau massif, nous pouvons calculer le hamiltonien $\mathbf{k} \cdot \mathbf{p}$ à six bandes de nano-objets. Partant du modèle à 6 bandes que nous avons présenté et sachant que nous souhaitons un modèle 3D, nous devons remplacer les k_j par $-i\frac{\partial}{\partial j}$ pour les trois directions ($j = x, y, z$). Ce hamiltonien est ensuite discrétisé selon la méthode des différences finies et nous obtenons ainsi une matrice de dimension $6 \times N_x \times N_y \times N_z$ de la forme :

$$H = \begin{bmatrix} [\alpha_1] & [\beta] & 0 & \cdots & 0 \\ [\beta^+] & [\alpha_2] & [\beta] & \ddots & \vdots \\ 0 & [\beta^+] & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & [\alpha_{N_x-1}] & [\beta] \\ 0 & \cdots & 0 & [\beta^+] & [\alpha_{N_x}] \end{bmatrix} \quad (4.21)$$

avec $[\alpha]$ décrivant les plans perpendiculaires au transport et $[\beta]$ couplant deux plans successifs de dimension $6 \times N_y \times N_z$. Nous pouvons maintenant appliquer ce hamiltonien à l'étude des transistors à nanofil de type p.

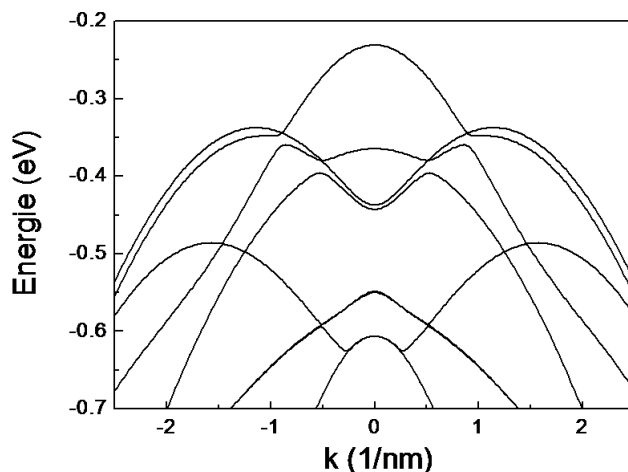


FIG. 4.2: Structure de bandes de valence d'un nanofil de silicium orienté [100] de section transverse $2 \times 2 \text{ nm}^2$.

4.2 Résultats de simulation

Nous avons simulé le fonctionnement d'un transistor à nanofil de silicium *Gate-All-Around* de type *p* dont la section transverse est $2 \times 2 \text{ nm}^2$. Le nanofil est confiné selon les directions [010] et [001]. Le transport s'effectue selon [100]. La longueur de la source, du drain et du canal de conduction est de 8 nm. La tension de drain a été fixée à -0.4 V et seule la tension de grille V_G varie. Le dopage de la source et du drain est de 10^{20} cm^{-3} . Étant donné qu'aucune méthode semi-empirique n'est satisfaisante pour décrire la structure de bandes des trous dans le SiO_2 amorphe, la pénétration des fonctions d'onde dans l'oxyde n'a pas été prise en compte. L'oxyde est seulement décrit dans l'équation de Poisson et son épaisseur est de 1 nm. Les paramètres de Luttinger (γ_1 , γ_2 et γ_3) du silicium connus dans la littérature sont ceux du silicium massif. Ils sont généralement utilisés pour décrire la bande de valence de structures peu confinées. Dans le cas de nanofils dont l'épaisseur ne dépasse pas quelques nanomètres, ces paramètres ne sont plus valables [20]. Dans le cas que nous avons étudié, c'est-à-dire pour un nanofil de section transverse $2 \times 2 \text{ nm}^2$, les paramètres de Luttinger ont été ajustés sur la bande de valence obtenue à l'aide d'un calcul en liaisons fortes [21]. Les paramètres obtenus sont : $\gamma_1 = 2.97$, $\gamma_2 = 0.65$ et $\gamma_3 = 1.045$. Les autres termes de la paramétrisation sont le couplage spin-orbite $\Delta = 0.044 \text{ eV}$, l'influence de la bande de conduction $E_p = 20 \text{ eV}$ et la bande

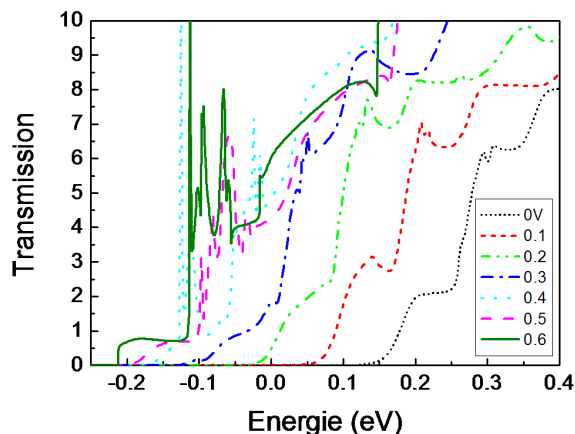


FIG. 4.3: Coefficient de transmission pour $-V_G$ variant de 0 à 0.6 V par pas de 0.1 V pour un transistor à nanofil de type p sans impureté. Il apparaît clairement que l’allure de la transmission devient de plus en plus “piquée” avec l’augmentation de la tension de grille. $V_{DS} = -0.4$ V.

interdite $E_G = 1.12$ eV. La structure de bandes du fil obtenue avec ces paramètres est représentée sur la figure 4.2.

