



HAL
open science

Modélisation et simulation du transport quantique dans les transistors MOS nanométriques

Marc Bescond

► **To cite this version:**

Marc Bescond. Modélisation et simulation du transport quantique dans les transistors MOS nanométriques. Matière Condensée [cond-mat]. Université de Provence - Aix-Marseille I, 2004. Français. NNT: . tel-00008075

HAL Id: tel-00008075

<https://theses.hal.science/tel-00008075>

Submitted on 14 Jan 2005

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THESE

présentée devant

L'UNIVERSITE DE PROVENCE (AIX-MARSEILLE I)

par

Marc BESCOND

pour obtenir

LE GRADE DE DOCTEUR

Ecole Doctorale : Physique, Modélisation et Sciences pour l'Ingénieur

Spécialité : Physique et modélisation des systèmes complexes

MODELISATION ET SIMULATION DU TRANSPORT QUANTIQUE DANS LES TRANSISTORS MOS NANOMETRIQUES

Directeur de thèse : Michel LANNOO

Soutenue le 26 novembre 2004 devant la commission d'examen :

| | | |
|-------------|---------------------------|-------------------------------------|
| Rapporteurs | Christophe DELERUE | IEMN-CNRS, Lille |
| | Philippe DOLLFUS | IEF-CNRS, Orsay |
| Examineurs | Asen ASENOV | Dev. Modelling Group, Univ. Glasgow |
| | Martin STÄDELE | INFINEON, München |
| | Michel LANNOO | L2MP-CNRS, Marseille |
| | Jean-Luc AUTRAN | L2MP, Univ. Aix-Marseille I |
| Invités | Jacques GAUTIER | CEA-LETI, Grenoble |
| | Didier GOGUENHEIM | L2MP, ISEN-Toulon |

Remerciements

Ce travail a été effectué au Laboratoire Matériaux et Microélectronique de Provence (L2MP), dirigé par Monsieur MICHEL LANNOO, dans l'équipe des Dispositifs Ultimes sur Silicium dirigée par Monsieur JEAN-LUC AUTRAN. Je les remercie de m'avoir accueilli.

Je remercie Monsieur JEAN-LUC AUTRAN d'avoir bien voulu présider le Jury chargé de juger ma thèse.

Messieurs CHRISTOPHE DELERUE et PHILIPPE DOLLFUS se sont intéressés à ce travail et ont accepté d'en être les rapporteurs. Qu'ils trouvent l'expression de ma profonde gratitude.

Messieurs ASEN ASENOV et MARTIN STÄDELE me font également un grand honneur en se déplaçant depuis Glasgow et Munich respectivement pour participer au Jury.

Je remercie Monsieur DIDIER GOGUENHEIM de m'avoir accueilli durant ces trois années dans son équipe enseignante et d'avoir bien voulu être invité à ce Jury.

Que Monsieur JACQUES GAUTIER reçoive toute ma gratitude pour s'être intéressé à mon travail et pour avoir accepté d'être invité à ma soutenance.

Je tiens à remercier de nouveau M. LANNOO qui, en dépit de ses lourdes responsabilités au département Sciences Physiques et Mathématiques du CNRS, a assuré la direction de ma thèse, et m'a offert le maximum de son temps disponible pour faire avancer les problèmes qui se sont présentés tout au long de ces trois années.

J'exprime ma profonde reconnaissance à J.L. AUTRAN, pour son aide, son enthousiasme et son souci d'assurer le déroulement de mon travail dans les meilleures conditions.

Que NICOLAS CAVASSILAS reçoive toute ma gratitude pour s'être intéressé quotidiennement à mon travail et pour les nombreuses discussions enrichissantes que nous avons eues.

Je remercie également LAURENT RAYMOND dont les compétences en traitement numérique m'ont secourues à plusieurs reprises.

DANIELA MUNTEANU a suivi régulièrement mon travail et a été d'une grande aide pour la relecture d'articles. Je lui en suis très reconnaissant.

Merci aussi à EVELYNE SANTACROCE pour avoir géré autant de d'ordres de mission et de bons de transport en gardant le sourire.

De nombreuses personnes du laboratoire m'ont encouragé, conseillé, motivé, bref accompagné. A cet égard, j'aimerais remercier l'équipe de Physique Théorique du L2MP, et plus particulièrement Gilbert, Steffen, Anne-Marie, Roland et Pierre... l'équipe de Fiabilité, Conception et moléculaire: Didier, Alain, Edith, Stéphane, Lionel, Philippe, Claude, et Jaqueline... et mes collègues thésards ou anciens thésards: Thierry, Xavier, Fabrice, Eric, Samir, Olivier, Samuel, Simon, Vincent, Céline, Pascal, Alexandre, Valéry, Nicolas, Juliano et Karim qui commence sa thèse. Je lui souhaite bonne chance.

Je remercie le C.N.R.S, la région Provence-Alpe-Côte d'Azur et la société Gemplus pour le soutien financier au travers de la bourse de thèse BDI.

Merci enfin, à ma famille et à mes amis pour tout ce que je leur dois et parce qu'ils ont été le soutien essentiel de ces trois années.

Table des matières

| | |
|--|-----------|
| Notations et constantes | 7 |
| Introduction | 9 |
| 1 Evolution de la microélectronique : rencontre avec le quantique | 11 |
| 1.1 Miniaturisation des transistors | 12 |
| 1.1.1 Pourquoi réduire la taille des transistors ? | 12 |
| 1.2 Le MOSFET | 12 |
| 1.3 Les architectures émergentes | 17 |
| 1.3.1 Transistors à canal de silicium contraint | 17 |
| 1.3.2 Les transistors SOI | 18 |
| 1.3.3 Les transistors multi-grilles | 19 |
| 1.4 Les transistors à nanotube de carbone | 23 |
| 1.4.1 Nanostructures de carbone et nanotubes | 23 |
| 1.4.2 Les transistors à nanotubes de carbone | 30 |
| 1.5 Conclusion : quel avenir pour le transistor ? | 34 |
| 2 Méthode de calcul de la structure de bandes | 35 |
| 2.1 Introduction : système d'électrons indépendants | 36 |
| 2.2 Methodes <i>ab initio</i> | 37 |
| 2.2.1 Equations de HARTREE et approximation du champ moyen | 37 |

| | | |
|----------|--|-----------|
| 2.2.2 | Equations de HARTREE-FOCK | 38 |
| 2.2.3 | Fonctionnelle de la densité | 39 |
| 2.2.4 | Approximation GW | 41 |
| 2.2.5 | Méthodes numériques | 41 |
| 2.3 | Méthodes semi-empiriques | 44 |
| 2.3.1 | Pseudopotentiels empiriques | 45 |
| 2.3.2 | Liaisons fortes | 46 |
| 2.3.3 | L'auto-cohérence en liaisons fortes | 47 |
| 2.3.4 | Masse effective et théorie k.p | 48 |
| 3 | Transport électronique dans les nano-composants | 53 |
| 3.1 | Transport électronique dans un nano-MOSFET : rencontre avec le quantique . . | 54 |
| 3.1.1 | Inadéquation des modèles classiques : du régime diffusif au balistique . . . | 54 |
| 3.1.2 | Description physique d'un nanocomposant : cas général | 60 |
| 3.2 | Les limites du couplage faible | 61 |
| 3.2.1 | La théorie des perturbations | 61 |
| 3.2.2 | La théorie orthodoxe du courant tunnel | 63 |
| 3.3 | Au delà de la théorie des perturbations | 66 |
| 3.3.1 | Formule de Landauer | 67 |
| 3.3.2 | Intérêt de la matrice densité | 70 |
| 3.3.3 | Formalisme de la diffusion | 74 |
| 3.3.4 | Calcul des fonctions de Green et Self-énergie | 83 |
| 3.4 | Interactions électron-électron et auto-cohérence | 85 |
| 3.4.1 | Calculs auto-cohérents en théorie du champ moyen | 86 |
| 3.4.2 | Profil du potentiel auto-cohérent | 87 |
| 3.5 | Conclusion | 88 |

| | | |
|----------|---|------------|
| 4 | Modélisation atomique de transistors MOSFETs double-grille à canal de conduction confiné | 91 |
| 4.1 | Les différentes approches | 92 |
| 4.1.1 | L'approximation de la masse effective : fonctions d'ondes délocalisées | 92 |
| 4.1.2 | Les liaisons fortes : fonctions d'ondes localisées | 94 |
| 4.1.3 | Comparaison et limites des deux approches | 96 |
| 4.2 | Modélisation du transistor double-grille à canal unique de conduction | 97 |
| 4.2.1 | Présentation du modèle et aspects numériques | 97 |
| 4.2.2 | Résultats et discussion | 110 |
| 4.3 | Dépendance du courant tunnel source-drain en fonction de la longueur de grille et du travail de sortie de son métal | 117 |
| 4.4 | Conclusion | 122 |
| 5 | Modélisation 3D de MOSFETs multi-grilles | 123 |
| 5.1 | Approximation de l'espace des modes | 124 |
| 5.1.1 | Confinement quantique dans les architectures MOSFETs innovantes | 124 |
| 5.1.2 | Théorie | 125 |
| 5.1.3 | Justification de l'approche "mode-space" | 131 |
| 5.1.4 | Résolution de l'équation de Poisson 3D et auto-cohérence | 133 |
| 5.2 | Application au MOSFET Gate-All-Around | 135 |
| 5.3 | Comparaison des différentes architectures émergentes | 141 |
| 5.4 | Conclusion | 146 |
| 6 | Traitement et influence de défauts ponctuels dans les nano-transistors | 149 |
| 6.1 | Défauts ponctuels dans une chaîne à l'équilibre | 150 |
| 6.2 | Influence du défaut dans un MOSFET double-grille | 156 |
| 6.3 | Défauts ponctuels ionisés dans les MOSFETs multi-grilles | 160 |
| 6.4 | Résultats | 167 |

| | |
|--|------------|
| 6.5 Conclusion | 171 |
| Conclusion | 173 |
| Appendice 1. Les fonctions de Green | 189 |
| Appendice 2. Equivalence entre les deux formules de courant | 193 |

Notations et constantes

Constantes fondamentales

| | | |
|---|--------------|--|
| Module de la charge élémentaire de l'électron | e | $= 1.602189 \cdot 10^{-19} \text{ C}$ |
| Constante de BOLTZMANN | k_B | $= 1.38066 \cdot 10^{-23} \text{ J/K}$ |
| Constante de PLANCK | h | $= 6.62618 \cdot 10^{-34} \text{ Js}$ |
| Constante de PLANCK réduite | \hbar | $= 1.054589 \cdot 10^{-34} \text{ Js}$ |
| Masse de l'électron | m_e | $= 9.10953 \cdot 10^{-31} \text{ kg}$ |
| Célérité de la lumière dans le vide | c | $= 2.997925 \cdot 10^8 \text{ m/s}$ |
| Permittivité diélectrique du vide | ϵ_0 | $= 8.854 \cdot 10^{-12} \text{ F/m}$ |
| Perméabilité magnétique du vide | μ_0 | $= 4\pi \cdot 10^{-7} \text{ H/m}$ |

Propriétés du silicium

| | | |
|--|------------------------|---------------------------------|
| Maille élémentaire du réseau | a | $= 0.543 \text{ nm}$ |
| Energie du Gap à 300K | E_G | $= 1.12 \text{ eV}$ |
| Affinité électronique | χ_{Si} | $= 4.01 \text{ eV}$ |
| Permittivité relative | ϵ_{Si} | $= 11.9$ |
| Masse effective transverse | m_t | $= 0.19 m_e$ |
| Masse effective longitudinale | m_l | $= 0.92 m_e$ |
| Mobilité des électrons (matériau peu dopé) | μ_n | $= 0.15 \text{ m}^2/\text{Vs}$ |
| Mobilité des trous (matériau peu dopé) | μ_p | $= 0.045 \text{ m}^2/\text{Vs}$ |

Propriétés de la silice

| | | |
|-----------------------|---------------------------|--------------------|
| Energie du Gap | $E_{G_{\text{SiO}_2}}$ | $= 9.1 \text{ eV}$ |
| Affinité électronique | χ_{SiO_2} | $= 0.9 \text{ eV}$ |
| Permittivité relative | ϵ_{SiO_2} | $= 3.8$ |
| Masse des électrons | m_{SiO_2} | $= 0.5 m_e$ |

Propriétés de quelques métaux

| | | |
|---|--------------------|---------------------|
| Masse de l'électron | m | $= m_e$ |
| Travail de sortie de l'or | ϕ_{Au} | $= 5 \text{ eV}$ |
| Hauteur du niveau de Fermi de l'or | $E_{F \text{ Au}}$ | $= 5.5 \text{ eV}$ |
| Travail de sortie de l'aluminium | ϕ_{Al} | $= 4.1 \text{ eV}$ |
| Hauteur du niveau de Fermi de l'aluminium | $E_{F \text{ Al}}$ | $= 11.6 \text{ eV}$ |

Introduction

Depuis les années 60, la capacité d'intégration en microélectronique suit une évolution exponentielle, avec une densité de transistors par microprocesseur doublée tous les 3 ans. Cette évolution, connue sous le nom de loi de MOORE [1, 2], a conduit à la fabrication de microprocesseurs de plus de 400 millions de transistors et à la réalisation de mémoires dont la densité dépasse le Gbit. L'augmentation de la densité d'intégration et l'amélioration des performances sont rendues possibles par la diminution de la taille des transistors. La grandeur caractéristique d'un transistor qui distingue une génération de transistors de la suivante est la longueur de grille L_G . Actuellement la plus petite longueur de grille produite est de 37 nm et la réduction des dimensions demeure essentielle pour réaliser de nouvelles innovations technologiques. L'ITRS (International Technology Roadmap for Semiconductors) [3] dont les prévisions sont établies par les principales industries des semi-conducteurs annonce une longueur de grille de 9 nm en 2016. De telles dimensions génèrent des problèmes technologiques bien plus complexes et surtout différents de ceux rencontrés dans les transistors actuels. En plus de la nécessité d'optimiser les procédés de fabrication, il est indispensable de connaître les limites physiques de la technologie CMOS (pour l'anglais "Complementary Metal-Oxide-Semiconducteur"), au delà desquelles un transistor technologiquement parfait n'aurait plus des propriétés électriques tolérables. Bien que la réponse ne soit pas encore évidente, il apparaît clairement que de nouvelles architectures devront remplacer à plus ou moins longue échéance les transistors conventionnels. Il a déjà été décidé par exemple de ne plus réaliser un type de transistor valable pour toutes les applications, mais une architecture particulière pour les applications "hautes performances" (microprocesseurs) ou pour les applications "basses consommations" (téléphones portables). Le développement de transistors nanométriques ajoute aux problèmes expérimentaux, des défis théoriques qu'il est crucial de surmonter pour que perdure l'élan de miniaturisation. La diminution soutenue des dimensions accélère la rencontre de la microélectronique avec la mécanique quantique et d'autres lois régissent désormais le transport des électrons. La simulation des transistors a donc besoin de nouvelles théories et techniques de modélisation améliorant la compréhension physique des dispositifs de taille nanométrique. Dans ce contexte, les principaux objectifs de cette thèse sont : 1) d'identifier et de décrire les méthodes les plus pertinentes pour modéliser correctement ce type de composants, 2) d'examiner et de comprendre le transport de charges dans les nano-MOSFETs (pour l'anglais "Metal-Oxide Semiconductor Field-Effect-Transistors"), 3) d'explorer les différentes structures émergentes et de pointer les plus performantes.

Cette thèse s'articulera autour de quatre grands axes :

- Le chapitre 1 rappellera la problématique actuelle liée à la miniaturisation des transistors, présentera les réalisations les plus significatives des MOSFETs conventionnels (approche *Top-down*) et détaillera le cas prometteur d'une approche *Bottom-up* : le transistor à nanotube de carbone.

- Le chapitre 2 présentera les différentes méthodes de calcul des bandes d'énergie des matériaux alors que les théories du transport seront traitées dans le chapitre 3. Après avoir établi les limites des modèles classiques et semi-classiques, nous décrirons en détail la théorie de LANDAUER exprimée dans le formalisme des fonctions de GREEN hors-équilibre (NEGF pour "NonEquilibrium Green's Function") [4, 5]. A travers des exemples simples nous montrerons les surprenantes propriétés de transport d'un système mésoscopique en insistant sur la notion de transport balistique, de quantum de conductance et de résistance de contact.
- Dans le chapitre 4, nous appliquerons le formalisme NEGF à un transistor MOSFET double-grille à canal unique de conduction dans lequel l'axe source-canal-drain est modélisé par une chaîne atomique linéaire. Nous verrons qu'un tel système, ultimement confiné, conserve des propriétés électriques proches d'un MOSFET conventionnel. Nous étudierons ensuite l'influence de défauts ponctuels dans ce composant (chapitre 5) et discuterons l'impact sur les caractéristiques de courant du type de défaut et de sa localisation le long du canal.
- Ce formalisme (NEGF) sera ensuite étendu à un système à trois dimensions et servira à l'étude d'architectures émergentes de transistors plus réalistes (chapitre 6). Ainsi, nous verrons l'évolution du courant et des charges en fonction de la structure des transistors multi-grilles (tri-gate, gate-all-around etc...). Nous traiterons également l'influence des défauts ponctuels avant de resumer l'ensemble des résultats obtenus et de présenter les perspectives à envisager pour compléter le présent travail.

Chapitre 1

Evolution de la microélectronique : rencontre avec le quantique

La diminution de la taille des transistors MOSFETs vers l'échelle nanométrique s'accompagne de nouveaux effets électrostatiques et quantiques. Afin de retarder leur apparition, plusieurs stratégies sont envisagées à plus ou moins longue échéance.

Ce chapitre se propose de présenter les principaux effets issus de la réduction des dimensions et de décrire les limites des solutions technologiques déjà disponibles (section (1.1)). Nous discutons dans la section (1.3) des structures émergentes de MOSFETs à plusieurs grilles, susceptibles de poursuivre la miniaturisation dans les années à venir. Au delà, les futures générations de transistors pourraient être construites à partir de molécules. La section (1.4) étudie donc en détail le cas prometteur du transistor à nanotubes de carbone pressenti comme le successeur direct du MOSFET. La section (1.5) résume et ouvre sur les perspectives.