4.2.1 Transistor de type p à canal intrinsèque

Nous avons considéré dans un premier temps un transistor à canal intrinsèque. La figure 4.3 représente le coefficient de transmission pour des tensions de grille allant de 0 à 0.6 V. Pour faciliter la compréhension les sous-bandes, les coefficients de transmission et les densités d’état seront représentées en terme d’énergie de trous. Pour de faibles valeurs de $|V_G|$, la transmission forme des marches du fait de la discrétisation des modes transverses de conduction dans les nanofils [60]. Les marches sont de moins en moins nettes et le nombre de pics augmente avec $-V_G$. L’apparition de ces pics résulte du couplage multibandes dû à la variation rapide du potentiel sur de courtes distances. La figure 4.4 représente la densité d’états locale (LDOS) sur laquelle nous avons superposé deux pseudo sous-bandes et la transmission afin de mieux comprendre l’apparition de ces pics. La structure de bandes de chaque extrémité du dispositif a également été représentée. Nous voyons sur la structure de bandes côté source pour une énergie d’environ -0.1 eV la présence de croisements ou anti-croisements d’où résulte une LDOS importante (en rouge). En longeant

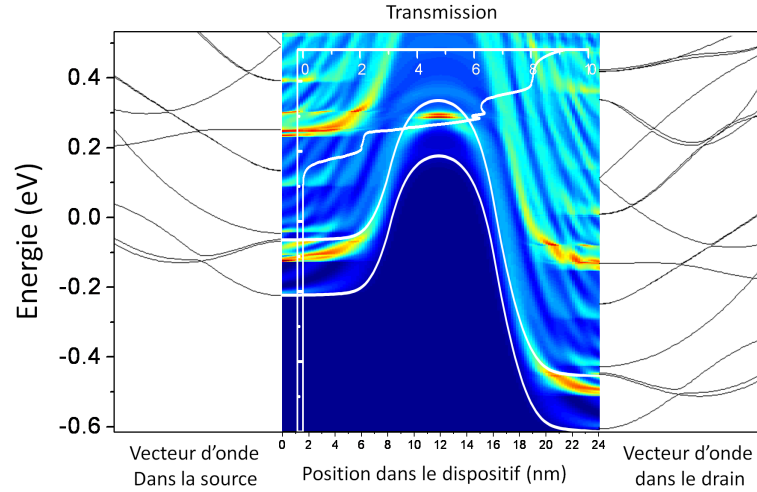


FIG. 4.4: Densité d'états locale, pseudo sous-bandes et transmission dans un MOSFET à nanofil de type p intrinsèque. La structure de bandes de chaque extrémité est aussi représentée. $V_{DS} = -0.4$ V et $V_G = 0$ V.

les pseudo sous-bandes, cette forte densité se retrouve au dessus de la barrière de potentiel et correspond aux pics dans la transmission (à une énergie d'environ 0.3 eV). A fort $|V_G|$, l'alignement en énergie des bandes dans le canal avec celles dans la source entraîne des interactions inter-bandes. Le couplage entre sous-bandes dans les transistors à nanofils de type p est un phénomène qui ne pouvait être observé dans les dispositifs de type n avec un modèle simple bande. Ce phénomène devrait amplifier les perturbations engendrées par la rugosité de surface, les impuretés ionisées ou les interactions avec les phonons qui apparaissent dans les transistors de type n. A noter que d'un point de vue numérique, la présence de ces pics perturbe l'auto-cohérence Schrödinger-Poisson. Il faut donc affiner le maillage en énergie afin d'aboutir à un potentiel stable.

4.2.2 Transistor de type p avec impureté ionisée

Traitement des impuretés ionisées

Dans cette partie, nous allons nous intéresser à la présence d'une impureté ionisée dans le canal de conduction d'un MOSFET à nanofil de type p [61]. Pour simuler la présence d'une impureté ionisée dans le canal, il faut prendre en compte l'énergie

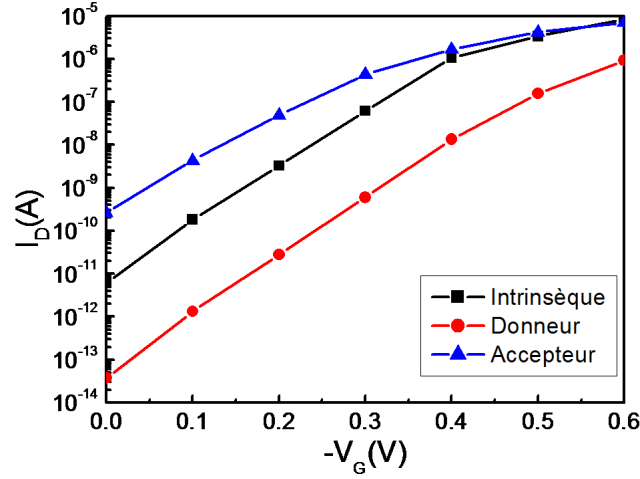


FIG. 4.5: Caractéristiques $I_D - V_G$ d'un composant avec trois configurations de dopage : sans impureté (carrés), impureté donneur (ronds) et impureté accepteur (triangles).

potentielle d'interaction coulombienne à longue distance combinée à un potentiel sur le site de l'impureté qui définit sa structure chimique [62]. Nous négligeons dans le présent travail la nature chimique de l'impureté et nous ne considérons que le potentiel d'interaction à longue distance. Les impuretés ionisées sont donc modélisées en ajoutant une charge (positive pour un donneur ou négative pour un accepteur) sur le site de l'impureté dans la résolution de l'équation de Poisson. En effet, le traitement du potentiel chimique de chaque type d'impureté nécessiterait une comparaison détaillée avec des données expérimentales et des calculs *ab initio*. Une telle étude n'a pour l'instant pas été publiée et est au-delà des objectifs de cette thèse.

Les variations du potentiel ne sont pas suffisamment lentes pour utiliser l'approche de l'espace des modes. Il faut, comme dans le chapitre précédent (*notch*) utiliser une approche tridimensionnelle dans l'espace réel où tous les modes de conduction sont couplés. Le calcul du courant est donc numériquement très lourd. En effet, la matrice du hamiltonien discrétisé H décrite dans l'expression 4.21 est maintenant de dimension $6 \times N_x \times N_y \times N_z$ (N_x , N_y et N_z étant les nombres de points respectivement selon x , y et z). Nous utilisons une nouvelle fois l'algorithme de récursion décrit au chapitre 2.4 afin de réduire le temps de calcul.