1.1 Miniaturisation des transistors

1.1.1 Pourquoi réduire la taille des transistors ?

La croissance de l'industrie des semi-conducteurs dépend pour l'instant de sa capacité à miniaturiser les transistors. L'objectif de la démarche est de délivrer de meilleures performances à moindre coût. Des circuits plus petits réduisent la surface globale de la puce électronique et permettent donc de produire plus de transistors sur un même wafer sans impacter sur le prix de fabrication. Le coût des circuits diminue ainsi d'un facteur deux tous les 18 mois. Les performances électriques des composants sont également améliorées (tableau(1.1)). En diminuant la dimension des MOSFETs, le temps de passage de l'état "off" à l'état "on" diminue linéairement du fait de l'évolution du temps de réponse intrinsèque \cong longueur de canal/vitesse des porteurs. Un autre avantage est la réduction de la consommation de puissance, utile pour augmenter la durée d'autonomie des systèmes mobiles mais aussi pour améliorer la fiabilité des systèmes hautes performances. Des puces plus petites consomment moins de puissance, donc moins d'énergie est utilisée pour chaque opération. En conséquence, le produit puissance-temps de réponse est réduit. Enfin, la rapidité de transmission de l'information dans un circuit intégré est limitée par la vitesse de l'impulsion électrique. Pour pouvoir augmenter la rapidité globale d'une opération, il faut réduire les distances géométriques, et empiler un maximum de données d'information dans un minimum d'espace pour les rapprocher. Cette évolution permet à un large public d'accéder à des services plus performants, moins chers et souvent nouveaux. Des objets technologiques innovants sont ainsi produits et créent de nouveaux marchés dont les retombées financières sont réinvesties dans la course à l'intégration (figure(1.1)).

| Année | 2004 | 2007 | 2010 | 2013 | 2016 |
|---|------|------|------|------|------|
| Noeud technologique (nm) | 90 | 65 | 45 | 32 | 22 |
| Longueur de grille physique (nm) | 37 | 25 | 18 | 13 | 9 |
| Tension d'alimentation (V) | 1.2 | 1.1 | 1 | 0.9 | 0.8 |
| Epaisseur d'oxyde équivalent (nm) | 1.2 | 0.9 | 0.7 | 0.6 | 0.5 |
| Tension de seuil en saturation (V) | 0.2 | 0.18 | 0.15 | 0.11 | 0.10 |
| Courant de fuite nominal des NMOS ($\mu\text{A}/\mu\text{m}$) | 0.05 | 0.07 | 0.1 | 0.3 | 0.5 |
| Courant conducteur nominal des NMOS ($\mu\text{A}/\mu\text{m}$) | 1110 | 1510 | 1900 | 2050 | 2400 |
| Temps de réponse intrinsèque des NMOS (ps) | 0.95 | 0.64 | 0.39 | 0.26 | 0.15 |

TAB 1.1 : Prévisions ITRS des caractéristiques des MOSFETs ultimes (high performances) [3].

1.2 Le MOSFET

Le principe de fonctionnement d'un transistor MOS repose sur "l'effet de champ électrique", qui consiste à moduler de façon électrostatique une densité de charges mobiles dans un semicon-



FIG. 1.1: Cycle de croissance de l'industrie des semi-conducteurs.

ducteur. Cette modulation est provoquée par un champ électrique agissant entre deux électrodes séparées par un diélectrique, et perpendiculaire à la direction de mouvement de ces charges. Dans cette structure, la tension de grille (V_G) commande la quantité de charges et la tension source-drain ($V_{DS} = V_D - V_S$) les met en mouvement (figure (1.2)). La figure (1.3) représente l'évolution du courant de drain (I_D) en fonction de la tension de grille. Le transistor MOSFET se caractérise par deux régimes de fonctionnement. A fort V_G , le courant de drain est maximal ($=I_{ON}$) : c'est le régime de saturation. A faible V_G , les charges mobiles minoritaires sont repoussées dans le substrat, et le courant de drain diminue exponentiellement : c'est le régime sous le seuil ou "de blocage".

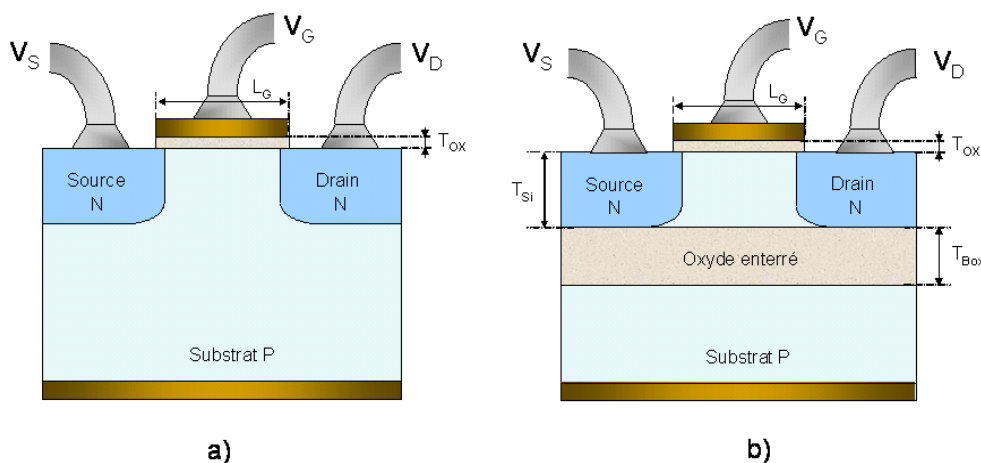


FIG. 1.2: a) Schéma d'un transistor MOS en silicium massif. b) Schéma d'un transistor MOS réalisé en technologie SOI. Les deux transistors représentés sont à enrichissement à canal d'électrons (NMOSFET).

Les paramètres importants du transistor sont les suivants :

- La tension sous le seuil (V_{TH}) est la tension de grille nécessaire à la formation d'un canal de conduction (couche d'inversion) entre la source et le drain. La tension de seuil typique d'un nanoMOSFET est comprise entre 0.2 et 0.4 V.

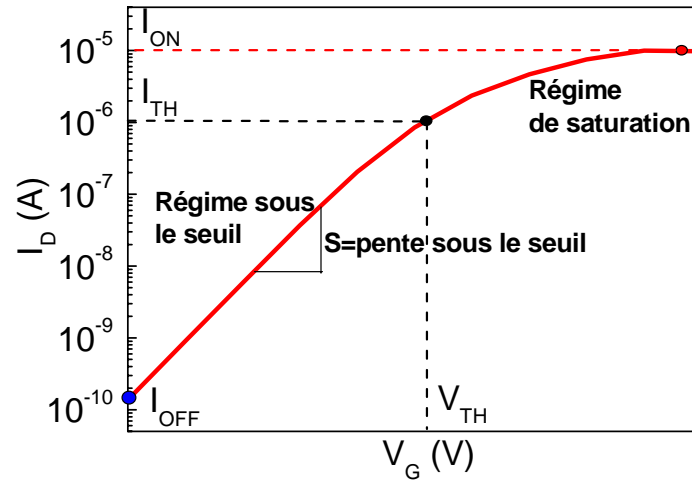


FIG. 1.3: Caractéristique I_D - V_G usuelle d'un transistor MOSFET. V_{TH} est la tension de seuil, I_{OFF} représente le courant du transistor dans l'état bloqué, et I_{ON} celui dans l'état passant. Attention à l'abus de langage : la pente sous le seuil S vaut en fait $\Delta V_G / \Delta I_D$ et est exprimée en mV/décade.

- La transconductance ou gain (g_m) est définie par :

$$g_m = \left(\frac{\partial I_D}{\partial V_G} \right)_{V_{DS}=cte}, \quad (1.1)$$

et doit être la plus élevée possible. La transconductance augmente très rapidement lorsque la longueur de canal devient inférieure à 100 nm.

- La conductance du canal (g_{DS}) donnée par :

$$g_{DS} = \left(\frac{\partial I_D}{\partial V_{DS}} \right)_{V_G=cte}. \quad (1.2)$$

- La pente sous le seuil : elle représente la tension de grille à appliquer (en régime sous le seuil) pour augmenter le courant de drain d'une décade :

$$S = \frac{dV_G}{d(\log I_D)}. \quad (1.3)$$

Un MOSFET est dans l'état "off" lorsque I_D est inférieur au dixième de sa valeur à la tension de seuil. Le paramètre S limite la réduction de la tension de seuil d'un nanoMOSFET puisque le courant à zéro volt, c'est-à-dire le courant de l'état "off" doit rester suffisamment faible pour ne pas consommer de puissance.

Il existe actuellement deux structures de composants largement étudiées et utilisées en technologie CMOS. La première est la structure silicium massif (ou "bulk"), dans laquelle le transistor est directement fabriqué sur un substrat semiconducteur (figure (1.2.a)). La seconde est appelée SOI (pour l'anglais "Silicon On Insulator") dans laquelle le transistor est réalisé sur une fine couche de silicium séparée du reste du substrat par une couche d'isolant (figure (1.2.b)). La structure bulk est relativement

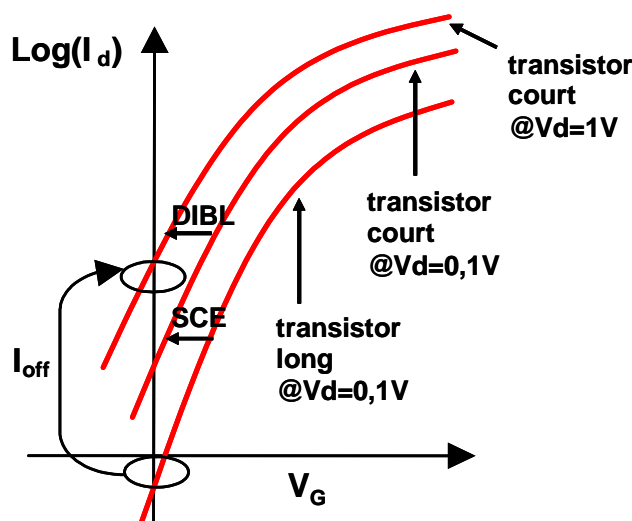


FIG. 1.4: Comparaison de la caractéristique sous le seuil d'un transistor long et d'un transistor court : mise en évidence des effets canaux courts.

simple à produire et constitue la brique de base de la plupart des circuits CMOS à ce jour.

Lorsque les dimensions du transistor diminuent, un équilibre doit s'établir entre deux critères : la performance du composant et sa fiabilité. La performance nécessite la suppression des effets canaux courts (SCE pour "Short Channel Effects" [6]). Ces derniers se traduisent tout d'abord par la réduction de la tension de seuil V_{TH} lorsque la longueur de grille décroît (figure (1.4)). La grille ne contrôle plus parfaitement les porteurs de charges du canal ; on parle de "roll-off" (ΔV_{TH}) de la tension de seuil. Un autre effet canaux courts, le DIBL (pour "Drain-Induced Barrier Lowering"), se manifeste par la réduction de la tension de seuil à fort V_{DS} [7]. Le *roll-off* dû aux effets canaux courts traduit l'abaissement de la barrière de potentiel du canal (figure(1.5)) qui engendre une dégradation de la pente sous le seuil (S) ainsi qu'une augmentation du I_{OFF} et donc de la puissance consommée au repos. Le DIBL quand à lui entraîne une dépendance de V_{TH} à la tension de drain et freine la réalisation de CMOS à l'échelle d'un circuit. Afin de retarder l'apparition de ces effets, les technologues disposent de plusieurs paramètres. Un dopage élevé du canal réduit tout d'abord l'épaisseur de la zone de déplétion à une région proche de la grille qui peut alors exercer avec plus d'efficacité son contrôle sur les charges en mouvement. En contre partie le libre parcours moyen des électrons dans le canal diminue et leur mobilité s'en trouve dégradée [8]. Une forte capacité d'oxyde C_{ox} améliore également le contrôle des charges ($Q = C_{ox} V_G$) et un oxyde mince à haute permittivité ϵ_{ox} ($C_{ox} = \epsilon_{ox}/T_{ox}$) facilite donc la diminution des dimensions. Néanmoins une réduction trop intense de l'épaisseur d'oxyde conduit à des courants de fuite de grille par courant tunnel [9] mettant à défaut la fiabilité. Afin d'atteindre une épaisseur inférieure au nanomètre, l'oxyde de silicium actuellement utilisé devra être remplacé par des matériaux à grande constante diélectrique, notamment par des oxydes binaires de métaux de transition comme le Zr, Hf, Y, et d'autres lanthanides [10]. Ces matériaux, aux structures électroniques complexes, sont intensivement étudiés, mais leur application

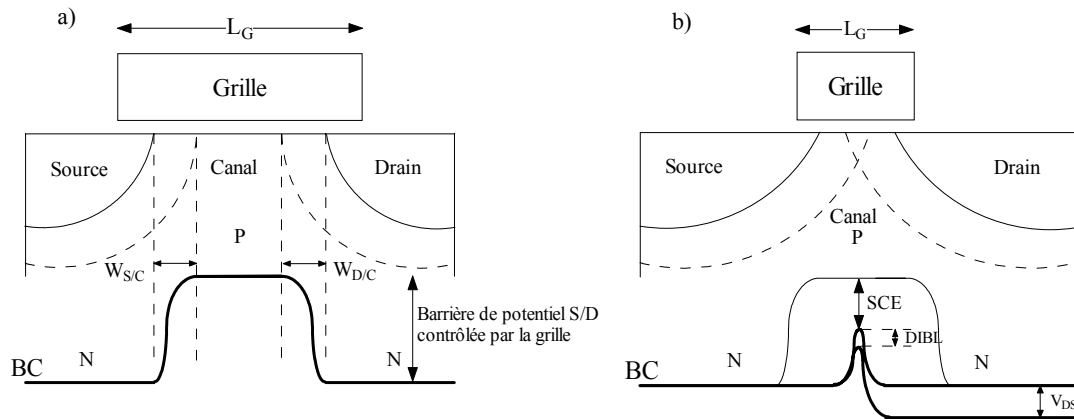


FIG. 1.5: Structure d'un transistor MOSFET bulk a) à canal long. b) à canal court. Dans le dernier cas, les effets canaux courts résultent du chevauchement des zones de charges d'espaces des régions dopées ($W_{S/C}$ et $W_{D/C}$).

définitive à la microélectronique ne sera pas immédiate. La réalisation de composants ultimes en silicium massif doit en conséquence faire intervenir des profils "compliqués" de dopage des régions de source, de canal, et de drain [11]. Des halos fortement dopés générés près des jonctions source/canal et drain/canal servent par exemple à écranter la zone active des champs électriques induits dans le drain (figure (1.6)). Lorsque la longueur de grille converge vers les dimensions ultimes la fusion de ces halos conduit à un nouveau dopage de canal très élevé qui permet ainsi de mieux résister à la dérive de V_{TH} . En conjuguant ce procédé avec d'autres technologies (dopage rétrograde notamment [12]), il a été démontré que la structure bulk pouvait poursuivre l'intégration jusqu'à des longueurs de grilles de 25 nm [11]. En-dessous de cette longueur, les composants bulk sont handicapés par de sévères fuites de la jonction p-n entre la source (ou le drain) et le substrat dues à la présence de forts champs électriques générés par l'implantation des halos.

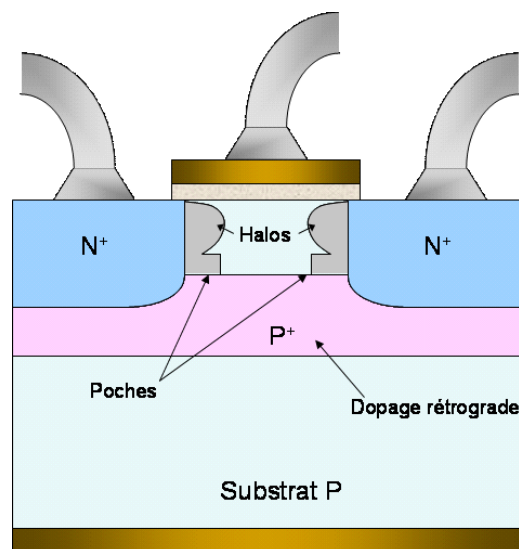


FIG. 1.6: Architecture d'un MOSFET à dopage rétrograde et halos.

1.3 Les architectures émergentes

1.3.1 Transistors à canal de silicium contraint

Une façon d'améliorer les performances du transistor MOS est d'introduire un matériau à haute mobilité au niveau du canal de la structure, de façon à augmenter notablement le courant I_{ON} : ce sont des transistors HEMT (pour "High Electron Mobility Transistor"). On peut donc soit changer la nature du semiconducteur dans la zone de transport (des canaux silicium-germanium ou silicium-carbone sont envisagés), soit utiliser un film de silicium contraint mécaniquement, ce qui présente l'avantage de conserver l'interface Si/SiO₂ (figure (1.7)).

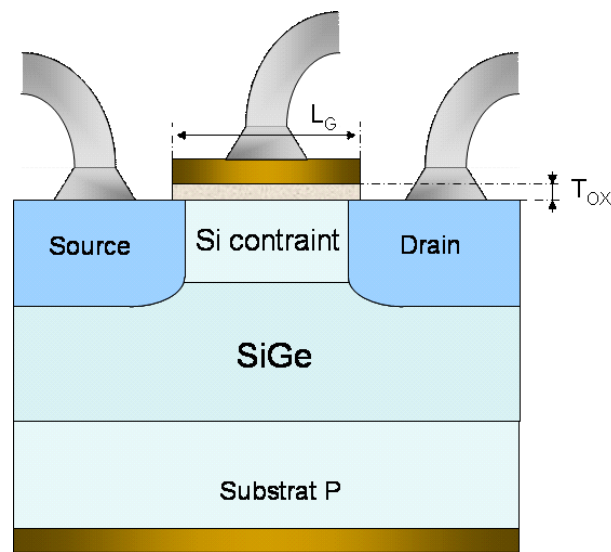


FIG. 1.7: Coupe schématique d'un HEMT à canal contraint.

La mise sous contrainte du film de silicium actif est obtenue en faisant croître celui-ci par épitaxie sur un substrat de silicium-germanium ($\text{Si}_{1-x}\text{Ge}_x$) relaxé. La différence de maille entre ces deux matériaux induit une contrainte bi-axiale dans le film de silicium, qui a pour effet de lever les dégénérescences dans la bande de conduction (vallées Δ) et dans la bande de valence (trous lourds). Ce "splitting" des niveaux d'énergie a pour conséquence une diminution des probabilités d'interaction entre les porteurs (électrons et trous) et les phonons intervalles. Dans le cas des électrons, il implique également une occupation préférentielle des deux vallées transverses de masse effective plus faible ($\sim 0.19 \times m_0$). La combinaison de ces deux effets explique donc les résultats expérimentaux (confirmés par des résultats théoriques) qui ont mis en évidence une augmentation de la mobilité des électrons jusqu'à 80% par rapport au silicium massif [13]. La situation est plus compliquée pour les trous, due à une structure de la bande de valence complexe et au fait que le "splitting" des niveaux d'énergie est moins marqué. Par ailleurs, l'application d'une contrainte fait varier les masses effectives des trous de façon différente, suivant que l'on considère des trous légers et des trous lourds. Enfin, l'hétérostructure Si/SiGe présente un offset de bandes défavorable au confinement des trous dans le film de silicium contraint, contrairement

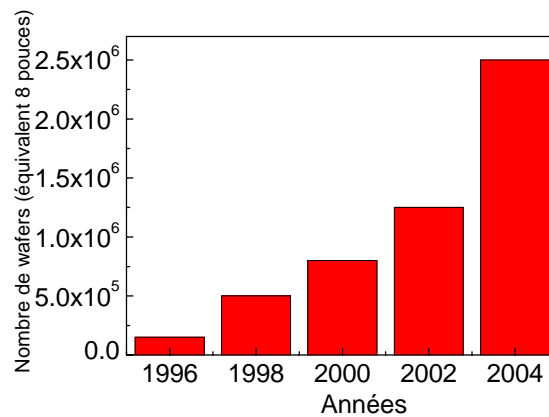


FIG. 1.8: *Estimation de la production annuelle de wafers SOI [17] [18].*

au cas des électrons bien localisés dans le puits de potentiel associé à la couche contrainte. Au final, le bénéfice du silicium contraint pour les transistors à canal p semble plus délicat à obtenir pour des transistors en silicium massif [14]. Pour les noeuds technologiques avancés, l'utilisation de silicium contraint sur isolant (SSOI pour Strained Silicon-On-Insulator), combinant l'apport d'une couche contrainte avec les avantages de la technologie SOI est envisagée et pourrait favoriser la réalisation de transistors contraints à canal p [15]. De nombreux travaux actuels portent sur la mise au point de procédés de fabrication de wafers SSOI et plusieurs industriels, tels IBM, envisageraient d'utiliser une telle technologie dès le noeud 65 nm en 2005 [16].

1.3.2 Les transistors SOI

Les structures SOI MOSFET (partiellement ou complètement déplétées) sont envisagées comme des candidats susceptibles de réussir l'intégration de composants de dimensions inférieures à 25 nm. La quantité estimée de wafers SOI produits depuis 8 ans est d'ailleurs en très nette augmentation (figure (1.8)) et l'utilisation du SOI, malgré le prix plus élevé des substrats, semble gagner de plus en plus d'industriels. Les avantages de la technologie SOI sont multiples. Premièrement le canal de conduction est mince et entièrement isolé. La profondeur des jonctions est parfaitement définie et limitée par l'épaisseur du canal de silicium (figure 1.9.b). L'aire des jonctions en SOI est considérablement réduite (les jonctions sont verticales) ce qui conduit à de faibles courants de fuite. De plus les effets parasites qui existent dans la technologie Si massif sont éliminés. Le plus fréquent est le thyristor parasite ou effet de "*latch-up*", qui consiste au déclenchement du thyristor parasite représenté sur la figure (1.9.a) par deux transistors bipolaires. Le déclenchement du *latch-up* est lié à la proximité des zones N^+ et P^+ de transistors NMOS et PMOS voisins. Sur le SOI, ces zones sont complètement isolées et le claquage entre ces deux zones est impossible quelle que soit la distance qui les sépare.

La faculté d'augmenter la densité d'intégration est une conséquence directe de la suppression de l'effet *latch-up* qui permet le rapprochement des régions N^+ et P^+ . L'absence de caisson conduit aussi à une simplification des niveaux d'interconnexions au-dessus des transistors.

L'oxyde enterré réduit ensuite les capacités parasites des régions source et drain (figure (1.9.b)). Les capacités source/substrat (C_{JS}) et drain/substrat (C_{JD}) dans la technologie SOI sont 4 à 7 fois plus faibles que sur silicium massif. Tandis que sur silicium massif C_{JS} (resp. C_{JD}) équivaut à la capacité d'une jonction polarisée en inverse, dans le SOI C_{JS} (resp. C_{JD}) est dominée par la capacité de l'oxyde enterré sous la source (resp. le drain), qui est beaucoup plus faible. La réduction des capacités parasites se répercute au niveau du circuit et améliore les performances en hautes fréquences des SOI MOSFETs par rapport au silicium massif [19].

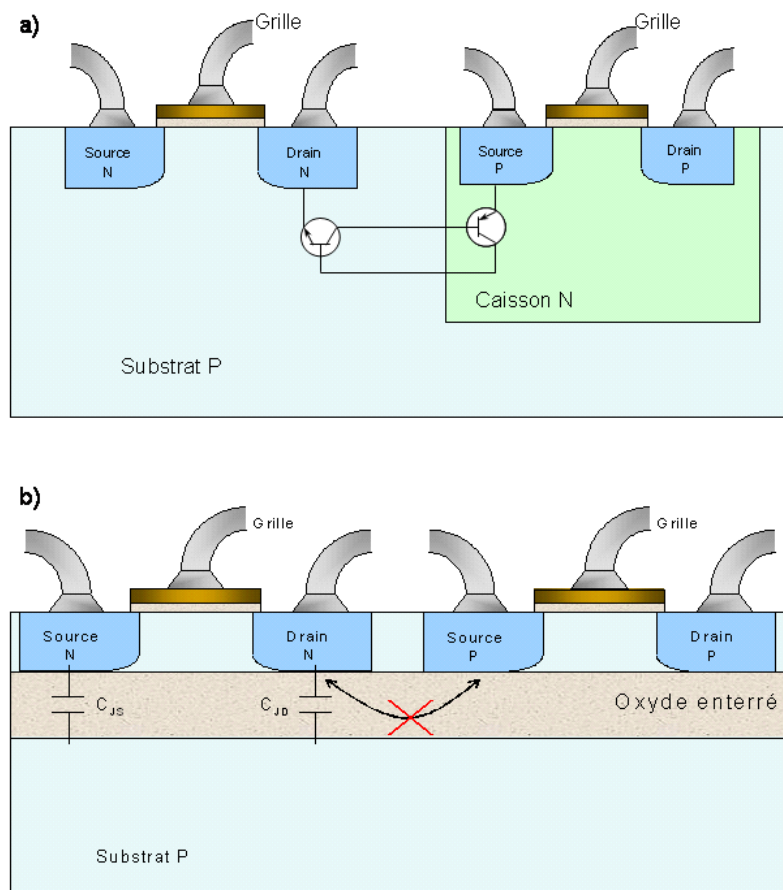


FIG. 1.9: Comparaison d'une structure CMOS réalisée (a) sur Si massif; (b) sur SOI. Illustration de l'élimination du phénomène de latch-up dans le CMOS sur SOI.

1.3.3 Les transistors multi-grilles

Cependant, les MOSFET SOI simple grille présentent un oxyde enterré très épais ($10 < T_{BOX} < 100$ nm) dans lequel pénètrent les lignes équi-potentielles de champ électrique de la source et du drain. Il s'en suit une perte de potentiel dans l'oxyde enterré et un mauvais contrôle électrostatique de la grille pour les petites dimensions : les effets canaux courts apparaissent et dégradent à nouveau la pente sous le seuil [21]. Depuis une dizaine d'années, la plupart des études semblent indiquer que les transistors SOI MOSFET à deux, trois, voir quatre grilles sont plus adaptés à la réduction ultime des dimensions. La multiplication des grilles contrôle mieux

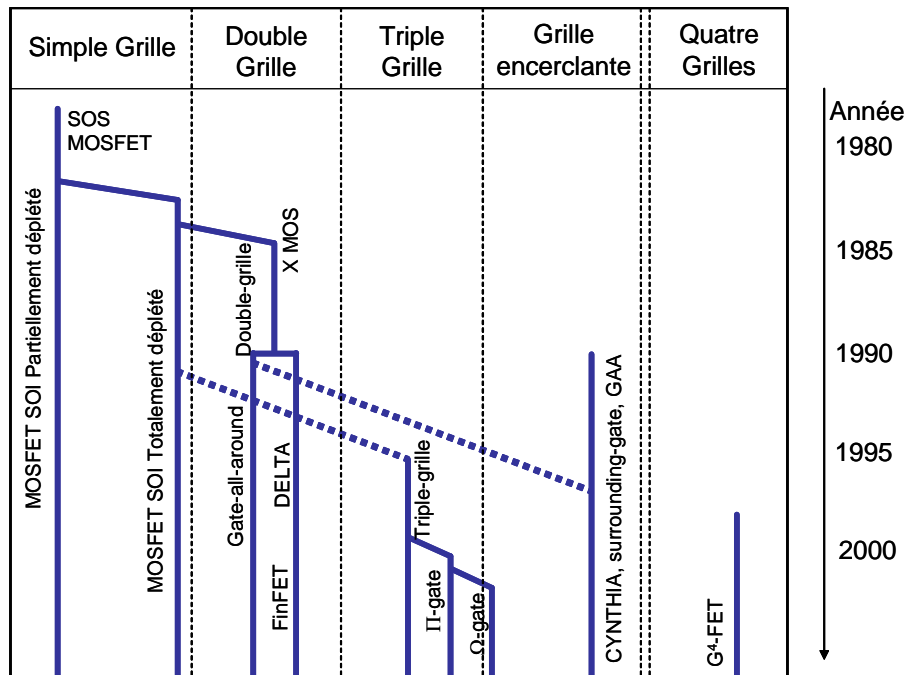


FIG. 1.10: Arbre généalogique de la technologie SOI.

le potentiel de canal et immunise le transistor contre les effets canaux courts observés dans une géométrie SOI simple grille [22, 23]. Pour ces géométries, un fort dopage de canal n'est plus nécessaire. L'utilisation d'une partie active en silicium ultra-fine et faiblement dopée stabilise les variations de la tension de seuil V_{TH} [24], améliore la mobilité du canal (diminution de la diffusion avec les impuretés ionisées [8]) et fait tendre la pente sous le seuil S vers la valeur idéale de 60 mV/décade.

Nous allons évoquer dans ce paragraphe les principales familles de transistors MOSFETs multi-grilles susceptibles de succéder au MOSFET *bulk*. Observons pour cela l'arbre généalogique des MOSFETs SOI représenté sur la figure (1.10). Le premier transistor SOI date de 1964. Ce composant était partiellement déplété et fabriqué sur un substrat en saphir (d'où l'appellation SOS pour Silicon-On-Sapphire). La première publication décrivant un MOSFET SOI double-grille remonte à 1984. Le composant reçu l'acronyme de XMOS à cause de sa ressemblance avec la lettre grecque Ξ [25]. Déjà cet article prédisait les bons comportements à faibles longueurs de grille d'un tel composant. Vint ensuite le MOSFET double-grille totalement déplété dont la première réalisation fut le DELTA MOSFET pour "*fully DEpleted Lean channel TrAnsistor*" [26]. A ce moment, le MOSFET double-grille, qui comprend un film horizontal de silicium "sandwiché" entre deux grilles, présente une amélioration spectaculaire du contrôle électrostatique des charges du canal. En revanche le non-alignement des grilles avant et arrière génère une capacité de recouvrement grille/S-D et des résistances séries [27].

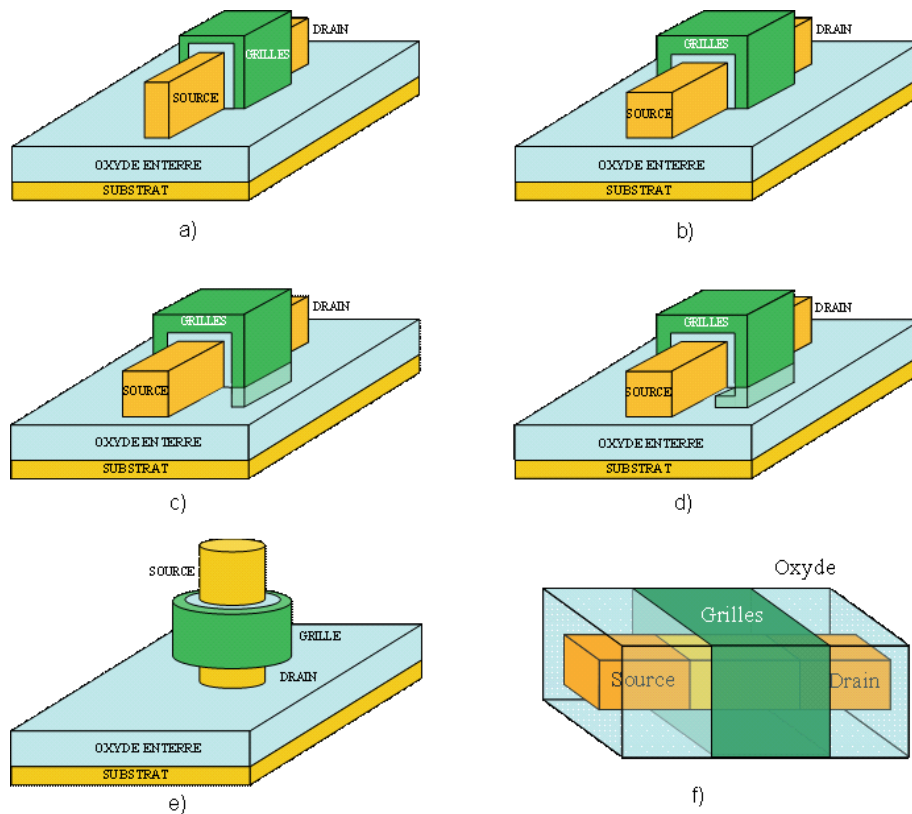


FIG. 1.11: *Différentes configurations de grilles pour des composants SOI : a) FinFET ; b) Triple grille ; c) Pi-gate MOSFET ; d) Oméga-gate MOSFET ; e) Surrounding gate ; f) Gate-All-Around MOSFET (structure GAA).*

Le FinFET (figure (1.11)), structure auto-alignée, est alors proposé comme une amélioration du transistor double-grille [28, 29]. Il est caractérisé par un fil de silicium déposé sur un oxyde enterré et entouré de trois grilles. De par l'épaisseur réduite de sa partie silicium, seules les deux grilles verticales exercent un véritable contrôle sur le canal. Sa relative simplicité de fabrication et sa compatibilité avec la technologie CMOS conventionnelle en font un candidat prometteur. En augmentant l'épaisseur du film de silicium, le FinFET devient le triple-grille (ou tri-gate) [30]. Les trois grilles contrôlent maintenant le transport des charges et réduisent toujours davantage les effets canaux courts. Cette architecture permet de relaxer les dimensions transverses (et donc de diminuer le coût de fabrication) tout en améliorant les performances électriques : l'augmentation de la section transversale du silicium diminue non-seulement la sensibilité à la variation d'épaisseur, mais réduit les résistances en série. Cependant l'oxyde enterré sur lequel est fabriqué le transistor constitue toujours un obstacle à l'optimisation des propriétés électriques. Des versions améliorées de la précédente structure sont les pseudo-quatre grilles comme le Pi-gate [31] et le Omega-gate [32]. Ces composants simulent, sans ajouter d'étapes technologiques, une quatrième grille par la pénétration des grilles verticales dans l'oxyde enterré. L'extension de grille de ces transistors ne sert pas seulement d'armature pour les lignes de champ électrique provenant du drain, mais elle améliore aussi le contrôle de la grille sur le canal. La structure ultime qui offre théoriquement le meilleur contrôle de la région de canal est le quadruple-grilles ou "surrounding-gate" transistor.

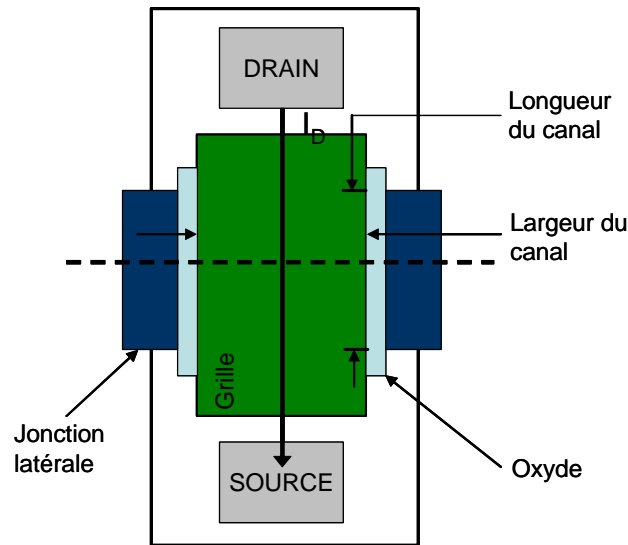


FIG. 1.12: Vue de dessus d'une structure G^4 -FET.

Deux architectures existent : le "surrounding-gate" à canal vertical, habituellement fabriqué en utilisant un îlot de silicium en forme de colonne [33], et le "surrounding-gate" à section rectangulaire aussi appelé Gate-All-Around (GAA) [34,35]. Enfin, le G^4 -MOSFET [36,37] est une structure plus exotique constituant un nouveau type de transistors à accumulation. Il combine deux grilles MOS verticales et 2 grilles à jonctions latérales afin d'encercler le canal (figure(1.12)). Des études restent encore à mener pour déterminer tous les avantages de ce type de composant.

| Références | Structure | L_G | I_{ON} | I_{OFF} |
|--|-----------|--------------|--------------------|------------------------|
| J. Kedzierski <i>et al.</i> (IBM) [40] | FinFET | $L_G=100$ nm | $1200 \mu A/\mu m$ | $<1 \mu A/\mu m$ |
| B. Yu <i>et al.</i> (AMD) [46] | FinFET | $L_G=10$ nm | $446 \mu A/\mu m$ | $100 \mu A/\mu m$ |
| F.-L. Yang <i>et al.</i> (TSMC) [47] | Oméga FET | $L_G=25$ nm | $1440 \mu A/\mu m$ | $8 \text{ nA}/\mu m$ |
| B.H. Lee <i>et al.</i> (IBM) [48] | SOI-FET | $L_G=60$ nm | $1000 \mu A/\mu m$ | $200 \text{ nA}/\mu m$ |
| S. Harrison <i>et al.</i> (ST) [49] | SON-GAA | $L_G=70$ nm | $1950 \mu A/\mu m$ | $254 \text{ nA}/\mu m$ |

TAB 1.2 : Publications récentes de composants ultimes réalisés en technologie SOI simple et multi-grilles (Les courants sont ceux de NMOS).

Plusieurs inconvénients restent caractéristiques de ces nouvelles architectures. De par la finesse du fil de silicium, les résistances des extensions source et drain peuvent être très élevées, limitant ainsi les performances du transistor. Des solutions ont été proposées, notamment en remplaçant les régions dopées par des contacts Schottky [38] ou des couches d'inversion [39]. De plus, l'utilisation de canal intrinsèque rend difficile l'ajustement de la tension de seuil V_{TH} . Ici aussi, plusieurs antidotes existent : varier le dopage des grilles en polysilicium [24], employer de nouveaux métaux [40] ou appliquer un offset de tension entre les différentes grilles de façon à mimer plusieurs travaux de sorties [41] sont des techniques envisageables. Les fluctuations des caractéristiques de courant induites par la présence, volontaire ou non de dopants constituent un autre problème associé aux petites géométries [42,43]. S'invitent enfin les effets quantiques (dûs

au confinement latéral), qui peuvent engendrer une sensibilité accrue de V_{TH} à l'épaisseur de la partie active [44, 45]. Les fluctuations de l'épaisseur de silicium doivent donc être rigoureusement contrôlées. Malgré ces difficultés, le potentiel d'intégration de structures multi-grilles ne doit pas être sous-estimé. D'ailleurs de nombreux transistors multi-grilles ont été réalisés et présentent des caractéristiques électriques bien supérieures à celles des transistors en silicium massif (voir tableau(1.2)).