Résultats avec impureté ionisée

La figure 4.5 représente les caractéristiques de courant $I_D - V_G$ d'un transistor à nanofil orienté [100] avec trois configurations de dopage différents : sans impureté, avec une impureté de type donneur et de type accepteur placée au milieu du canal de conduction, au centre de la section transverse.

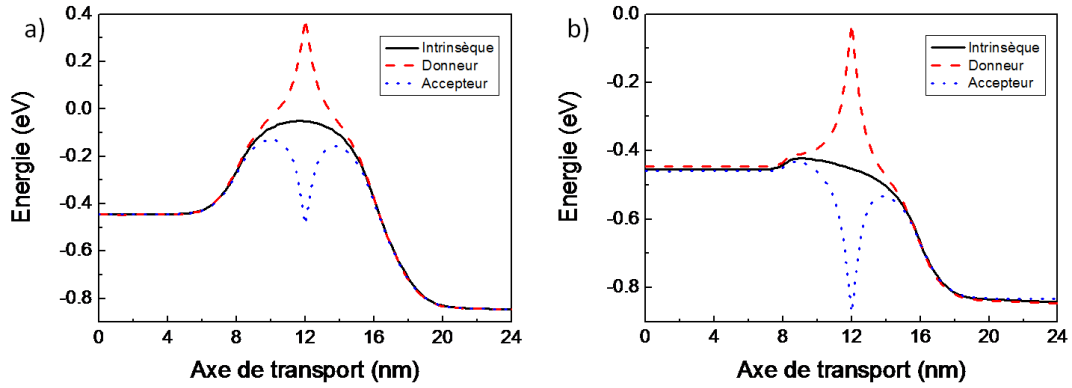


FIG. 4.6: Potentiel du nanofil sans impureté (ligne continue), avec une impureté de type donneur (tirets) et une impureté de type accepteur (pointillés) a) en régime OFF et b) en régime ON.

Tel un accepteur dans un dispositif de type n [59], un donneur dans un dispositif de type p induit un potentiel répulsif (figure 4.6) pour les trous environnants. L'augmentation de la hauteur de barrière de potentiel engendrée par le donneur dégrade le courant de plus de deux décades en régime *OFF* et d'une décade en régime *ON* (Fig. 4.5).

Cette dégradation s'interprète assez bien à l'aide de la figure 4.7 qui représente le coefficient de transmission en régime *OFF*. L'allure de la courbe avec impureté (pointillés) nous indique une dégradation de la transmission par rapport à la courbe sans impureté (ligne continue) due à l'augmentation de la hauteur de barrière. La proportion de courant tunnel est donc plus importante avec impureté. Ceci explique la forte diminution du courant *OFF*. En régime *ON*, l'augmentation de la hauteur de barrière due à l'impureté dégrade aussi le courant, mais le courant des trous provoque l'écrantage du potentiel coulombien qui conduit à un courant moins diminué par l'impureté en *ON* qu'en *OFF*.

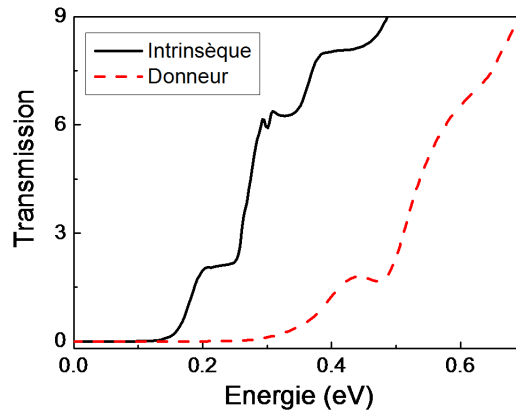


FIG. 4.7: Coefficient de transmission à $V_G = 0$ V pour un transistor à nanofil de type p sans impureté (ligne continue) et avec un donneur (pointillés) au milieu du canal de conduction.

La figure 4.8 représente la LDOS sur laquelle nous avons superposé des pseudo sous-bandes et la transmission, mais cette fois en présence d'une impureté ionisée de chaque type dans le canal. De façon générale, la LDOS montre des franges d'interaction dues aux réflexions des trous, injectés depuis les deux réservoirs, sur la barrière de potentiel. Nous voyons que ces franges sont plus marquées en présence d'un donneur (Fig. 4.8.a) que d'un accepteur (Fig. 4.8.b). Ceci est dû à l'augmentation de la hauteur de barrière engendrée par le donneur. Ce dernier réduit le courant de drain, mais son impact dépend fortement de sa position. Un donneur a plus d'impact s'il est positionné au centre de la section transverse que s'il est sur le bord de celle-ci. En positionnant l'impureté ionisée en bord de section transverse, nous obtenons un courant environ trois fois plus important que si elle est centrée (Fig. 4.9). Enfin, l'influence d'une impureté située près du bord de la section transverse a plus d'impact sur le courant dans un dispositif de type p que dans un dispositif de type n de même dimension [63]. La réduction plus importante du courant pour un dispositif de type p est dû au couplage inter-bandes plus important.