1.4 Les transistors à nanotube de carbone

Depuis quelques années, les nanotechnologies sont pressenties pour réussir le développement de techniques de fabrication dites "*Bottom-up*" peu coûteuses par opposition à la miniaturisation des transistors MOSFETs classiques de type "*Top-down*". Les nanotechnologies basées sur les nanotubes de carbone, du fait du très bon comportement électrique de ces matériaux, semblent être l'illustration actuelle la plus frappante de cette approche. Nous allons décrire dans les prochains paragraphes les propriétés physiques des nanotubes avant de présenter l'architecture des transistors à nanotube de carbone (CNTFET pour "Carbon NanoTube Field-Effect-Transistor").

1.4.1 Nanostructures de carbone et nanotubes

A l'état solide le carbone peut exister sous trois formes cristallines différentes nommées allotropes. Le diamant et le graphite sont les deux allotropes les plus connus du carbone. Le troisième allotrope du carbone regroupe les structures de type fullerènes. Il s'agit d'une nouvelle classe de matériaux formés par l'enroulement de plans graphitiques en sphère (fullerènes de type C_{60} [50] ou C_{70} semblables à un ballon de foot-ball) ou en cylindre (nanotubes).

Les différentes structures de nanotubes

On en distingue deux types :

- Les nanotubes mono-parois (SWNT pour "Single Walled NanoTubes") sont constitués d'un seul feuillet graphitique. Le nanotube possède une hélicité qui dépend de la façon dont les deux bords de la feuille de graphène vont se recouper. Chaque tube pourra être désigné par deux entiers (m,n) , valeurs dans la base unitaire du réseau hexagonal des composantes du vecteur reliant un même point le long de la circonférence [51]. Nous obtenons ainsi trois grandes catégories de nanotubes suivant le repliement du plan graphitique : ceux avec la structure "*armchair*" (m,m) , dans laquelle certaines liaisons C-C sont perpendiculaires à l'axe du nanotube ; ceux avec la structure "*zig-zag*" $(n,0)$ dans laquelle certaines liaisons C-C sont parallèles à l'axe ; et ceux avec une structure appelée chirale (m,n) présentant une symétrie miroir (figures(1.13) et (1.14)).

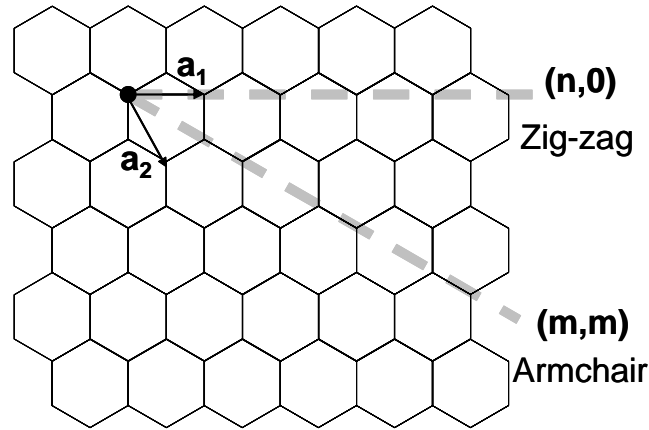


FIG. 1.13: Deux configurations d'enroulement d'un feuillet graphitique : zig-zag et armchair..

Les valeurs de m et n déterminent la chiralité du nanotube, c'est-à-dire la façon dont le plan de graphite est enroulé sur lui-même. La chiralité affecte la plupart des propriétés du nanotube : sa conduction, sa densité et sa taille. Par exemple, les valeurs m et n permettent de calculer le diamètre d du tube à l'aide de l'équation suivante :

$$d = \frac{0.142 \text{ nm}}{\pi} \sqrt{3(n^2 + m^2 + nm)}. \quad (1.4)$$

- Lorsque deux parois ou plus forment le tube, on parle de nanotube multi-parois (MWNT pour "Multi-Walled NanoTube"). L'espacement entre les parois d'une telle structure est d'environ 0.34 nm (figure (1.15)).

Il existe plusieurs méthodes servant à produire les nanotubes de carbone. La première méthode, dite de l'"arc électrique", date de 1991, année où SUMIO IJIMA¹ (NEC Fundamental Research Laboratory) fit la découverte des nanotubes de carbone multi-parois. Deux ans plus tard, il réussit à produire des nanotubes mono-paroi en modifiant légèrement sa méthode originale. Par la suite, de nombreux groupes développèrent de nouvelles techniques. Le tableau (1.3) présente quatre différents procédés de fabrication de nanotubes mono-paroi.

Cependant ces méthodes ne permettent pas l'obtention d'un matériau uniquement composé de nanotubes de carbone. Du carbone amorphe et des résidus de métaux de transition sont aussi présents dans les échantillons. De plus, il n'existe pas encore de méthode permettant de contrôler la chiralité des tubes produits. Cette limitation majeure ralentit considérablement le rythme des avancées dans le domaine des dispositifs électroniques à bases de nanotubes car il est impossible de produire uniquement des tubes métalliques ou uniquement semi-conducteurs.

¹Sumio Iijima a été le premier à observer un nanotube, mais certainement pas le premier à en produire. En fait, l'Homme de Néandertal a fabriqué des quantités minuscules de nanotubes dans le feu qui chauffait ses grottes. En effet, dissociés les uns des autres par la chaleur, les atomes de carbone se recombinaient aléatoirement dans la suie. Certains se regroupent en amas amorphes, d'autres forment des sphères ou des tubes.

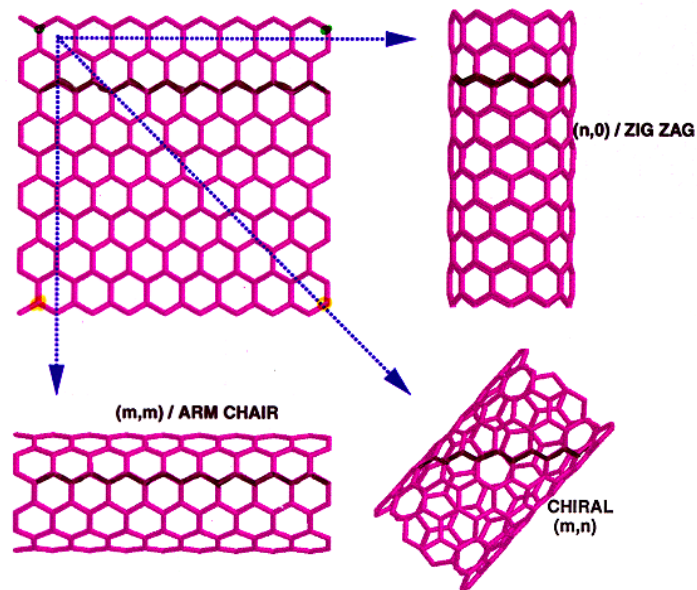


FIG. 1.14: Les trois structures d'un nanotube de carbone mono-paroi. Deux d'entre elles correspondent à des nanotubes à haute symétrie (zig-zag et armchair) alors que la troisième (chirale) présente une structure vis (droite ou gauche) comme l'ADN.

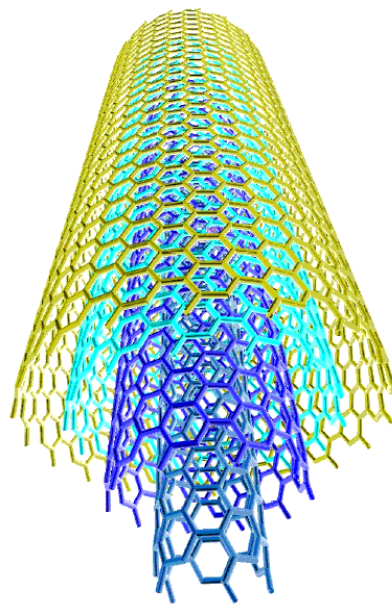


FIG. 1.15: Exemple d'une structure de nanotube multi-parois.

| Procédé | Description | Référence |
|-----------------------------|--|------------------------------|
| Arc électrique | Procédé catalytique basé sur la décharge par arc électrique entre deux électrodes de graphite ayant préalablement été contaminées par des métaux tels que Fe, Ni et Co. | Iijima [52, 53] |
| Ablation laser | Procédé catalytique utilisant des impulsions laser afin de sublimer un substrat de graphite. Le catalyseur (2% atomique) est intégré au substrat sous forme de clusters métalliques. | Thess <i>et al.</i> [54] |
| HiPco | Procédé catalytique en phase gazeuse et sous haute pression (10 à 100 atm). Le mélange gazeux est formé de CO et de Fe(CO) ₅ , le fer agissant comme catalyseur. | Nikolaev <i>et al.</i> [55] |
| Torche plasma atmosphérique | Procédé catalytique basé sur l'atomisation d'un mélange gazeux dans une torche plasma sous pression atmosphérique. Le mélange gazeux est formé de Ar, Fe(C ₅ H ₅) ₂ et C ₂ H ₄ . | Smiljanic <i>et al.</i> [56] |

TAB. 1.3 : Procédés de fabrication des nanotubes de carbone mono-paroi.

Propriétés optiques

Les propriétés optiques d'un nanotube sont très variées et peuvent s'étendre de celles d'un semiconducteur à celles d'un métal suivant son diamètre et sa chiralité.

HAMADA *et al.* [57] ont montré que les nanotubes de type $|n-m|=3l$, où l est un entier positif ou nul, étaient métalliques et que tous les autres tubes se comportaient comme des semiconducteurs. Par exemple, un nanotube de type *zig-zag* défini par le doublet (12,0) présente un gap de 8 meV alors que le nanotube (13,0) a un gap de 0.697 eV. Dans le cas de nanotubes semiconducteurs, la valeur du gap dépend du diamètre mais reste proche de 0.5 eV. WILDER *et al.* [58] ont proposé une formule analytique du gap d'un nanotube semiconducteur :

$$E_{gap} = 2y_0 a_{cc} / d, \quad (1.5)$$

où y_0 est l'énergie de recouvrement C-C ($\simeq 2.5$ eV), a_{cc} est la distance C-C entre deux plus proches voisins (0.142 nm), et d est le diamètre en nm. On trouve ainsi que les énergies sont comprises entre 0.4 eV et 0.7 eV et que la gap diminue quand le diamètre augmente.

Le changement abrupt des propriétés optiques en fonction de la chiralité reflète bien la complexité unique de ce matériau. Ce comportement est en partie dû à sa petite taille et aux effets quantiques qui régissent à cette échelle l'ensemble des propriétés physiques. Mais le graphite est lui aussi un matériau peu commun. Alors que la plupart des conducteurs électriques sont

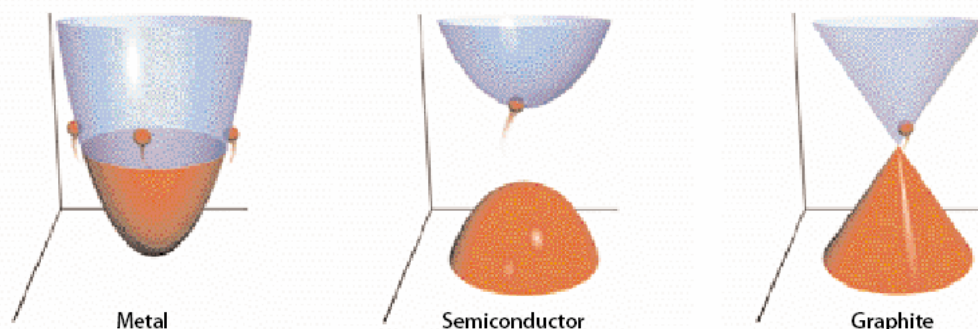


FIG. 1.16: Les propriétés électriques et optiques d'un matériau dépendent de l'écart entre l'énergie des états remplis (bande de valence en rouge) et celle des premiers états vides, disponibles pour accueillir un électron (bande de conduction en bleu). L'écart d'énergie dans les matériaux métalliques est nul alors que dans les semiconducteurs les électrons ont besoin d'un apport d'énergie extérieur (lumière ou champ électrique) pour atteindre les états de la bande de conduction. Le graphite se situe entre ces deux extrêmes car un faible apport d'énergie externe (température ambiante) suffit à exciter un électron d'une bande d'énergie vers la seconde [59].

soit métalliques soit semiconducteurs, le graphite est un des rares matériaux connus à être un semi-métal (figure (1.16)). En combinant les propriétés semi-métalliques du graphite avec les lois quantiques des niveaux d'énergie discrétisés, les nanotubes de carbone se présentent comme des conducteurs vraiment exotiques. Plus concrètement, la structure de bandes du nanotube est très proche de celle du graphite. Les conditions aux limites représentent la différence essentielle. En effet, l'opération d'enroulement impose des conditions aux limites périodiques pour décrire la propagation des électrons autour de la circonférence, qui peuvent interférer avec eux-mêmes constructivement ou destructivement. Les fonctions d'onde des électrons ne sont donc plus celles d'un système 2D infini, mais deviennent des ondes de Bloch aux vecteurs d'onde discrets. La figure (1.17) illustre la situation à l'aide de la première Zone de Brillouin (ZB) d'un nanotube de type *zig-zag* (6,0). Les lignes verticales correspondent aux vecteurs d'onde permis. Pour un nanotube de type $(n,0)$, les lignes verticales interceptent la ligne Γ -M dans la première ZB en n points. De plus, certaines bandes d'énergie du plan graphitique (π liantes et anti-liantes) sont dégénérées au point K. En conséquence si une ligne verticale passe par le point K, c'est-à-dire si n est un multiple de 3, le tube sera un métal. Un raisonnement similaire dans le cas de tubes *armchairs* et *chiraux* conduisent à la loi précédemment énoncée.

Aucun matériau connu à ce jour ne présente des propriétés optiques aussi modulables !

Propriétés électriques

Les propriétés de transport des nanotubes ont été récemment étudiées et ont soulevé quelques controverses. En 1998, FRANK *et al.* [60] ont mesuré la conductance de MWNT. Fixant des nanotubes à la place de la pointe d'un microscope SPM (pour "Scanning Probe Microscope"), ils contactèrent les fibres de nanotubes avec une surface de mercure. A faible polarisa-

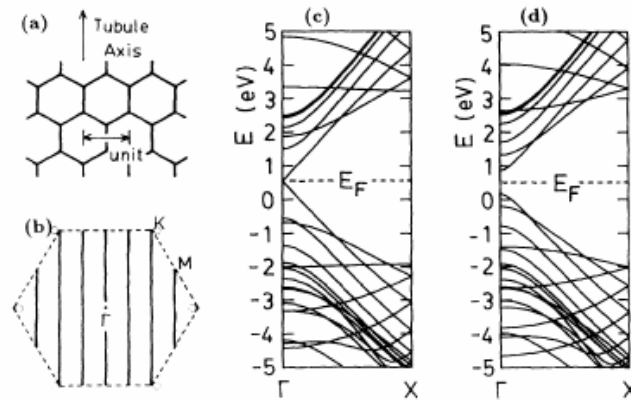


FIG. 1.17: (a) Configuration géométrique d'un tube $(n,0)$. (b) Première zone de Brillouin d'un plan graphitique (trait en pointillés), et vecteurs d'onde permis par la condition périodique le long de la circonférence pour $n=6$ (lignes continues). Structure de bande d'un nanotube (c) $(12,0)$ et (d) $(13,0)$. [57]

tion, leur résultats démontrèrent un comportement balistique avec une conductance quantifiée en $G_0 = 2e^2/h = (12.9 \text{ k}\Omega)^{-1}$ (le quantum de conductance [61], voir §3.3.1). Le coefficient du quantum de conductance prenait cependant des valeurs non-entières ($\simeq 0.5G_0$). En 1999, SANVITO *et al.* [62] ont démontré théoriquement que les valeurs inattendues de la conductance des MWNT mesurées par FRANK *et al.* provenaient du blocage de certains canaux de conductance par des réactions inter-couches. Comme nous le verrons dans le chapitre 3 la quantification de la conductance résulte des propriétés de guide d'onde électronique des fils extrêmement fins. Quand la longueur du conducteur est plus petite que le libre parcours moyen de l'électron, le transport devient balistique, auquel cas chaque mode du guide d'onde (ou canal conducteur) contribue avec G_0 à la conductance totale. La mesure de densités de courant des nanotubes est le second résultat apporté par FRANK *et al.*. Pour la première fois, les courants très élevés prévus par la théorie ont été vérifiés avec des densités supérieures à 10^7 A. cm^{-2} (pour comparaison, la densité de courant des supraconducteurs est typiquement de l'ordre de 10^5 A. cm^{-2}). Néanmoins, le MWNT ne constitue pas à ce jour un matériau idéal pour réaliser des transistors. Sa topologie en différentes sous-couches dans laquelle chaque tube possède un diamètre et une hélicité spécifique rend les propriétés électriques difficilement prévisibles. On lui préfère jusqu'à présent les SWNT qui, d'après PHAEDON AVOURIS (IBM T.J. Watson Research Center), pourraient atteindre des densités supérieures à $10^{13} \text{ A. cm}^{-2}$. Ces propriétés exceptionnelles de conduction ne sont possibles que grâce à la solidité des liaisons covalentes C-C (équivalentes à celle des liaisons covalentes du diamant). Notons que de tels courants vaporiseraient l'or ou le cuivre.

Nous allons dans le prochain paragraphe décrire les mécanismes du transport électronique intervenant dans un SWNT. Il a été constaté que l'augmentation de la conductance avec la tension appliquée n'est pas proportionnelle au nombre de sous-bandes qui participent au transport électrique. Une explication possible de ce comportement a été apportée par ANANTRAM [63, 64]. Dans le cas de contacts parfaits (sans réflexion), un électron injecté dans un nanotube peut subir trois processus (figure (1.18)) :

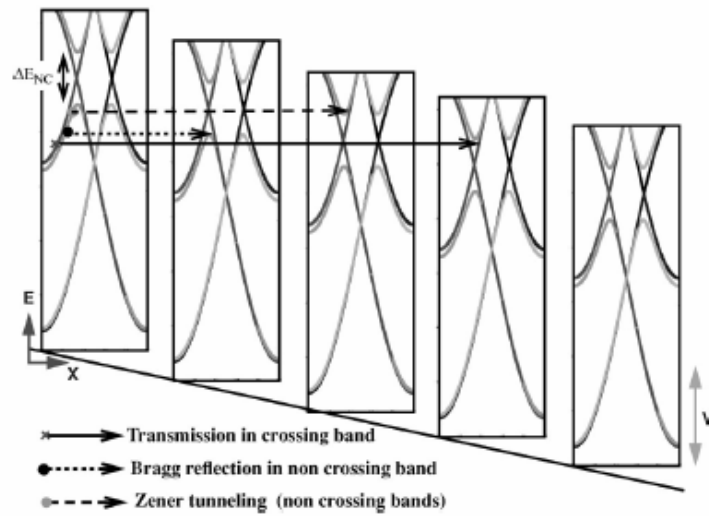


FIG. 1.18: Chaque boîte rectangulaire représente l'énergie en fonction du vecteur d'onde avec le bas de sous-bandes égal au potentiel électrostatique. Dans un but de clarté, seules quelques sous-bandes sont représentées. Les processus montrés sont la transmission directe (ligne continue), la réflexion de Bragg (ligne en pointillés) et le tunnel Zener inter sous-bande (ligne en tirets). [63]

- la transmission directe : l'électron est transmis à travers la sous-bande dans laquelle il a été injecté depuis le contact de gauche.

- la réflexion de Bragg : ces réflexions se manifestent quand le vecteur d'onde (\mathbf{k}) de l'électron injecté évolue vers une valeur pour laquelle la vitesse dans la sous-bande est nulle (rappel : $v(\mathbf{k}) = (1/\hbar) * (\partial E / \partial \mathbf{k})$).

- le tunnel Zener inter sous-bandes : il s'agit d'un effet tunnel entre sous-bandes qui apparaît pour de forts champs électriques extérieurs. L'espace entre deux sous-bandes qui ne se croisent pas ($= \Delta E_{NC}$ sur la figure (1.18)) décroît avec l'augmentation du diamètre du nanotube. Donc le tunnel Zener s'intensifie pour de grandes valeurs de diamètre du nanotube. L'importance relative de ces trois phénomènes dépend de l'énergie (et donc la tension appliquée), du profil de potentiel et du diamètre du nanotube. Considérant une évolution linéaire du potentiel le long du nanotube, l'application d'une tension faible ($< \Delta E_{NC} / e$) conduit à l'injection d'électrons dans les deux sous-bandes qui se croisent et pour lesquelles la probabilité de transmission est très élevée (≈ 1). Des électrons sont également injectés dans les bandes qui ne se croisent pas. Ils peuvent en principe contribuer au courant si les états finaux dans lesquels ils doivent être transmis sont disponibles. En l'absence de diffusion inter-sous-bande, ils subissent néanmoins majoritairement des réflexions de Bragg et les caractéristiques de courant sont principalement déterminées par la transmission directe. La conductance totale G est dans ce cas seulement quantifiée par deux canaux conducteurs : $G = 4e^2/h$. Lorsque la tension appliquée est supérieure à $\Delta E_{NC} / e$, les électrons de la première sous-bande au-dessous du centre de bande du contact gauche peuvent "tunneler" vers des états de la première sous-bande au-dessus du centre de bande du contact droit par tunnel Zener. La conductance totale dépasse alors $4e^2/h$.

1.4.2 Les transistors à nanotubes de carbone

Comme nous l'avons expliqué, les applications pour l'électronique moléculaire concernent principalement les SWNT. La structure retenue pour réaliser des transistors à nanotubes de carbone (CNTFET pour "Carbon NanoTube Field-Effect-Transistor") est actuellement la suivante : métal/nanotube semiconducteur/métal. En effet, le comportement des jonctions p-n des nanotubes diffère fortement de celui des jonctions dans les semi-conducteurs classiques. Le dopage des nanotubes est néanmoins connu et se réalise deux façons :

- Dopage par des impuretés substitutionnelles (typiquement B et N [65])
- Dopage par insertion d'atomes alcalins ou halogènes à l'intérieur du tube [66].

Deux longueurs caractérisent une jonction p-n dans un nanotube. La largeur de la région déplétée et la diminution logarithmique de la queue coulombienne due à l'écrantage des charges (qui peut s'étendre sur plusieurs microns). Un dopage non-intentionnel peut donc avoir beaucoup d'influence sur les caractéristiques électriques du nanotube et une précision atomique est nécessaire pour espérer réaliser des contacts de bonne qualité : un grand défi qui pour l'instant reste difficile à relever.

Les CNTFETs se schématisent donc par deux barrières Schottky connectées par un nanotube. APPENZELLER *et al.* ont vérifié expérimentalement que les propriétés du nanotube ne limitent pas les caractéristiques du transistor [67]. En revanche, les barrières Schottky des interfaces nanotube/métal, et leur réponse aux champs appliqués déterminent leurs performances électriques. Un CNTFET fonctionne donc en modulant le courant tunnel de la barrière Schottky de la source. Pour une valeur donnée de la tension de drain V_{DS} , l'augmentation du champ de grille conduit à une augmentation exponentielle du courant dans le nanotube. C'est la conséquence de la croissance exponentielle de la transmission à travers une barrière Schottky lorsque son épaisseur diminue. En se rapprochant de la tension de seuil, les charges s'accumulent dans le tube et l'évolution des bandes d'énergie avec la tension de grille ralentit. L'épaisseur de la barrière Schottky du contact de source ne varie plus de façon aussi significative que dans le régime sous le seuil, et l'augmentation du courant tunnel s'atténue (Figure (1.19)). Ainsi, bien que les caractéristiques des MOSFETs conventionnels et celles des CNTFETs soient assez proches, la physique est bien différente. Ces composants ne sont pas simplement une version miniature des transistors traditionnels. En effet la longueur du canal a , en première approximation, peu d'influence sur les propriétés électriques du composant (pente sous le seuil par exemple) et l'allure des courbes de courant d'un CNTFET reflète principalement la réponse des barrières Schottky de la source et du drain au champ de grille. Dans ce contexte, l'épaisseur de l'oxyde de grille joue un rôle essentiel, puisqu'elle contrôle la longueur de la barrière Schottky [68, 69]. Les premiers CNTFETs ont été fabriqués en déposant un nanotube entre deux contacts et présentaient une grille arrière dont l'épaisseur d'oxyde atteignait plusieurs dizaines de nanomètres [70–73] (figure (1.20)). Le faible couplage grille/nanotube nécessitait une forte tension de grille ($>5V$) pour rendre le transistor passant. L'autre inconvénient de la géométrie "ouverte" du CNTFET est

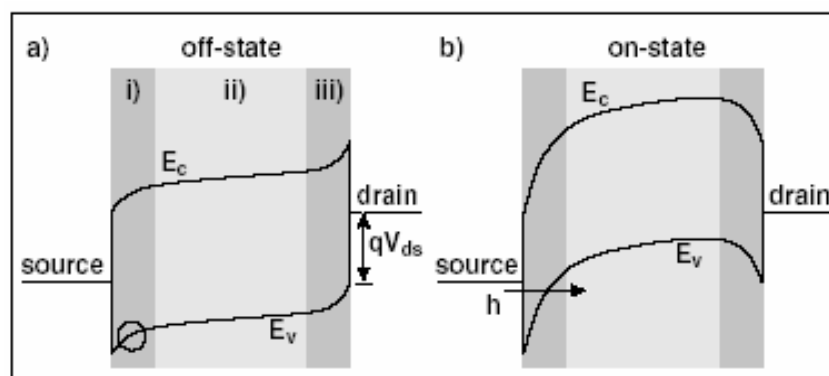


FIG. 1.19: Réponse qualitative de la bande de conduction et de valence à la tension de grille pour une tension source-drain négative. (a) Description du transistor dans l'état "off". (b) Evolution pour une tension de grille supérieure à la tension de seuil [67].

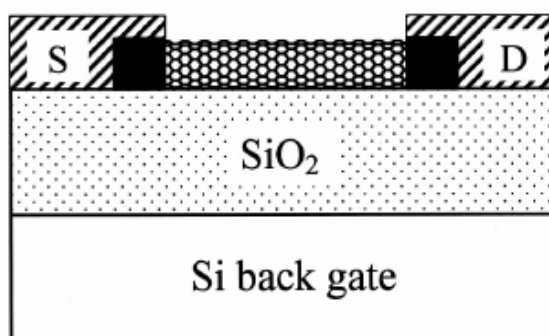


FIG. 1.20: Représentation schématique d'un transistor à nanotube de carbone avec grille arrière [72].

que le contact avec l'air conduit à des caractéristiques de type-p (due à l'adsorption d'oxygène par le réseau graphitique). Plus récemment, WIND *et al.* [74] ont présenté une géométrie à grille avant ("top-gate") similaire à celle des MOSFETs conventionnels, avec une grille au-dessus du canal et séparée de ce dernier par une fine couche d'oxyde de silicium (15-20 nm). Le nanotube est totalement incorporé dans l'isolant de grille (figure(1.21)), ce qui offre une meilleure capacité de couplage. Une faible variation de la tension de grille suffit désormais pour passer d'un régime à l'autre : pertinent pour les opérations à hautes fréquences. Des courants I_{ON} supérieurs à 2000 $\mu\text{A}/\mu\text{m}$ ont été atteints, ce qui est très supérieur aux meilleurs dispositifs CMOS innovants déca-nanométriques.

Bien que les performances électriques de ce transistor soient très prometteuses, la suppression totale des barrières Schottky permettrait d'injecter plus d'électrons et conduirait à des résultats spectaculaires. La lutte contre les barrières d'énergie dans les composants de la microélectronique est chargée d'histoire. Déjà dans les vénérables tubes à vide (figure (1.22)) les électrons devaient franchir une barrière d'énergie pour passer du fil métallique au vide.

Dans les composants à semiconducteurs classiques comme les MOSFETs, ce problème a été contourné en remplaçant le fil métallique par une région de silicium dopé par des impuretés

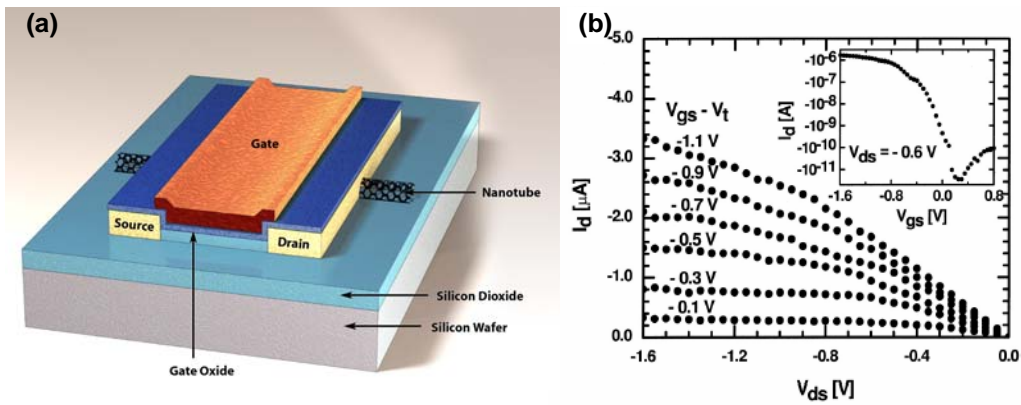


FIG. 1.21: (a) CNTFET à grille avant ("top gate"). (b) Caractéristiques de courant d'un CNTFET de type-p avec une grille en titane et une épaisseur d'oxyde de 15 nm. La tension de seuil vaut -0.5 V. Figure insérée : Caractéristique de transfert d'un CNTFET pour $V_{DS} = -0.6$ V. [74, 75]

appropriées. Dans le cas des nanotubes nous avons expliqué que la réalisation de contacts efficaces nécessitait un dopage extrêmement précis non-envisageable dans l'état de l'art de la technologie actuelle. L'autre alternative est d'utiliser une combinaison judicieuse entre fil métallique et nanotube semiconducteur. Récemment JAVEY *et al* [77] ont fabriqué un CNTFET à contact ohmique, qui présentait des transmissions balistiques à travers le canal. Bien que les raisons pour lesquelles l'architecture choisie (réservoirs palladium en contact avec un nanotube de 3 nm de diamètre) élimine la barrière Schottky restent encore méconnues, ils ont pu étudier les propriétés de transport de ce composant [78]. Le transport d'un nanotube ne serait pas balistique, mais quasi-balistique. La diffusion par émission de phonons optiques (dont le libre parcours moyen serait de 15 nm et l'énergie égale à 0.2 eV) engendrerait des réflexions pour de fortes tensions de drain ($V_{DS} > 0.5$ V). Ainsi les nanotubes de longueur supérieure à 10-15 nm sont respectivement balistiques et quasi-balistiques à faible et forte polarisation de drain.

La compréhension et la réalisation de CNTFETs performants semblent presque acquises. Les transistors sont désormais ambipolaires, c'est-à-dire qu'ils présentent un comportement symétrique pour les trous et les électrons [79]. Le prochain défi est désormais de développer l'électronique moléculaire au delà du composant isolé et de l'intégrer sur une puce électronique. MENON *et al.* [80] ont étudié théoriquement les jonctions de nanotubes de carbone en T et Y (figure (1.23)). Ces composants à trois terminaux peuvent tout d'abord servir de connections entre composants, mais proposent aussi des fonctions complémentaires des composants à deux terminaux comme le contrôle d'opérations transitoires (modulateur de courant) qui seront nécessaires à la réalisation de circuits électroniques moléculaires avancés. La fabrication de tels composants n'est pas une gageure puisque des connections moléculaires en forme de X ont été réalisées en irradiant des nanotubes croisés avec un canon à électrons [81]. En poursuivant l'irradiation, il est envisageable de supprimer un bras du X afin d'obtenir une jonction en T ou Y.

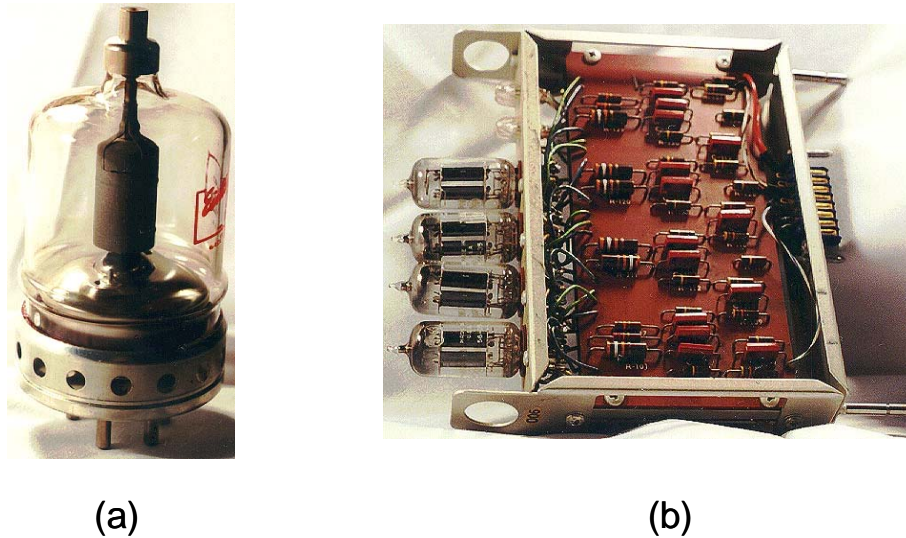


FIG. 1.22: a) Triode de puissance. Comparable aux transistors de puissance actuels, elle fut utilisée sur les premiers ordinateurs. b) Compteur à base de tubes à vide. Quatre tubes à vide sont nécessaires pour compter et stocker les 4 bits définissant un nombre décimal. Chaque tube à vide à deux triodes qui prennent des états 1 et 0 [76].

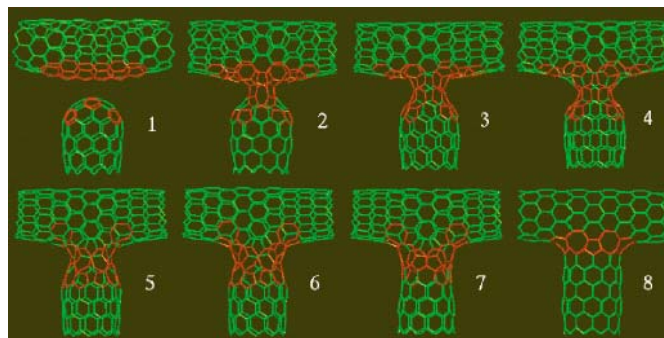


FIG. 1.23: Résultats de simulation montrant les différentes étapes de formation d'une jonction en T $(9,0)$ - $(10,0)$ - $(9,0)$ [80].

D'autres équipes ont déjà fabriqué des circuits électriques en incorporant des CNTFETs. LIU *et al.* [72] ont réalisé un circuit inverseur à partir de nanotubes dopés chimiquement alors que BATCHOLD *et al.* ont développé plusieurs portes logiques [82]. Enfin, des chercheurs des universités de Berkeley et Stanford ont dernièrement créé le premier circuit intégré de silicium qui intègre des nanotubes de carbone [83]. Le circuit, qui comprend plus de 2000 nanotubes, confirme la compatibilité du processus de synthèse des nanotubes de carbone avec la technologie du silicium, laissant ainsi la porte ouverte à de futures applications industrielles. L'entreprise de semiconducteurs Infineon a d'ailleurs présenté en mai 2004 ses premiers transistors en nanotubes de carbone dont la production en série est prévue pour 2007.

1.5 Conclusion : quel avenir pour le transistor ?

La longueur de grille des transistors, L_G , est passée progressivement de 10 μm en 1960 à 0.1 μm en 2000, permettant d'intégrer $(100)^2 = 10\,000$ fois plus de transistors sur une puce de surface donnée. Cette croissance de la densité d'intégration est le coeur de l'évolution informatique. Combien de temps cette diminution pourra continuer ? Dans 20 ans construirons-nous des transistors de longueur $L_G \simeq 0.01 \mu\text{m}$? Peut être, mais chacun s'accorde à dire que le chemin de la miniaturisation sera bien plus difficile, maintenant que nous atteignons l'échelle atomique. Pouvons-nous cependant imaginer la future brique élémentaire de l'électronique ? Rappelons tout d'abord que beaucoup de prévisions passées semblent aujourd'hui très cocasses. Par exemple en 1961, les scientifiques affirmaient qu'aucun transistor ne pouvait atteindre une longueur de grille de 10^{-5} m, alors qu'une puce d'Intel possède actuellement des centaines de millions de dispositifs 100 fois plus petits... Néanmoins les prévisions modernes sur la limite des dimensions devraient être plus fiables, puisque basées sur des principes de physique fondamentale. D'après les grandes industries des semiconducteurs (IBM, Infineon) la technologie MOSFET est condamnée à s'éteindre à plus ou moins longue échéance. Son successeur direct semble être le transistor à nanotube de carbone. Avant d'obtenir une rupture totale de technologie une première étape pourrait être la réalisation de systèmes hétérogènes combinant la technologie CMOS à de nouveaux composants moléculaires [84].