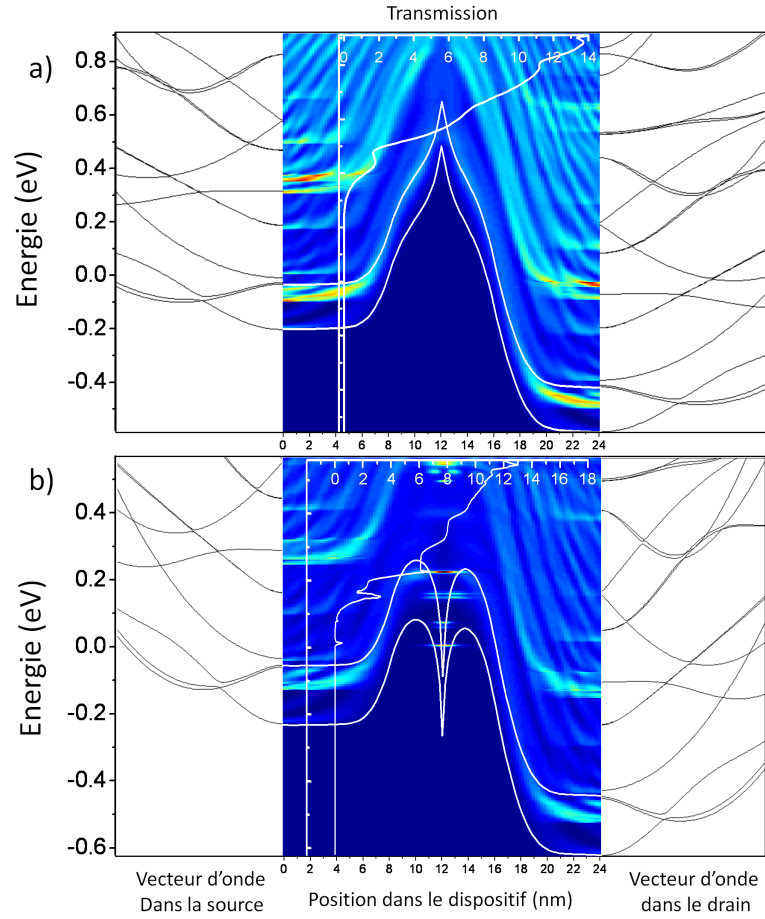


FIG. 4.8: Densité d'états locale, pseudo sous-bandes et transmissions dans un MOS-FET à nanofil de type p avec a) un donneur et b) un accepteur au milieu du canal de conduction. La structure de bandes de chaque extrémité est aussi représentée. $V_{DS} = -0.4$ V et $V_G = 0$ V.

La présence d'un accepteur dans un dispositif de type p crée un potentiel attractif pour les trous environnants. Ce puits de potentiel entraîne l'apparition d'états quasi-localisés où des interférences quantiques peuvent se produire. Nous pouvons voir sur la figure 4.8.b que les pics et les creux du coefficient de transmission correspondent aux énergies des états quasi-localisés qui apparaissent sur la LDOS. Pour mettre en avant ce comportement, nous avons tracé sur la figure 4.10.a les coefficients de transmission d'un dispositif sans impureté et avec un défaut de type accepteur au milieu de la section transversale. Nous voyons que les courbes de transmission sont décalées en énergie (ΔE). En effet, le potentiel attractif créé par l'accepteur a réduit

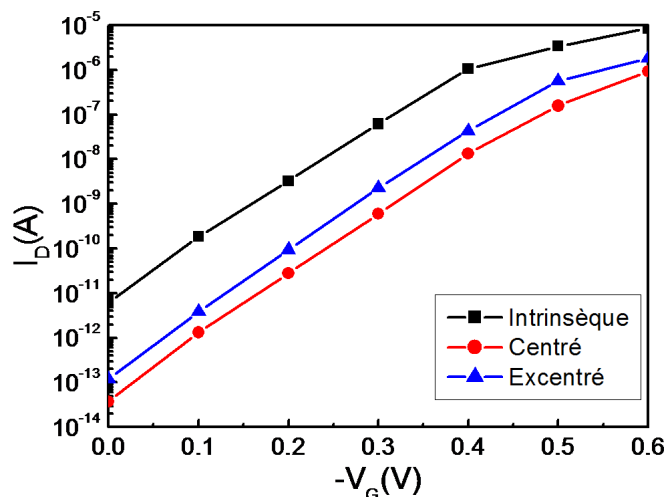


FIG. 4.9: Caractéristiques $I_D - V_G$ d'un composant avec trois configurations de dopage : sans impureté (carrés), avec impureté de type donneur au centre de la section transverse (ronds) et avec impureté de type donneur au bord de la section transverse (triangles).

la hauteur de barrière de potentiel de ΔE (Fig. 4.10.b). L'apparition des creux (effet anti-résonnant) et des pics (effet résonnant) est expliqué schématiquement sur la figure 4.10.c. Les pics sont le résultat d'un effet tunnel résonnant entre un trou injecté depuis la source dans une bande et un état quasi-localisé engendré par cette même bande. Les creux sont le résultat d'interactions anti-résonnantes entre un trou injecté depuis la source dans une bande et un état quasi-localisé engendré par une autre bande. Ce phénomène appelé résonance de Fano apparait également dans les dispositifs de type n mais seulement pour de fortes tensions de grille [59]. Pour un transistor de type p, cet effet est visible dès $V_G = 0$ V à cause du fort couplage inter-bandes. Comme dans les dispositifs de type n, la présence d'un accepteur entraîne une augmentation du courant *OFF* (Fig. 4.5). En régime *ON* il devient difficile de distinguer les interférences dues à l'accepteur et celles dues au couplage inter-bandes également présent dans un canal intrinsèque.