A plus long terme, des transistors ne transmettant qu'un électron à la fois sont envisagés. Un tel dispositif, appelé SET (pour "Single Electron Transistor"), ne pourrait prendre que deux états "on" ou "off" suivant qu'il contienne un électron ou pas et serait uniquement utilisé dans des codes binaires (et non pour amplifier). En nous projetant encore dans le temps, on peut imaginer des dispositifs qui tireraient profit des effets quantiques. Un électron pourrait être codé par trois niveaux : à la place du simple "on" ou "off", il pourrait y avoir aussi "quelque part entre on et off" [85]. Ce type de composant ouvrirait la voie à de nouveaux algorithmes dits "quantiques". Les ordinateurs quantiques qui en dériveraient auraient une puissance de calcul quasi-infinie !!! Mais ne nous emportons pas, la conjugaison des connaissances en physique, matériaux, composants et circuits, devrait assurer le développement de transistors MOSFETs jusqu'en 2015 [86]. Ainsi le chemin vers l'échelle atomique sera sinueux, incertain et dépendra de multiples paramètres : technologique, économique, juridique... Durant les 30 dernières années nous avons su quoi faire : réduire la taille des transistors. Les avancées industrielles nous permettent toujours de maintenir l'allure, mais la limite d'intégration est proche. Quand la limite sera atteinte les choses devront changer. Pourrons-nous un jour produire des composants fiables de quelques plans atomiques ? Ou découvrirons-nous des architectures pour lesquelles la structure de base ne sera plus le transistor ? Peut être, seul le temps nous le dira.

Mais il y a une chose qui apparaît inévitable : nous allons devoir apprendre à modéliser et à décrire les composants de taille atomique.

Chapitre 2

Méthode de calcul de la structure de bandes

Lorsque les dimensions du composant diminuent, l'ensemble des propriétés physiques du système se discrétisent. Les niveaux d'énergie propres de la structure se quantifient et une connaissance précise de la structure électronique devient indispensable. Ces dernières années de nouvelles méthodes sont apparues liées à des développements théoriques et à l'augmentation de la puissance de calcul des ordinateurs. Elles se divisent en deux catégories : les méthodes ab initio, qui reposent entièrement sur les principes fondamentaux de la mécanique quantique, et les méthodes semi-empiriques dont l'hamiltonien à une particule est paramétré de manière à retrouver les résultats ab initio. Dans ce chapitre, nous décrivons l'ensemble de ces techniques en soulignant leurs avantages et leurs limites.

2.1 Introduction : système d'électrons indépendants

Considérons dans un premier temps un hypothétique système de N électrons indépendants dont l'hamiltonien s'écrit :

$$H = \sum_{i=1}^N h(\mathbf{x}_i), \quad (2.1)$$

où les \mathbf{x}_i contiennent les coordonnées d'espace et de spin ($\mathbf{x}_i = \mathbf{r}_i, \xi$). Chaque hamiltonien individuel $h(\mathbf{x}_i)$ est identique et possède les mêmes solutions :

$$h(\mathbf{x}) \varphi_k(\mathbf{x}) = \varepsilon_k \varphi_k(\mathbf{x}). \quad (2.2)$$

Le vecteur propre ψ d'une telle configuration résulte du simple produit des états à un électron :

$$\psi = \prod_{k=1}^N \varphi_k(\mathbf{x}_k), \quad (2.3)$$

son énergie étant égale à la somme des valeurs propres correspondantes :

$$E = \sum_{k=1}^N \varepsilon_k. \quad (2.4)$$

En réalité, les électrons ne sont pas indépendants et l'hamiltonien d'un système à N électrons est donné par (si les noyaux atomiques sont fixes) :

$$H = \sum_{k=1}^N h_k + \frac{1}{2} \sum_{k=1}^N \sum_{k'(\neq k)=1}^N v(\mathbf{r}_k, \mathbf{r}_{k'}), \quad (2.5)$$

où les h_k sont les hamiltoniens mono-électroniques composés de l'énergie cinétique de l'électron k et de l'énergie potentielle due à l'interaction entre cet électron et les noyaux. Le second terme représente l'interaction répulsive électron-électron :

$$v(\mathbf{r}_k, \mathbf{r}_{k'}) = \frac{e^2}{|\mathbf{r}_k - \mathbf{r}_{k'}|}. \quad (2.6)$$

Ce sont bien sûr les termes de (2.6) qui interdisent la factorisation de H comme dans le cas d'électrons indépendants. La difficulté principale des calculs de structure électronique est liée au fait qu'il est impossible d'obtenir des solutions générales exactes de l'hamiltonien [Eq.(2.5)]. Il est alors nécessaire de recourir à des approximations plus ou moins heureuses suivant les propriétés du système étudié afin de se ramener à un ensemble d'équations mono-électroniques numériquement tractables.

2.2 Methodes *ab initio*

2.2.1 Equations de Hartree et approximation du champ moyen

Une première solution pour résoudre les problèmes à N corps fut proposée par HARTREE [87, 88], qui a considéré que chaque électron pouvait être traité séparément comme se déplaçant dans le champ électrostatique moyen dû aux autres électrons. La solution totale du système (Ψ_{Har}), établie à partir de la méthode des variations, est alors égale au produit des solutions de l'équation de SCHRÖDINGER mono-électronique de chaque particule k (φ_k) :

$$\Psi_{Har}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = \varphi_1(\mathbf{x}_1) \cdot \varphi_2(\mathbf{x}_2) \dots \varphi_N(\mathbf{x}_N), \quad (2.7)$$

où les $\varphi_k(\mathbf{x}_k)$ sont solutions de l'équation de SCHRÖDINGER mono-électronique :

$$\left[h(\mathbf{x}_k) + \sum_{k \neq k'} \int v(\mathbf{r}_k, \mathbf{r}_{k'}) |\varphi_{k'}(\mathbf{x}_{k'})|^2 d\mathbf{x}_{k'} \right] \varphi_k(\mathbf{x}_k) = \varepsilon_k^H \varphi_k(\mathbf{x}_k). \quad (2.8)$$

Il est plus adapté de rajouter et de retrancher le terme en $k = k'$ dans l'équation(2.8), ce qui donne :

$$\left[h(\mathbf{x}) + V_H(\mathbf{x}) - \int v(\mathbf{r}, \mathbf{r}') |\varphi_k(\mathbf{x}')|^2 d\mathbf{x}' \right] \varphi_k(\mathbf{x}) = \varepsilon_k^H \varphi_k(\mathbf{x}), \quad (2.9)$$

$$V_H(\mathbf{x}) = \int v(\mathbf{r}, \mathbf{r}') n(\mathbf{x}') d\mathbf{x}', \quad (2.10)$$

$$n(\mathbf{x}) = \sum_{k'} n_{k'} |\varphi_{k'}(\mathbf{x})|^2, \quad (2.11)$$

où $n(\mathbf{x})$ est la densité électronique, n_k sont les nombres d'occupation des états (0 pour les états vides, 1 pour les états occupés) et V_H est ce qu'on appelle habituellement le potentiel de HARTREE résultant de l'énergie potentielle électrostatique "moyenne" d'interaction de l'électron k avec tous les électrons du système (dont lui-même). Cette approche nécessite la résolution d'autant d'équations mono-électroniques que de particules dans le système. Le potentiel d'une équation dépendant des solutions des autres équations, la difficulté de ce calcul réside dans l'auto-cohérence ("self-consistency" en anglais). L'idée est d'arriver de façon itérative à des solutions approchées, en partant de solutions d'essais ($\varphi_1^0(x_1), \dots, \varphi_N^0(x_N)$). En posant $V(\mathbf{x}) = V_H(\mathbf{x}) - \int v(\mathbf{r}, \mathbf{r}') |\varphi_k(\mathbf{x}')|^2 d\mathbf{x}'$, on a :

$$\begin{array}{ccccccc} \varphi_1^0 & & \varphi_1^1 & & \varphi_1^2 & & \varphi_1^i \\ \vdots & \rightarrow V^1 \text{ résolution} & \vdots & \rightarrow V^1 \text{ résolution} & \vdots & \rightarrow V^2 \dots & \vdots & \rightarrow V^i \\ \varphi_N^0 & & \varphi_N^1 & & \varphi_N^2 & & \varphi_N^i \end{array} \quad (2.12)$$

Cette méthode étant parfois trop violente, une alternative plus souple est de pondérer le nouveau potentiel V_H^i avec le précédent V_H^{i-1} . L'auto-cohérence est atteinte lorsque la densité de charge totale n'évolue plus d'une itération à l'autre. L'application des équations de HARTREE aux

atomes isolés est assez réaliste, mais s'avère peu productive pour la description des solides.

2.2.2 Equations de Hartree-Fock

Nous avons vu dans le paragraphe précédent que les équations de HARTREE peuvent être obtenues en choisissant comme fonction d'essai du problème à N corps un produit de N fonctions mono-électroniques. Le problème principal de cette approche est que la fonction d'onde n'est pas antisymétrisée, comme doivent l'être toutes les particules fermioniques¹. Une amélioration de la méthode de HARTREE, développée par SLATER et FOCK en 1930, consiste donc à prendre comme fonction d'essai une fonction satisfaisant déjà au principe de symétrisation. On parle alors d'approximation de HARTREE-FOCK. Dans ce contexte, la fonction d'essai, également basée sur des fonctions mono-électroniques, est donnée par un déterminant de SLATER :

$$\Psi_{AS} = \frac{1}{\sqrt{N!}} \begin{vmatrix} \varphi_1(x_1) & \cdots & \varphi_1(x_N) \\ \varphi_2(x_1) & \cdots & \varphi_2(x_N) \\ \vdots & & \vdots \\ \varphi_N(x_1) & \cdots & \varphi_N(x_N) \end{vmatrix}. \quad (2.13)$$

L'application de la méthode variationnelle (ajustement des fonctions mono-électroniques de manière à minimiser l'énergie totale du système) à la fonction Ψ_{AS} donne un nouvel ensemble d'équations à une particule :

$$[h(\mathbf{x}) + V_H(\mathbf{x})] \varphi_k(\mathbf{x}) + \int \Sigma_{\mathbf{x}}(\mathbf{x}, \mathbf{x}') \varphi_k(\mathbf{x}') d\mathbf{x}' = \varepsilon_k^{HF} \varphi_k(\mathbf{x}), \quad (2.14)$$

$\Sigma_{\mathbf{x}}$ correspondant à un potentiel d'échange (non local) :

$$\Sigma_{\mathbf{x}}(\mathbf{x}, \mathbf{x}') = -v(\mathbf{r}, \mathbf{r}') \sum_l n_l \varphi_l(\mathbf{x}) \varphi_l^*(\mathbf{x}'), \quad (2.15)$$

où n_l est l'occupation de l'état φ_l (0 ou 1). Ce terme est sans équivalent classique. Le caractère antisymétrique de la fonction d'essai par rapport à la permutation de deux particules interdit aux deux électrons de se trouver au même endroit au même instant. V_H dans l'équation (2.14) est le potentiel de HARTREE défini en (2.9). La procédure de HARTREE-FOCK reproduit donc les équations de HARTREE avec des termes correctifs d'échange. Tout comme la méthode de HARTREE, la méthode de HARTREE-FOCK fait intervenir une boucle d'auto-cohérence. La fonction d'essai permet de calculer un potentiel moyen auquel sont soumis les électrons. Ce dernier fournit ensuite les nouvelles fonctions d'onde individuelles et ainsi de suite. L'itération s'arrête lorsque la convergence est atteinte, i.e. lorsque l'énergie obtenue reste constante à la précision demandée. La méthode de HARTREE-FOCK permet de traiter les propriétés magnétiques. Son application aux solides covalents donne des bandes d'énergie en général de bonne qualité, mais

¹En effet, l'indiscernabilité en mécanique quantique de deux particules identiques (même masse, charge électrique et spin), conduit au principe de PAULI qui stipule que la solution physique d'un système à plusieurs particules fermioniques est antisymétrique par permutation de deux particules quelconques [89].

surestime largement la largeur de la bande interdite. Ceci est principalement dû aux effets de corrélation dont l'importance est discutée dans la paragraphe suivant.

2.2.3 Fonctionnelle de la densité

Les problèmes rencontrés dans la méthode de HARTREE-FOCK proviennent des effets de corrélation. Les mouvements des électrons ne sont pas corrélés, puisque le potentiel vu par chaque électron est calculé comme une valeur moyenne sur les fonctions d'onde des autres électrons. Une étape importante vers le traitement de ces effets a été effectuée grâce aux méthodes de la fonctionnelle de la densité. L'idée de base est de remplacer le terme d'échange, complexe et non-local, de la méthode de HARTREE-FOCK par une approximation de densité moyenne, calculée localement comme si le système était un gaz parfait d'électrons.

La théorie de la fonctionnelle de la densité a été élaborée dans les années 1960 par Hohenberg et Kohn [90]. Elle dérive notamment du fait que deux potentiels extérieurs V_{ext} et V'_{ext} qui diffèrent de plus d'une constante ne peuvent donner la même densité d'électrons $n(\mathbf{r})$. Il existe donc une relation de bijection entre V_{ext} et $n(\mathbf{r})$ qui confère à $n(\mathbf{r})$ la faculté de complètement déterminer les propriétés de l'état fondamental du système. Il s'agit d'une reformulation de l'équation de SCHRÖDINGER pour le problème de N fermions en interaction qui offre en principe la solution exacte de l'état fondamental, moyennant la connaissance de la fonctionnelle de la densité électronique $F[n(\mathbf{r})]$ du système :

$$E[n(\mathbf{r})] = F[n(\mathbf{r})] + \int n(\mathbf{r}) V_{ext}(\mathbf{r}) d\mathbf{r}, \quad (2.16)$$

où $F[n(\mathbf{r})]$ comprend l'énergie cinétique ($T_0[n(\mathbf{r})]$) et le potentiel d'interaction électron-électron. Toutes les propriétés du système peuvent être obtenues à partir de la connaissance de cette fonctionnelle qui est minimale pour la densité de l'état fondamental. Bien qu'exacte cette écriture ne nous permet pas d'applications réelles puisque $F[n(\mathbf{r})]$ reste à déterminer. La démarche de Kohn et Sham [91] fut alors de relier le système de N électrons interagissants à un système fictif qui néglige ces interactions mais qui conduise à la même densité électronique $n(\mathbf{r})$:

$$\begin{aligned} E[n(\mathbf{r})] &= T_0[n(\mathbf{r})] + \int n(\mathbf{r}) V_{ext}(\mathbf{r}) d\mathbf{r} + \frac{1}{2} \int d\mathbf{r} \int d\mathbf{r}' \frac{n(\mathbf{r}) n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + E_{xc}[n(\mathbf{r})] \\ &= T_0[n(\mathbf{r})] + \int n(\mathbf{r}) \left(V_{ext}(\mathbf{r}) + \frac{1}{2} V_H(\mathbf{r}) \right) d\mathbf{r} + E_{xc}[n(\mathbf{r})], \end{aligned} \quad (2.17)$$

où V_H représente à nouveau l'énergie potentielle de HARTREE, c'est-à-dire le potentiel coulombien classique obéissant à l'équation de POISSON ($\Delta V_H + 4\pi e^2 n(\mathbf{r}) = 0$), V_{ext} est le potentiel extérieur dû aux noyaux tandis que $E_{xc}[n(\mathbf{r})]$ est la partie d'échange-corrélation qui contient toutes les contributions à N corps de l'énergie totale. L'équation (2.17) est exacte et à la base de toutes les méthodes dérivées de la fonctionnelle de densité.

La solution d'un tel système est immédiate : la fonction d'onde multi-électronique est simple-

ment un produit antisymétrisé de fonctions d'onde mono-électroniques φ_k obéissant à l'ensemble d'équations de SCHRÖDINGER :

$$\left[-\frac{\hbar^2}{2m_e} \Delta + V(\mathbf{r}) \right] \varphi_k(\mathbf{r}) = \varepsilon_k \varphi_k(\mathbf{r}), \quad (2.18)$$

où le potentiel extérieur de KOHN-SHAM $V(\mathbf{r})$ renferme le terme d'échange-corrélation $V_{xc}(\mathbf{r})$ à déterminer :

$$V(\mathbf{r}) = V_{ext}(\mathbf{r}) + V_H(\mathbf{r}) + V_{xc}(\mathbf{r}), \quad (2.19)$$

et où l'on a posé $V_{xc}(\mathbf{r}) = \frac{\delta E_{xc}[n(\mathbf{r})]}{\delta n(\mathbf{r})}$.

La densité de charges est donnée par la relation usuelle :

$$n(\mathbf{r}) = \sum_{k=1}^N |\varphi_k(\mathbf{r})|^2. \quad (2.20)$$

En réalité, l'équation(2.18) donne un ensemble d'équations à un corps appelées équations de KOHN-SHAM [91], couplées et non-linéaires car le potentiel $V(\mathbf{r})$ dépend aussi de la densité. Ces équations doivent être résolues par une méthode auto-cohérente.

Jusqu'à présent la formulation que nous venons de donner est exacte. Le terme d'échange-corrélation qui est la partie inconnue de l'hamiltonien a été repoussé dans le terme $V_{xc}(\mathbf{r})$. La seule approximation portera donc sur ce terme, qui ne peut être connu exactement. L'approximation la plus courante est celle de la densité locale (LDA pour "Local Density Approximation") qui consiste à écrire :

$$E_{xc}[n(\mathbf{r})] = \int d\mathbf{r} n(\mathbf{r}) \varepsilon_{xc}[n(\mathbf{r})], \quad (2.21)$$

où ε_{xc} est l'énergie d'échange-corrélation par électron d'un gaz de densité uniforme $n(\mathbf{r})$. Cette énergie d'échange-corrélation est souvent estimée à l'aide de calculs de type Monte-Carlo quantique. La fonctionnelle d'échange-corrélation est dans ce cas uniquement fonction de la densité au point \mathbf{r} (d'où la désignation "densité locale"). L'approximation LDA caractérise en général correctement les propriétés structurales des molécules et des solides covalents. Les énergies d'excitations sont en revanche mal restituées par cette approche. Par exemple, la largeur de la bande interdite des isolants et semiconducteurs est largement sous-estimée. Elle est égale à 0.6 eV pour le silicium au lieu de 1.2 eV expérimentalement. L'origine de ces erreurs provient de la discontinuité que subit le potentiel d'échange-corrélation au passage du gap du matériau (c'est-à-dire quand on ajoute un électron au solide). Ce phénomène, qui traduit un comportement extrêmement non local du potentiel d'échange-corrélation², laisse apparaître les limites d'une approximation type LDA.

²Lorsqu'un électron est ajouté au solide, le potentiel d'échange-corrélation varie dans tout le système.

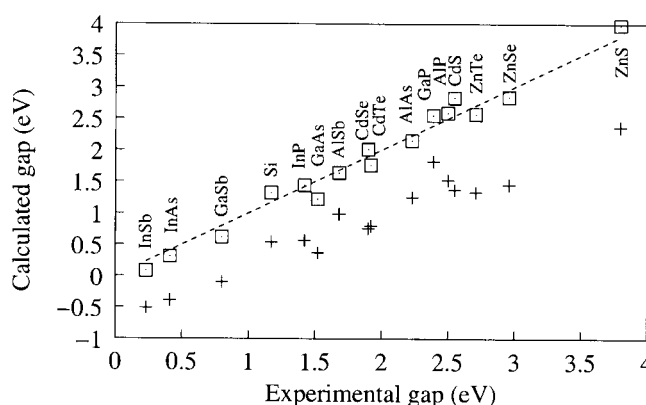


FIG. 2.1: Gap calculés en LDA (+) et en GW (carré) en fonction des gaps expérimentaux. Les résultats concernant les semiconducteurs II-VI proviennent de la référence [92] et les autres de [93].

2.2.4 Approximation GW

Dans le cas d'isolants et de semiconducteurs massifs, il faut avoir recours à une méthode plus sophistiquée permettant d'améliorer les résultats des théories de la fonctionnelle de la densité en évaluant l'opérateur de self-énergie électronique $\Sigma(\mathbf{x}, \mathbf{x}', \varepsilon)$ qui contient l'ensemble des effets d'échange et de corrélation. La difficulté réside dans la détermination de l'opérateur self-énergie qui peut être estimé en utilisant l'approximation GW. Cette dernière consiste à prendre le premier ordre en théorie des perturbations de l'opérateur $\Sigma(\mathbf{x}, \mathbf{x}', \varepsilon)$ en fonction de l'interaction dynamique écrantée $W(\mathbf{r}, \mathbf{r}', \omega)$ entre les électrons :

$$W(\mathbf{r}, \mathbf{r}', \omega) = \int \epsilon^{-1}(\mathbf{r}, \mathbf{r}'', \omega) v(\mathbf{r}'', \mathbf{r}') d\mathbf{r}'', \quad (2.22)$$

où $\epsilon^{-1}(\mathbf{r}, \mathbf{r}'', \omega)$ est la fonction diélectrique inverse à la pulsation ω . En pratique, la méthode la plus employée pour traiter un système dans l'approximation GW est de calculer les énergies en LDA, puis de corriger ces énergies en traitant la différence entre la self-énergie et le potentiel d'échange et de corrélation en perturbation au premier ordre. Les résultats obtenus dans l'approximation GW pour les solides covalents ou ioniques sont en général de très bonne qualité en ce qui concerne la largeur de bande interdite et la structure de bande dans son ensemble (figure (2.1)).

2.2.5 Méthodes numériques

L'ensemble des théories décrites précédemment calculent la structure électronique des matériaux dont les électrons peuvent se répartir en deux catégories : les orbitales de coeur et les orbitales de valence. Il existe différentes techniques pour supprimer les problèmes numériques liés aux comportements antagonistes de ces deux types d'orbitales.

Pseudopotentiels

Tous les électrons d'un atome ne jouent pas le même rôle. Ceux des couches internes (électrons de coeur) ne participent pas directement aux liaisons chimiques, alors que les électrons de la couche de valence sont les plus actifs. En réalité, les états de coeur dans un solide ou une molécule sont très proches de ceux des atomes libres : ils sont extrêmement ramassés autour du noyau et leurs intercorrélations sont très faibles. Par conséquent, il est parfois avantageux de remplacer les électrons de coeur par des potentiels effectifs. En ne traitant explicitement que les électrons de valence, on ne perd pratiquement aucune information sur les propriétés du solide, mais on réduit de façon très significative le volume des calculs à effectuer, surtout si le système étudié contient des atomes lourds. La théorie des pseudopotentiels génère des méthodes permettant d'éliminer les états de coeur des calculs et de concentrer les moyens numériques sur les états de valence qui sont plus simples à décrire. Considérons une équation de SCHRÖDINGER à résoudre de la forme :

$$(T + V) |\varphi\rangle = \varepsilon |\varphi\rangle. \quad (2.23)$$

L'état propre de valence $|\varphi\rangle$ est orthogonal aux états de coeur $|c\rangle$ produit par le même potentiel V . La fonction d'onde $|\varphi\rangle$ doit donc être fortement oscillante au voisinage du coeur atomique si l'on souhaite que son produit scalaire avec la fonction d'onde $|c\rangle$, compacte et regroupée autour de l'atome, soit nul. En pratique, la description dans une base d'ondes planes d'une fonction oscillante est impossible à réaliser (un trop grand nombre d'ondes planes serait nécessaire pour décrire convenablement les oscillations). Cette constatation est à la base de l'approximation du pseudopotentiel qui remplace le potentiel du noyau V (en $1/r$) écranté par les électrons de coeur, par un pseudopotentiel beaucoup plus "doux" V_{ps} (qui ne présente pas de divergence). Exprimons tout d'abord l'état de valence sous la forme :

$$|\varphi\rangle = (1 - P) |\psi\rangle, \quad (2.24)$$

où P est le projecteur sur les états de coeur :

$$P = \sum_c |c\rangle \langle c|. \quad (2.25)$$

$|\varphi\rangle$ est automatiquement orthogonal aux états de coeur et le pseudo-état $|\psi\rangle$ ne doit pas satisfaire à l'exigence d'orthogonalité. L'équation de SCHRÖDINGER devient :

$$(T + V) (1 - P) |\psi\rangle = \varepsilon (1 - P) |\psi\rangle. \quad (2.26)$$

Parce que les états de coeur $|c\rangle$ sont des vecteurs propres de l'hamiltonien $T + V$ associés à l'énergie ε_c , l'équation (2.26) peut s'écrire sous la forme :

$$\left[T + V + \sum_c (\varepsilon - \varepsilon_c) |c\rangle \langle c| \right] |\psi\rangle = \varepsilon |\psi\rangle. \quad (2.27)$$

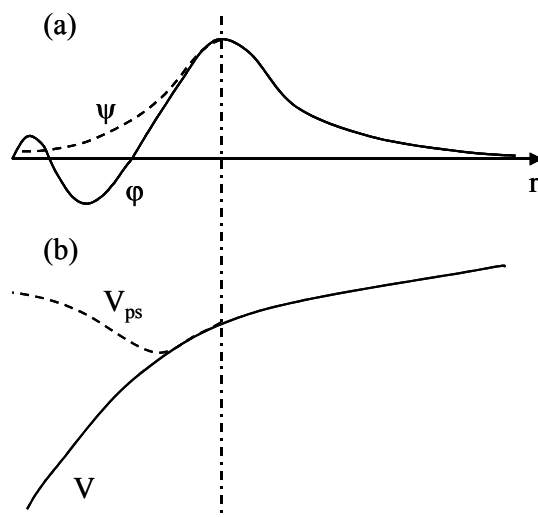


FIG. 2.2: a) Allure générale de la fonction d'onde et de la pseudo-fonction d'onde (trait pointillé). b) Allure générale du potentiel et du pseudopotentiel (trait pointillé).

Le potentiel V est alors remplacé par le pseudopotentiel suivant :

$$V_{ps} = V + \sum_c (\varepsilon - \varepsilon_c) |c\rangle \langle c|, \quad (2.28)$$

où le nouveau terme a été introduit pour tenir compte des effets des états de coeur sur les fonctions d'onde des électrons de valence. L'équation de SCHRÖDINGER devient :

$$(T + V_{ps}) |\psi\rangle = \varepsilon |\psi\rangle. \quad (2.29)$$

Les valeurs propres de cette équation restent clairement égales aux énergies des états de valence $|\varphi\rangle$ (le terme $\varepsilon - \varepsilon_c$ s'annule pour des valeurs propres associées aux états de coeur). V_{ps} est un opérateur non-local, dont la définition n'est pas unique (on peut lui ajouter n'importe quelle combinaison linéaire d'orbitales de coeur sans changer les valeurs propres). Cette propriété permet d'optimiser le pseudopotentiel de manière à rendre les pseudo-fonctions $|\psi\rangle$ aussi monotones que possible (voir figure(2.2)). Les pseudopotentiels sont souvent obtenus par des calculs *ab initio* et sont ajustés de manière à être applicables à des systèmes variés (critère de transférabilité).

Bases

La résolution des équations mono-électroniques en *ab initio* repose toujours sur le développement des fonctions propres suivant une base. Les bases les plus courantes se divisent en deux catégories : les ondes planes délocalisées et les orbitales atomiques localisées.

De par leur nature, les ondes planes sont particulièrement efficaces pour décrire des systèmes périodiques tels que les cristaux. Le théorème de Bloch montre qu'une fonction d'onde se dé-

compose naturellement sur une base d'ondes planes de vecteur d'onde \mathbf{k} :

$$\varphi_{\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}} c_{\mathbf{k}+\mathbf{G}} e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}}, \quad (2.30)$$

où \mathbf{G} est un vecteur du réseau réciproque. La base d'ondes planes est complète (peut décrire tous les états) mais infinie. Cependant, les coefficients sont plus importants pour les ondes planes d'énergie cinétique faible. On peut donc tronquer le développement et ne prendre en compte que les ondes planes dont l'énergie cinétique $\frac{\hbar^2(\mathbf{k}+\mathbf{G})^2}{2m}$ est inférieure à une certaine énergie (on parle de "cutoff"). Notons que l'utilisation des ondes planes est en général couplée à une description des atomes en pseudopotentiels qui ne conserve que les orbitales délocalisées.

Les bases d'orbitales localisées sont elles plus particulièrement utilisées pour décrire les systèmes finis comme les molécules. Les fonctions propres sont écrites comme une combinaison linéaire d'orbitales atomiques (CLOA, qui devient en anglais LCAO pour "Linear Combination of Atomic Orbitals") :

$$\varphi = \sum_{i,\alpha} c_{i\alpha} \phi_{i\alpha}, \quad (2.31)$$

où $\phi_{i\alpha}$ est la α -ième orbitale de l'atome i . En pratique, la base est tronquée. Les énergies propres sont alors données par l'équation séculaire :

$$\det(H - ES) = 0, \quad (2.32)$$

où H est la matrice de l'hamiltonien exprimé dans la base atomique réduite et S est la matrice de recouvrements :

$$S_{i\alpha,j\beta} = \langle \phi_{i\alpha} | \phi_{j\beta} \rangle. \quad (2.33)$$

2.3 Méthodes semi-empiriques

Nous avons décrit dans le paragraphe précédent le développement considérable de méthodes *ab initio* basées sur les premiers principes de la mécanique quantique dans lesquelles aucun paramètre d'ajustement n'était utilisé. En dépit des résultats obtenus et malgré l'élégance de ces approches, il est nécessaire de constamment passer d'une méthode à l'autre en fonction du matériau traité et des propriétés que l'on souhaite étudier. Un autre inconvénient réside dans la complexité de mise en oeuvre de ces théories dont la première conséquence se traduit par un besoin boulimique de puissance numérique. Aussi leur utilisation se restreint à la description de systèmes de très petites dimensions (≈ 100 atomes) et n'est pas appropriée à l'étude des cellules mésoscopiques que sont les puits quantiques ou les transistors. Les méthodes semi-empiriques (contenant des paramètres ajustés à partir de calculs *ab initio* ou de données expérimentales) semblent au contraire toutes indiquées pour ce genre d'études, car elles permettent d'obtenir des descriptions simples, pédagogiques et performantes de problèmes complexes.

2.3.1 Pseudopotentiels empiriques

Il s'agit d'une approche dérivée de la méthode des pseudopotentiels décrite précédemment, mais dans laquelle les pseudopotentiels sont supposés locaux pouvant s'écrire comme des sommes de contributions atomiques :

$$V(\mathbf{r}) = \sum_{j,\alpha} v_j(\mathbf{r} - \mathbf{R}_j - \mathbf{r}_\alpha), \quad (2.34)$$

où j somme sur les mailles positionnées en \mathbf{R}_j , et α indexe les atomes situés en \mathbf{r}_α dans la maille. Les éléments de la matrice V entre deux ondes planes s'expriment de la façon suivante :

$$\langle \mathbf{k} + \mathbf{G} | V | \mathbf{k} + \mathbf{G}' \rangle = \frac{1}{\Omega} \sum_{\alpha} e^{i(\mathbf{G}' - \mathbf{G}) \cdot \mathbf{r}_\alpha} \int v_{\alpha}(\mathbf{r}) e^{i(\mathbf{G}' - \mathbf{G}) \cdot \mathbf{r}} d\mathbf{r}, \quad (2.35)$$

où Ω est le volume de la cellule unité. Supposons que plusieurs atomes de la maille soient identiques. La somme sur α peut se décomposer en une somme sur des groupes β d'atomes identiques dont la position est spécifiée par un second index γ ($\mathbf{r}_\alpha = \mathbf{r}_{\gamma\beta}$). Notant n le nombre d'atomes dans une maille unitaire, on obtient :

$$\langle \mathbf{k} + \mathbf{G} | V | \mathbf{k} + \mathbf{G}' \rangle = \sum_{\beta} S_{\beta}(\mathbf{G}' - \mathbf{G}) v_{\beta}(\mathbf{G}' - \mathbf{G}), \quad (2.36)$$

où S_{β} et v_{β} sont respectivement les facteurs de structure et de forme des espèces atomiques correspondantes définis par :

$$S_{\beta}(\mathbf{G}) = \frac{1}{n} \sum_{\gamma} e^{i\mathbf{G} \cdot \mathbf{r}_{\gamma\beta}}, \quad (2.37)$$

et

$$v_{\beta}(\mathbf{G}) = \frac{n}{\Omega} \int v_{\beta}(\mathbf{r}) e^{i\mathbf{G} \cdot \mathbf{r}} d\mathbf{r}. \quad (2.38)$$

Si $v_{\beta}(\mathbf{r})$ sont des potentiels *lisses*, leur transformée $v_{\beta}(\mathbf{G})$ diminuera rapidement avec $|G|$ de telle sorte qu'une bonne approximation est de tronquer leur calcul à une valeur maximale G_c . Par exemple, la structure de bande de semiconducteurs tétravalents comme le silicium peut être assez bien reproduite en utilisant seulement les trois plus basses composantes de FOURIER $v(|\mathbf{G}|)$ du pseudopotentiel atomique [94]. Il y a en conséquence deux valeurs limites pour $|\mathbf{G}|$: la première, G_M , limite le nombre d'ondes planes considérées (≈ 50) et donc la taille de la matrice de l'hamiltonien, tandis que la seconde, G_c , limite le nombre de composantes de FOURIER des facteurs de forme.

En pratique, la méthode des pseudopotentiels empiriques gère les facteurs de forme $v_{\beta}(\mathbf{G})$ comme des paramètres et les ajuste de manière à obtenir des structures de bande satisfaisantes. L'utilisation d'un pseudopotentiel empirique n'est pas pleinement justifiée puisque d'après l'équation (2.27) des opérateurs de projection devraient intervenir. Cette approche offre néanmoins une excellente description des structures de bandes de matériaux massifs dans toute la première zone de Brillouin et peut décrire des systèmes contenant jusqu'à 2000 atomes.

2.3.2 Liaisons fortes

L'approximation des liaisons fortes est une méthode variationnelle dont la fonction d'essai est une combinaison linéaire d'orbitales atomiques (CLOA) [95] :

$$\Psi(\mathbf{r}) = \sum_{i=1}^N \sum_{\alpha=1}^m c_{i,\alpha} \phi_{i,\alpha}(\mathbf{r} - \mathbf{R}_i), \quad (2.39)$$

où $\phi_{i,\alpha}$ est l'orbitale α de l'atome i en position \mathbf{R}_i . Cette méthode, en principe exacte si l'on considère une infinité d'orbitales par atome, permet de trouver à partir d'une combinaison linéaire d'orbitales d'atomes isolés (solutions de l'équation de SCHRÖDINGER des atomes isolés), la solution de l'équation de SCHRÖDINGER du solide ou de la molécule par l'ajustement des coefficients de chaque orbitale $c_{i,\alpha}$. Généralement, la base d'orbitales atomiques utilisée est minimale, c'est-à-dire réduite aux orbitales de valence de l'atome isolé. Par exemple, pour le silicium on utilise les orbitales $3s$, $3p_x$, $3p_y$, $3p_z$ et pour l'hydrogène $1s$. L'équation aux valeurs propres de SCHRÖDINGER se met sous la forme matricielle suivante :

$$H\Psi = \varepsilon S\Psi, \quad (2.40)$$

où Ψ est le vecteur des $n = N \times m$ coefficients $c_{i,\alpha}$. H est la matrice $n \times n$ de l'hamiltonien. Cette dernière contient deux types de termes :

- les termes diagonaux ou intra-atomiques $H_{i\alpha,i\alpha} = \langle \phi_{i,\alpha} | H | \phi_{i,\alpha} \rangle$, qui décrivent l'énergie de l'orbitale α de l'atome i dans le système.

- les termes inter-atomiques $H_{i\alpha,j\beta} = \langle \phi_{i,\alpha} | H | \phi_{j,\beta} \rangle$ qui sont les "intégrales de saut" c'est-à-dire qu'ils décrivent l'interaction entre deux orbitales d'atomes différents. Les orbitales étant localisées, les termes $H_{i\alpha,j\beta}$ diminuent rapidement avec la distance entre les atomes i et j . En conséquence, seuls les termes entre premiers, seconds voire mais c'est plus rare troisièmes voisins sont conservés.

Enfin dans l'équation (2.40) S est la matrice de recouvrement :

$$S_{i\alpha,j\beta} = \langle \phi_{i,\alpha} | \phi_{j,\beta} \rangle. \quad (2.41)$$

L'équation (2.40) représente un système de n équations à n inconnues qui n'accepte de solutions non triviales (tous les coefficients $c_{i,\alpha} = 0$) que si l'opérateur linéaire H qui à une fonction Ψ associe la fonction $\varepsilon S\Psi$ n'est pas bijectif, c'est-à-dire si :

$$\det |H - \varepsilon S| = 0. \quad (2.42)$$

L'équation (2.42) fournit les énergies propres du système. L'approximation des liaisons fortes consiste à négliger les recouvrements entre orbitales :

$$\langle \phi_{i,\alpha} | \phi_{j,\beta} \rangle = \delta_{i,j} \delta_{\alpha,\beta}. \quad (2.43)$$

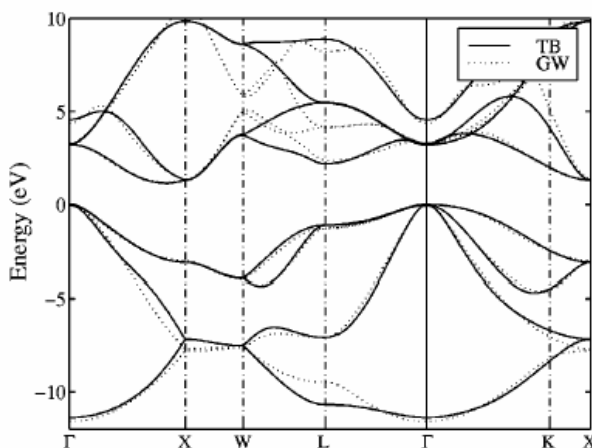


FIG. 2.3: Structures de bande du silicium calculées en théorie des liaisons fortes [99] et en GW [100].