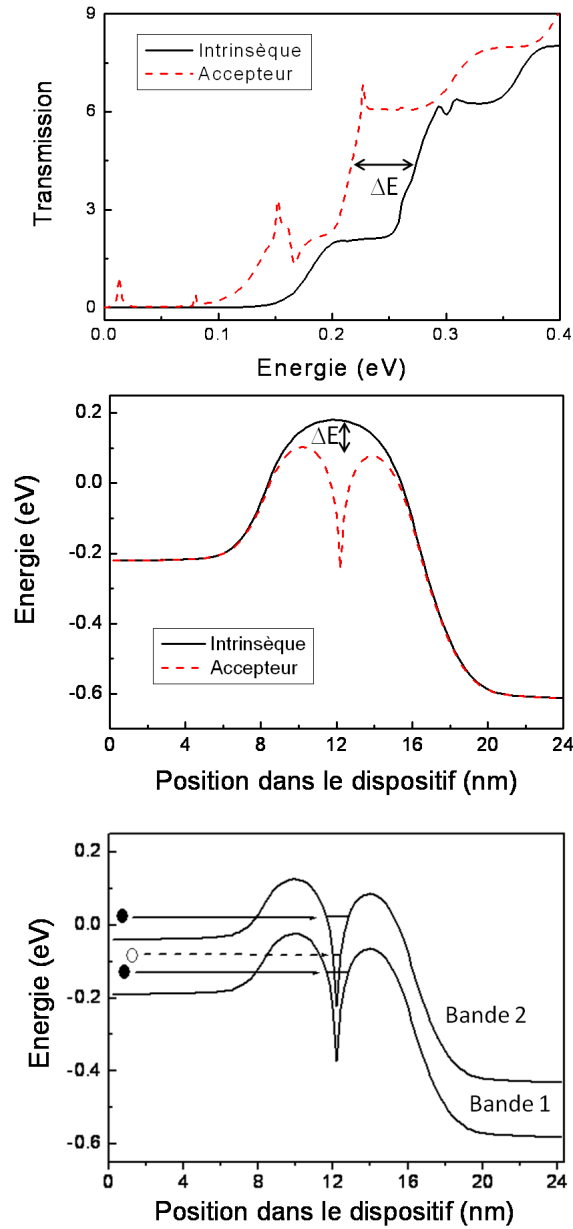


FIG. 4.10: a) Coefficient de transmission pour un transistor à nanofil de type p sans impureté (ligne continue) et avec un donneur (pointillés) au milieu du canal de conduction, b) Pseudo sous-bande le long du dispositif sans impureté (ligne continue) et avec un donneur (pointillés), et c) Représentation schématique du transport de trous à travers un potentiel attractif de Coulomb. Transmission résonnante (flèche continue) et anti-résonnante (flèche pointillée). $V_{DS} = -0.4$ V et $V_G = 0$ V.

4.2.3 Conclusion

Nous avons présenté un code de transport quantique en 3D dans l'espace réel permettant de simuler les transistors à nanofil de type p. Nous avons ainsi étudié les propriétés de transport de la bande de valence et l'influence d'une impureté ionisée sur celles-ci. Pour ce faire, nous avons calculé le coefficient de transmission, la densité locale d'états et le courant pour différentes tensions de grille et différents dopages. Nous avons montré que le couplage inter-bandes joue un rôle très important dans la bande de valence même pour un canal intrinsèque, notamment en régime *ON*. Nous avons également observé qu'en présence d'un défaut accepteur, les résonances et les anti-résonances dues aux états quasi-localisés apparaissent pour toutes les tensions de grille. Au final, le couplage inter-bandes et la présence d'impuretés rendent imprévisible les caractéristiques de transport dans les transistors à nanofil de type p. Il est fort probable que la présence aléatoire d'une impureté dans le canal entraîne encore plus de variabilité que dans un dispositif de type n. Il est donc indispensable de prendre en compte le caractère multibandes de la bande de valence afin de décrire correctement le transport quantique des trous dans les transistors à nanofil.

Conclusion Générale

La diminution effrénée des dimensions est telle que la longueur de grille des MOSFETs devrait, d'ici peu, descendre en dessous de la barre des 10 nm. Les effets quantiques seront tellement importants que l'architecture "classique" du MOSFET ne devrait plus fonctionner correctement. Ainsi, l'étude des dispositifs de demain nous a amené à modéliser et simuler des transistors ayant une architecture innovante. Nous nous sommes consacré, dans cette optique, à l'étude de MOSFET à nanofil de silicium.

Afin d'atteindre cet objectif, nous avons commencé par présenter les méthodes de physique théorique que nous avons utilisées pour simuler les dispositifs.

Le premier chapitre est consacré au transport électronique dans les nanocomposants. Nous avons brièvement défini les limites de l'approche classique avant d'introduire l'approche quantique. Cette approche implique notamment de considérer l'effet tunnel, que nous avons expliqué et qui introduit la notion de probabilité de passage. Nous avons également étudié l'approche de Landauer qui décrit le système comme étant composé d'un conducteur balistique relié à chaque extrémité à un réservoir macroscopique semi-infini à l'équilibre thermodynamique. Nous avons montré comment calculer le courant avec la formule de Landauer et mis en évidence la notion de quantum de conductance. La formule de Landauer nous dit que le calcul du courant nécessite la connaissance de la transmission qui dépend du potentiel du conducteur. Mais le potentiel et la densité de charges étant dépendants l'un de l'autre, nous avons montré, à l'aide d'un exemple simple, l'importance d'effectuer un calcul auto-cohérent des équations de Schrödinger et Poisson. Cet exemple nous a aussi permis de mettre en évidence la notion de résistance des réservoirs.

Le deuxième chapitre est consacré au formalisme des fonctions de Green hors-équilibre. Nous avons commencé par voir que d'un point de vue mathématique une fonction de Green est une solution élémentaire d'une équation différentielle linéaire à coefficients constants. Nous avons ensuite défini la fonction de Green retardée d'un système quantique. Par la suite, nous avons introduit le concept de *self-energy* de contact qui permet notamment de transformer un opérateur de taille infinie en un opérateur de taille finie et donc de le traiter numériquement le problème. Les *self-energies* introduisent une partie imaginaire au hamiltonien qui provoque l'élargissement des niveaux d'énergie. Cet élargissement qui implique la limitation du

temps de vie sur les niveaux est indispensable au transport. Nous avons ensuite présenté l'équation de Dyson qui permet de formaliser n'importe quel système perturbé. Nous avons aussi montré comment calculer la densité de porteurs dans un système isolé avant de passer à un système 3D multibandes ouvert. Afin de pouvoir traiter un tel cas, nous avons montré comment calculer les *self-energies* à l'aide de la décomposition de Schur. Enfin, nous avons montré comment calculer la densité de charge et le courant dans le formalisme des fonctions de Green, et comment réduire le temps de ce calcul (extrêmement lourd en 3D) grâce à un algorithme de récursion basé sur l'équation de Dyson.

Les deux autres chapitres présentent les résultats de simulation effectués à l'aide des outils théoriques évoqués précédemment.

Le troisième chapitre est consacré à l'étude d'une architecture innovante permettant d'améliorer le fonctionnement des transistors à nanofil "classiques", le *notch-MOSFET*. Le but de cette architecture est de diminuer le courant *OFF* en réduisant le courant tunnel tout en préservant le courant *ON*. Ceci est réalisé en rajoutant un *notch* d'oxyde à la fin du canal de conduction. Nous avons effectué plusieurs simulations avec différentes configurations du *notch*. Nous avons fait varier la hauteur, la longueur et la configuration (d'un seul côté et de deux côtés) de celui-ci. Nos résultats montrent, comme escompté, une diminution du courant *OFF* en présence d'un *notch*. Nous voyons de plus que le rapport I_{ON}/I_{OFF} augmente avec l'augmentation de la hauteur ou de la longueur. Un résultat inattendu est l'augmentation du courant *ON* pour certaines configurations du *notch*. Ceci est due à l'apparition d'états localisés quasi-résonnants dans la zone pré-*notch*. Les meilleures performances ont été obtenues pour un "notch simple" de longueur $L = 1$ nm et de hauteur $h = 0.4$ nm. Cette configuration permet d'améliorer de 32% le rapport I_{ON}/I_{OFF} sans réduction du courant *ON*.

Le quatrième chapitre est consacré à l'étude des transistors à nanofil de type p. Nous avons pour cela choisi d'utiliser un hamiltonien $\mathbf{k} \cdot \mathbf{p}$ à six bandes. En effet, une approche multibandes est indispensable à cause de la forte non-parabolicité de la bande de valence. Nous avons fait une brève introduction à la théorie $\mathbf{k} \cdot \mathbf{p}$ et avons montré comment l'utiliser pour modéliser les nanocomposants. Nous avons commencé par simuler des transistors à canal intrinsèque. Nous avons observé des phénomènes propres au caractère multibandes du hamiltonien. En effet, la transmission laisse apparaître des pics, de plus en plus nombreux quand $|V_G|$ augmente, dus

aux couplages inter-bandes. Nous avons également simulé des transistors avec une impureté ionisée au milieu du canal. Dans le cas d'une impureté donneur, l'augmentation de la hauteur de barrière entraîne une diminution du courant, très importante en régime *OFF* (environ deux décades) et moins importante en régime *ON* (environ une décade) à cause de l'écrantage. Par ailleurs, nous montrons que la position de l'impureté joue un rôle important. Une impureté située au bord de la section transversale a moins d'impact sur le transport qu'une impureté au centre avec un courant trois fois plus grand dans le premier cas. La présence d'une impureté accepteur crée un potentiel attractif. Le puits de potentiel ainsi créé entraîne l'apparition d'états quasi-localisés. Nous avons observé des pics de transmission résultants d'effets tunnel résonnants et des creux résultants d'interactions anti-résonnantes correspondants à ces états. Tous ces phénomènes rendent indispensable la prise en compte du caractère multibandes pour décrire le transport des trous dans les transistors à nanofil.

Le code que nous avons utilisé pour simuler les MOSFETs de type n, dans le cadre de l'étude du *notch*, a été fortement modifié afin de pouvoir simuler des MOSFETs de type p. Il est envisageable de le modifier à nouveau pour d'autres études comme l'impact de la rugosité de surface, de l'orientation [64], des contraintes mécaniques [65] ou de l'utilisation d'autres matériaux semiconducteurs dans les transistors de type p. Ces modifications sont *a priori* relativement facile à mettre en place. La rugosité nécessite de modifier la géométrie du canal, le changement d'orientation nécessite de recalculer le hamiltonien 3D et le changement de matériau se fait en utilisant les paramètres de Luttinger adéquats. D'autres modifications plus complexes sont aussi envisageables comme la prise en compte des interactions électron-phonon qui ont déjà été implémentés au sein de notre équipe dans un code 2D. Nous pouvons aussi envisager un passage à un hamiltonien $\mathbf{k} \cdot \mathbf{p}$ à huit bandes pour modéliser simultanément la bande de valence et de conduction. Ceci permettrait d'étudier des dispositifs tels que les TFETs ou encore de s'orienter vers le photovoltaïque.

Bibliographie

- [1] M. Bescond, K. Nehari, J.L. AuTRAN, N. Cavassilas, D. Munteanu, and M. Lannoo, "3D quantum modeling and simulation of multiple-gate nanowire MOS-FETs," *IEDM Tech. Digest IEEE*, 617, San Francisco (USA), (2004).
- [2] R. E. Bank, D. J. Rose et W. Fichtner, "Numerical methods for semiconductor device simulation," *IEEE Trans. Electron Devices* **30**, 1031 (1983).
- [3] C. Jacoboni et L. Reggiani, "The Monte Carlo method for the solution of charge transport in semiconductors with applications to covalent materials," *Rev. Mod. Phys.* **55**, 645 (1983).
- [4] M. Le Bellac et F. Montessagne, "Thermodynamique Statique : équilibre et hors équilibre : cours, exercices et problèmes résolus," *Dunod* (2001).
- [5] Y. Apanovich, "Steady-state and transient analysis of submicron devices using energy balance and simplified hydrodynamic models," *Trans. Electron Devices* **13** (1994).
- [6] Z. Yu, R. W. Dutton et R. A. Kielh, "Circuit/device modeling at the quantum level," *IEEE Trans. Electron Device* **47**, 1819 (2000).
- [7] R. Landauer, "Spatial variation of currents and fields due to localized scatterers in metallic conduction," *IBM J. Res. Dev.* **1**, 223 (1957).
- [8] R. Landauer, "Conductance determined by transmission : probes and quantized constriction resistance," *J. Phys. : Condens. Matter* **1**, 8099 (1989).
- [9] M. Büttiker, Y. Imry, R. Landauer et S. Pinhas, "Generalized many-channel conductance formula with application to small rings," *Phys. rev. B* **31**, 6207 (1985).
- [10] Z. Ren, R. Venugopal, S. Goasguen, S. Datta et M.S. Lundstrom, "nanoMOS 2.5 : A two dimensional simulator for quantum transport in double-gate MOSFETs," *IEEE Trans. Electron Dev.* **50**, 1914 (2003).

- [11] S. Datta, "Nanoscale modeling : the Green's function method," *Superlattices Microstruct.* **28**, 253 (2000).
- [12] N. J. Higham et H. M. Kim, "Numerical analysis of a quadratic matrix equation," *IMA J. Num. Analysis* **20**, 499 (2000).
- [13] A. Martinez, N. Seoane, A.R. Brown, J.R. Barker et A. Asenov, "3-D Nonequilibrium Green's Function Simulation of Nonperturbative Scattering From Discrete Dopants in the Source and Drain of a Silicon Nanowire Transistor," *IEEE Trans. on Nanotechnology* **8**, 603 (2009).
- [14] S. Datta, "Quantum Transport : Atom to Transistor," *Cambridge University Press, Cambridge, England* (2005).
- [15] N. Cavassilas, M. Bescond et J.L. Autran, "Improvement of current-control induced by oxide notch in very short field-effect transistor," *Applied Physics Letters* **87**, 073509 (2005).
- [16] M. V. Fischetti, Z. Ren, P. M. Solomon, M. Yang et K. Rim, "Six-band $k \cdot p$ calculation of the hole mobility in silicon inversion layers : Dependence on surface orientation, strain, and silicon thickness," *J. Appl. Phys.* **94**, 1079 (2003).
- [17] M. De Michielis, D. Esseni, Y. L. Tsang, P. Palestri, L. Selmi, A. G. O'neill et S. Chattopadhyay, "A Semianalytical Description of the Hole Band Structure in Inversion Layers for the Physically Based Modeling of pMOS Transistors," *IEEE Trans. Electron Devices* **54**, 2164 (2007).
- [18] E. X. Wang, P. Matagne, L. Shifren, B. Obradovic, R. Kotlyar, S. Cea, M. Stettler et M. D. Giles, "Physics of Hole Transport in Strained Silicon MOSFET Inversion Layers," *IEEE Trans. Electron Devices* **53**, 1840 (2006).
- [19] M. Luisier et G. Klimeck, "Atomistic full-band simulations of silicon nanowire transistors : Effects of electron-phonon scattering," *Phys. Rev. B* **80**, 155430 (2009).
- [20] M. Shin, "Full-quantum simulation of hole transport and band-to-band tunneling in nanowires using the $k \cdot p$ method," *J. Appl. Phys.* **106**, 054505 (2009).
- [21] M. Shin, S. Lee et G. Klimeck, "Computational Study on the Performance of Si Nanowire pMOSFETs Based on the Method," *IEEE Trans. Electron Devices* **57**, 2274 (2010).

-
- [22] M. Luisier, “Phonon-limited and effective low-field mobility in n- and p-type [100]-, [110]-, and [111]-oriented Si nanowire transistors,” *Appl. Phys. Lett.* **98**, 032111 (2011).
- [23] M. V. Fischetti et S. E. Laux, “Band structure, deformation potentials, and carrier mobility in strained Si, Ge, and SiGe alloys,” *J. Appl. Phys.* **80**, 2234 (1996).
- [24] D. Rideau, M. Feraille, M. Michailat, Y. M. Niquet, C. Tavernier et H. Jaouen, “On the validity of the effective mass approximation and the Luttinger $k \cdot p$ model in fully depleted SOI MOSFETs,” *Solid-State Electron* **53**, 452 (2009).
- [25] W. Frensley, “Boundary conditions for open quantum systems driven far from equilibrium,” *Rev. Mod. Phys.* **62**, 745 (1990).
- [26] S. Martinie, “Modélisation analytique du transport quasi-balistique, quantique et non-stationnaire pour les circuits à base de nanotransistors multigrilles,” *PhD report* (2009).
- [27] M. Gritsch, “Numerical modelling of silicon-on-insulator MOSFETs,” *PhD report* (2002).
- [28] K. Blotekjaer, “Transport equations for electrons in two-valley semiconductors,” *Trans. Electron Devices* **17**, 38 (1970).
- [29] Y. Apanovich, E. Lyumkis, B. Polsky, A. Shur, P. Blakey, “Steady-state and transient analysis of submicron devices using energy balance and simplified hydrodynamic models,” *Trans. CAD* **13**, 38 (1994).
- [30] D. K. Ferry et S. M. Goodnick, “Transport in Nanostructures,” *Cambridge University Press, Cambridge, England* (1997).
- [31] A. Wettstein, A. Schenk, W. Fichtner, “Quantum Device-Simulation with the Density-Gradient Model on Unstructured Grids,” *Trans. Electron Devices* **48**, 279 (2001).
- [32] D. Bohm, “A suggested interpretation of the quantum theory in terms of hidden variables,” *Phys. Rev.* **85**, 166 (1952).
- [33] M. G. Ancona, “Quantum correction to the equation of state of an electron gas in a semiconductor,” *Phys. Rev. B* **39**, 9536 (1989).
- [34] D. Munteanu, “Modélisation et simulation numérique des dispositifs Multi-grilles,” *Rapport HDR* (2008)

- [35] A. S. Spinelli, A. Benvenuti, A. Pacelli, "Self-consistent 2-D model for quantum effects in n-MOS transistors," *Trans. Electron Devices* **45**, 1342 (1998).
- [36] D. Querlioz, Jérôme Saint-Martin, Van-Nam Do, Arnaud Bournel, Philippe Dollfus, "A study of quantum transport in end-of-roadmap DG-MOSFETs using a Fully Self-Consistent Wigner Monte Carlo Approach," *Trans. Nanotechnology* **5**, 737 (2006).
- [37] M. A. Jaud, "Approche du potentiel effectif pour la simulation Monte Carlo du transport électronique avec effets de quantification dans les dispositifs MOSFETs," *PhD report* (2006).
- [38] J. Saint-Martin, D. Querlioz, A. Bournel, P. Dollfus, "Efficient multi sub-band Monte Carlo simulation of nano-scaled double gate MOSFETs," *SISPAD*, 216 (2006).
- [39] J. H. Davies, "The physics of low-dimensional semiconductors" *Cambridge University Press Cambridge, UK* (1998).
- [40] M. Bescond, N. Cavassilas, K. Kalna, K. Nehari, L. Raymond, J.L. Autran, M. Lannoo, A. Asenov, "Ballistic transport in Si, Ge and GaAs nanowire MOSFETs," *IEDM Tech. Digest IEEE*, Washinton (USA), (2005).
- [41] C. Kittel, "Physique de l'état solide," *Dunod* (2007).
- [42] B. J. Van Wees, H. Van Houten, C. W. J. Beenakker, J. G. Williamson, L. P. Kouwenhoven, D. Van der Marel et C. T. Foxon, "Quantized conductance of point contacts in a two-dimensional electron gas," *Phys. Rev. Lett.* **60**, 848 (1988).
- [43] P. L. Pernas, A. Martin-Rodero et F. Flores, "Electrochemical-potential variations across a constriction," *Phys. Rev. B* **41**, 8553 (1990).
- [44] V. Mujica, A. E. Roitberg et M. Ratner, "Molecular wire conductance : Electrostatic potential spatial profile," *J. Chem. Phys.* **112**, 6834 (2000).
- [45] E. N. Economou, "Green's Functions in Quantum Physics," *Springer-Verlag, Berlin* (1979).
- [46] C. Delerue et M. Lannoo, "Nanostructures, theory and modeling," *Springer, Berlin* (2004).
- [47] R. D. Mattuck "Aguide to Feynman Diagrams in the Manu-Body Problem" *2nd ed, Dover Publication*.

-
- [48] A. Svizhenko, M. P. Anantram, T. R. Govindan, B. Biegel et R. Venugopal, "Two-dimensional quantum mechanical modeling of nanotransistors," *J. Appl. Phys.* **91**, 2343 (2002).
- [49] K. Nehari, N. Cavassilas, J. L. Autran, M. Bescond, D. Munteanu et M. Lannoo, "Influence of band structure on electron ballistic transport in silicon nanowire MOSFET's : An atomistic study," *Solid-State Electronics* **50**, 716 (2006).
- [50] J. Wang, A. Rahman, A. Ghosh, G. Klimeck et M. Lundstrom, "On the validity of the parabolic effective-mass approximation for the I-V calculation of silicon nanowire transistors," *IEEE Trans. Electron Devices* **52**, 7 (2005).
- [51] S. Jin, T.-W. Tang et M. V. Fischetti, "Theoretical Study of Carrier Transport in Silicon Nanowire Transistors Based on the Multisubband Boltzmann Transport Equation," *IEEE Trans. Electron Devices* **55**, 3 (2008).
- [52] K. Nehari, N. Cavassilas, F. Michelini, M. Bescond, J. L. Autran et M. Lannoo, "Full band study of current across silicon nanowire transistors," *Appl. Phys. Lett.*, **90**, 132112 (2007).
- [53] R. Venugopal, Z. Ren, S. Datta, M.S. Lundstrom and M.S. Jovanovic, "Simulating quantum transport in nanoscale transistors : real versus mode-space approaches," *J. Appl. Phys.* **97**, 3730 (2002).
- [54] N. Pons, N. Cavassilas, F. Michelini, L. Raymond et M. Bescond, "Original shaped nanowire metal-oxide-semiconductor field-effect transistor with enhanced current characteristics based on three-dimensional modeling," *J. Appl. Phys.* **106**, 053711 (2009).
- [55] J. M. Luttinger et W. Kohn, "Motion of Electrons and Holes in Perturbed Periodic Fields ," *Phys. Rev.*, **97**, 869 (1955).
- [56] E. O. Kane, "Band structure of indium antimonide," *Journal of the Physics and Chemistry of Solids*, **1** , 249 (1957).
- [57] G. F. Koster, R. G. Wheeler, J. O. Dimmock et H. Statz, "The Properties of the Thirty-Two Point Groups", *MIT Press, Cambridge massachusetts USA* (1963).
- [58] M. Tiersten, "Acoustic-Mode Scattering of Holes," *IBM J. Res. Dev.* **5**, 122 (1961).

- [59] M. Bescond, M. Lannoo, L. Raymond et F. Michelini, “Single donor induced negative differential resistance in silicon n-type nanowire metal-oxide-semiconductor transistors,” *J. Appl. Phys.* **107**, 093703 (2010).
- [60] A. Svizhenko, P. W. Leu et K. Cho, “Effect of growth orientation and surface roughness on electron transport in silicon nanowires,” *Phys. Rev. B* **75**, 125417 (2007).
- [61] N. Pons, N. Cavassilas, M. Bescond, F. Michelini, L. Raymond, M. Lannoo, “Three-dimensional $k \cdot p$ quantum simulation of p-type nanowire MOS transistors : Influence of ionized impurity,” *Proceedings of IWCE, Pisa*, 89 (2010).
- [62] M. Lannoo et J. Bourgoin, “Point defects in semiconductors,” *Springer-Verlag, Berlin* (1981).
- [63] M. Bescond, M. Lannoo, F. Michelini, L. Raymond, and M.G. Pala, “Influence of Ionized Impurities in Silicon Nanowire MOS Transistors,” *Proceedings of the International Workshop on Computational Electronics, Beijing*, 137 (2009).
- [64] N. Cavassilas, N. Pons, F. Michelini, M. Bescond, “Multiband quantum transport simulations of ultimate p-type double-gate transistors : Influence of the channel orientation,” *Appl. Phys. Lett.*, **96**, 102102 (2010).
- [65] F. Michelini, N. Pons, N. Cavassilas, M. Bescond, M. Szczap, “Spin-orbital strain effects in semiconductors,” *Proceedings of IWCE, Pisa*, 147 (2010).