La principale qualité de la méthode est sa simplicité. Avec quelques paramètres, on peut décrire la structure électronique de solides ou molécules très complexes. De plus, son application n'est pas limitée à de très petits systèmes comme en calculs *ab initio*. Par exemple, des études ont été réalisées sur des clusters de silicium de plus de 100000 atomes. D'autre part, sa simplicité permet une interprétation plus facile des résultats.

L'autre qualité est la diversité de problèmes physiques que les liaisons fortes permettent de traiter. Il est possible de les utiliser pour des problèmes aussi divers que les molécules organiques [96], les nanotubes [97, 98], les semi-conducteurs [99] (figure (2.3)), les défauts dans semi-conducteurs [101]...

2.3.3 L'auto-cohérence en liaisons fortes

La méthode des liaisons fortes peut être rendue auto-cohérente en couplant l'équation de SCHRÖDINGER à celle de POISSON. Ce type d'approche est équivalent à un calcul mené dans l'approximation de HARTREE, dans laquelle l'effet sur l'électron étudié de l'ensemble des électrons du système est pris en compte en considérant un potentiel macroscopique moyen de répulsion. En liaisons fortes les charges sont supposées ponctuelles et localisées sur les sites atomiques. L'équation de SCHRÖDINGER permet, à partir du potentiel extérieur V_{ext} et du potentiel à l'équilibre V_{eq} de calculer les charges sur chaque site (module au carré au point \mathbf{r} de la fonction d'onde). L'expression de la densité électronique $n(\mathbf{r})$ peut se simplifier en :

$$n(\mathbf{r}) = \sum_{i,\alpha} n_{i,\alpha} |\phi_{i,\alpha}(\mathbf{r})|^2, \quad (2.44)$$

$n_{i,\alpha}$ étant la population électronique de l'orbitale $\phi_{i,\alpha}$. L'équation de POISSON quand à elle fournit à partir des charges un nouveau potentiel (voir figure (2.4)). La modélisation de l'écrantage (\equiv résolution de l'équation de POISSON) repose sur les intégrales de COULOMB, usuellement

définies par :

$$\begin{aligned} \int |\phi_{i,\alpha}(\mathbf{r})|^2 v(\mathbf{r}, \mathbf{r}') |\phi_{j,\beta}(\mathbf{r}')|^2 d\mathbf{r}d\mathbf{r}' &= v_{ij} = \frac{e^2}{R_{ij}} \text{ si } i \neq j, \\ &= U_i^{at} \text{ si } i = j, \end{aligned} \quad (2.45)$$

où U_i^{at} est l'énergie intra-atomique de COULOMB, calculée à partir des fonctions d'ondes atomiques, et $v(\mathbf{r}, \mathbf{r}') = e^2/|\mathbf{r} - \mathbf{r}'|$ représente l'interaction entre un électron en \mathbf{r} et un autre électron en \mathbf{r}' . Ce dernier induit une modification de la densité électronique δn_{ind} qui en retour crée un potentiel V_{ind} . Le potentiel total s'exprime alors :

$$W = v + V_{\text{ind}}. \quad (2.46)$$

Nous allons maintenant calculer le potentiel total du système W en \mathbf{r} créé par une charge test en \mathbf{r}' :

$$W(\mathbf{r}, \mathbf{r}') = v(\mathbf{r}, \mathbf{r}') + \int v(\mathbf{r}, \mathbf{r}'') \delta n_{\text{ind}}(\mathbf{r}'', \mathbf{r}') d\mathbf{r}'', \quad (2.47)$$

où $\delta n_{\text{ind}}(\mathbf{r}'', \mathbf{r}')$ définit la modification des charges en \mathbf{r}'' due à la charge test introduit en \mathbf{r}' . Dans le formalisme des liaisons fortes, la charge test est nécessairement localisée sur un atome j , c'est-à-dire distribuée selon $|\phi_{j,\alpha}(\mathbf{r})|^2$. Multipliant l'équation (2.47) par cette distribution et intégrant sur \mathbf{r}' , nous obtenons :

$$W(\mathbf{r}, \mathbf{R}_j) = v(\mathbf{r}, \mathbf{R}_j) + \int v(\mathbf{r}, \mathbf{r}'') \delta n_{\text{ind}}(\mathbf{r}'', \mathbf{R}_j) d\mathbf{r}''. \quad (2.48)$$

Nous développons ensuite $\delta n_{\text{ind}}(\mathbf{r}'', \mathbf{R}_j)$ sur chaque orbitale au carré comme décrit dans l'équation (2.44) pour écrire :

$$W(\mathbf{r}, \mathbf{R}_j) = v(\mathbf{r}, \mathbf{R}_j) + \sum_k v(\mathbf{r}, \mathbf{R}_k) \delta n_k(\mathbf{R}_j). \quad (2.49)$$

l'étape finale consiste à multiplier (2.49) par $|\phi_{i,\alpha}(\mathbf{r})|^2$ et à intégrer sur \mathbf{r} . Le potentiel total est ainsi discrétisé sur chaque atome i, j :

$$W_{ij} = v_{ij} + \sum_k v_{ik} \delta n_{kj}. \quad (2.50)$$

La boucle auto-cohérente prend fin lorsque la convergence est atteinte, c'est-à-dire lorsque le potentiel total et les charges ne fluctuent plus (ou presque) d'une itération à une autre.

2.3.4 Masse effective et théorie \mathbf{k}, \mathbf{p}

Le concept de la masse effective près d'un extrémum de bande est très efficace pour décrire les hétérostructures et les nanostructures lorsqu'elles ne sont pas dans un état hors-équilibre trop prononcé. Cette approche est directement obtenue via la méthode \mathbf{k}, \mathbf{p} que nous présentons

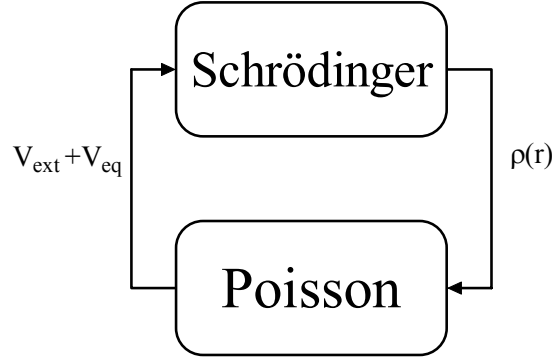


FIG. 2.4: Boucle d'auto-cohérence SCHRÖDINGER-POISSON.

ci-après.

La méthode $\mathbf{k}\cdot\mathbf{p}$ permet de reproduire les principales bandes d'un matériau massif cristallin au voisinage d'un point \mathbf{k}_0 situé dans la première zone de BRILLOUIN [102]. Elle consiste à résoudre l'équation de SCHRÖDINGER avec une fonction de BLOCH et à supposer connues les solutions aux extrema de bandes :

$$\left\{ \frac{p^2}{2m_e} + V \right\} e^{i\mathbf{k}\cdot\mathbf{r}} u_{\mathbf{k}}(\mathbf{r}) = \varepsilon(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}} u_{\mathbf{k}}(\mathbf{r}). \quad (2.51)$$

L'opérateur quantité de mouvement étant défini par : $\mathbf{p}[\psi] = -i\hbar \overrightarrow{\text{grad}}[\psi]$, l'équation (2.51) peut se réécrire sous la forme :

$$\left\{ \frac{(\mathbf{p} + \hbar\mathbf{k})^2}{2m_e} + V \right\} u_{\mathbf{k}}(\mathbf{r}) = \varepsilon(\mathbf{k}) u_{\mathbf{k}}(\mathbf{r}) \quad (2.52)$$

Afin de résoudre (2.52), nous pouvons développer la partie périodique inconnue $u_{\mathbf{k}}(\mathbf{r})$ sur la base des solutions correspondantes à un point \mathbf{k}_0 donné, que nous noterons $u_{n,\mathbf{k}_0}(\mathbf{r})$:

$$u_{\mathbf{k}}(\mathbf{r}) = \sum_n c_n(\mathbf{k}) u_{n,\mathbf{k}_0}(\mathbf{r}). \quad (2.53)$$

Les solutions associées à l'équation (2.53) sont les vecteurs propres et les valeurs propres de la matrice de terme général :

$$A_{n,n'}(\mathbf{k}) = \langle u_{n,\mathbf{k}_0} | \frac{(\mathbf{p} + \hbar\mathbf{k})^2}{2m_e} | u_{n',\mathbf{k}_0} \rangle. \quad (2.54)$$

Nous utilisons maintenant le fait que $u_{n,\mathbf{k}_0}(\mathbf{r})$ soit une fonction propre de l'équation (2.52) pour $\mathbf{k} = \mathbf{k}_0$, avec l'énergie $\varepsilon_n(\mathbf{k}_0)$. Une forme plus simple de l'équation (2.54) est alors :

$$A_{n,n'}(\mathbf{k}) = \left\{ \varepsilon_n(\mathbf{k}_0) + \frac{\hbar^2}{2m_e} (\mathbf{k} - \mathbf{k}_0)^2 \right\} \delta_{n,n'} + \frac{\hbar(\mathbf{k} - \mathbf{k}_0)}{m_e} \mathbf{p}_{nn'}(\mathbf{k}_0), \quad (2.55)$$

avec :

$$\mathbf{p}_{n,n'}(\mathbf{k}_0) = \langle u_{n,\mathbf{k}_0} | \mathbf{p} | u_{n',\mathbf{k}_0} \rangle. \quad (2.56)$$

La diagonalisation de la matrice $A(\mathbf{k})$ donnée par l'équation (2.55) peut fournir la structure de bande exacte. Cependant, l'intérêt de cette méthode est qu'elle constitue un point de départ "inné" pour un développement en perturbations. Illustrons cette disposition à travers le cas d'un extrémum simple non-dégénéré. Considérons une bande d'énergie $\varepsilon_n(\mathbf{k})$ présentant un extrémum en $\mathbf{k} = \mathbf{k}_0$ et étudions son évolution pour des valeurs du vecteur d'onde proches de \mathbf{k}_0 . Le dernier terme de (2.55) peut être alors envisagé comme une perturbation, nous permettant ainsi de calculer la différence $\varepsilon_n(\mathbf{k}) - \varepsilon_n(\mathbf{k}_0)$ par une théorie de second ordre appliquée à la matrice $A(\mathbf{k})$. Nous obtenons³ :

$$\varepsilon_n(\mathbf{k}) = \varepsilon_n(\mathbf{k}_0) + \frac{\hbar^2}{2m_e} (\mathbf{k} - \mathbf{k}_0)^2 + \frac{\hbar^2}{m_e^2} \sum_{n' \neq n} \frac{[(\mathbf{k} - \mathbf{k}_0) \cdot \mathbf{p}_{nn'}][(\mathbf{k} - \mathbf{k}_0) \cdot \mathbf{p}_{n'n}]}{\varepsilon_n(\mathbf{k}_0) - \varepsilon_{n'}(\mathbf{k}_0)}, \quad (2.57)$$

qui représente le développement au second ordre en \mathbf{k}_0 conduisant à la définition des masses effectives. Le dernier terme de l'équation (2.57) est un tenseur. En notant α son principal axe, on obtient l'expression générale des masses effectives m_α^* :

$$\frac{m_e}{m_\alpha^*} = 1 + \frac{2}{m_e} \sum_{n' \neq n} \frac{|(p_\alpha)_{nn'}|^2}{\varepsilon_n(\mathbf{k}_0) - \varepsilon_{n'}(\mathbf{k}_0)}. \quad (2.58)$$

Ceci montre qu'en présence de deux bandes interagissantes, la plus élevée en énergie possède une masse effective positive alors que l'opposé est vrai pour la seconde.

L'approximation de la masse effective est souvent appliquée aux semi-conducteurs pour décrire le voisinage du minimum de bande de conduction. Dans cette approximation, la bande de conduction est parabolique ($\varepsilon(\mathbf{k}) = \hbar^2 \mathbf{k}^2 / (2m_e^*)$, où la masse effective m_e^* est mesurée expérimentalement). Dans les hétérostructures et nanostructures de semiconducteur, la méthode $\mathbf{k} \cdot \mathbf{p}$ substitue au réseau atomique un milieu continu caractérisé par des potentiels de confinements pour les électrons et pour les trous. Par exemple, l'approximation de la masse effective pour la bande de conduction d'un nanocristal homogène donne un résultat bien connu : chaque fonction propre $\psi_i(\mathbf{r})$ de H est le produit de la fonction de BLOCH qui varie rapidement (typiquement sur des distances inter-atomiques) par une fonction enveloppe $\phi_i(\mathbf{r})$ lentement variable. Celle-ci est solution d'une équation de SCHRÖDINGER pour une particule de masse effective m_e^* dans un potentiel de confinement $V(\mathbf{r})$:

$$-\frac{\hbar^2}{2m_e^*} \Delta \phi_i(\mathbf{r}) + V(\mathbf{r}) \phi_i(\mathbf{r}) = \varepsilon_i \phi_i(\mathbf{r}). \quad (2.59)$$

$V(\mathbf{r}) = 0$ à l'intérieur et $V(\mathbf{r}) = V_0$ à l'extérieur du nanocristal. En général, la limite $V_0 = +\infty$ est considérée. La durée du calcul est donc indépendante du nombre d'atomes N et de la taille du système. C'est là un atout majeur de la méthode $\mathbf{k} \cdot \mathbf{p}$, qui a largement contribué à la compréhension des nanostructures de semiconducteurs. Néanmoins, les modèles $\mathbf{k} \cdot \mathbf{p}$ prédisent tous des énergies de confinement trop élevées dans les petits nanocristaux de semiconducteurs,

³Rappel [89] : La correction de l'énergie fondamentale au second ordre due à une perturbation V s'écrit :

$$\varepsilon_n^{(2)} = \sum_{n' \neq n} \frac{|\langle u_{n', \mathbf{k}_0} | V | u_{n, \mathbf{k}_0} \rangle|^2}{\varepsilon_n(\mathbf{k}_0) - \varepsilon_{n'}(\mathbf{k}_0)}.$$

parce que leur description du matériau massif est limitée et parce qu'ils ne permettent pas d'appliquer des conditions aux limites réalistes.

Chapitre 3

Transport électronique dans les nano-composants

Les recherches sur les propriétés de transport dans les composants nano-électroniques (théoriques et expérimentales) ont été stimulées au cours des dernières années par la fabrication de transistors Métal-Oxyde-Semiconducteur, dont la longueur de grille atteint désormais les 10 nm [103–114]. D'autre part, des expériences ont réussi à mesurer le courant I à travers des systèmes quantiques simples tels que des molécules [115–118] ou des plots semiconducteurs [119, 120].

L'objectif de ce chapitre est de montrer l'évolution des méthodes de modélisation "conventionnelles" des MOSFETs, d'établir leurs limites pour simuler les effets de la miniaturisation (section (3.1)) et de présenter les nouvelles approches quantiques, basées sur les calculs de structure électronique développés en physique théorique [121]. Nous expliquons dans la section (3.2) les limites de la théorie des perturbations avant de décrire en détail la vision du transport développée par LANDAUER [122] et l'expression de son modèle dans le formalisme des fonctions de GREEN hors-équilibre (section (3.3)). Nous étudierons enfin l'importance de l'auto-cohérence dans de tels systèmes et concluons.

3.1 Transport électronique dans un nano-MOSFET : rencontre avec le quantique

Le transistor à effet de champ peut se représenter par une résistance modulable (le canal) connectée à des réservoirs d'électrons (source et drain). La grille ajuste la résistance du canal, alors que les réservoirs permettent d'y faire circuler un courant. Dans cette partie nous allons démontrer les limites des méthodes classiques et quantiques usuelles pour décrire le transport de charges dans les nano-composants, avant de les comparer au formalisme, plus complet, des fonctions de GREEN hors-équilibre.

3.1.1 Inadéquation des modèles classiques : du régime diffusif au balistique

L'équation de transport de BOLTZMANN est le point de départ de la majorité des modèles classiques. Dans cette approche une fonction de distribution $f(\mathbf{k}, \mathbf{r}, t)$ décrit la densité de particules $n(\mathbf{k}, \mathbf{r}, t)$ en terme de leur moment \mathbf{k} et de leur position \mathbf{r} dans l'espace à un temps t :

$$n(\mathbf{r}, t) = \int f(\mathbf{k}, \mathbf{r}, t) d\mathbf{k}^3. \quad (3.1)$$

Le long de la trajectoire d'une particule, la dérivée par rapport au temps de la fonction de distribution doit être nulle :

$$\frac{d}{dt} f(\mathbf{k}, \mathbf{r}, t) = 0. \quad (3.2)$$

La résolution exacte de l'équation (3.2) présentant des difficultés majeures, plusieurs hypothèses simplificatrices sont faites. En considérant que :

- les processus de diffusion sont élastiques,
- les forces extérieures sont constantes dans tout le système.
- seul le champ électrique \mathbf{E} contribue aux forces extérieures \mathbf{F}_E (pas le champ magnétique),
- la fonction de distribution est symétrique dans l'espace des moments,

l'équation de transport de BOLTZMANN devient [123] :

$$\frac{\partial f}{\partial t} + \frac{\mathbf{F}_E}{\hbar} \frac{\partial f}{\partial \mathbf{k}} + \mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{r}} = \left(\frac{\partial f}{\partial t} \right)_{\text{coll}}, \quad (3.3)$$

où $\left(\frac{\partial f}{\partial t} \right)_{\text{coll}}$ est le terme des collisions qui se décompose en trois composantes dues à la densité des porteurs n , à leur vitesse \mathbf{v} et leur énergie ε :

$$\left(\frac{\partial f}{\partial t} \right)_{\text{coll}} = \left(\frac{dn}{dt} \right)_{\text{coll}} + \left(\frac{d\mathbf{v}}{dt} \right)_{\text{coll}} + \left(\frac{d\varepsilon}{dt} \right)_{\text{coll}}, \quad (3.4)$$

où \mathbf{v} est la vitesse de groupe des particules ($=d\mathbf{r}/dt$).

Si l'on ne conserve que les seules contributions de génération-recombinaison, de relaxation de moment et d'énergie, les termes de collisions s'expriment :

$$\left(\frac{dn}{dt}\right)_{\text{coll}} = -R_g, \quad \left(\frac{d\mathbf{v}}{dt}\right)_{\text{coll}} = -\frac{\mathbf{v}}{\tau_p}, \quad \text{et} \quad \left(\frac{d\varepsilon}{dt}\right)_{\text{coll}} = -\frac{w - w_0}{\tau_w}, \quad (3.5)$$

où R_g est le taux de génération-recombinaison, τ_p est le temps de relaxation du moment, τ_w est le temps de relaxation de l'énergie, w est l'énergie moyenne de l'électron et w_0 est l'énergie de l'électron à l'équilibre à une température θ_0 .

La technique la plus populaire pour résoudre l'équation de transport de BOLTZMANN est la méthode de MONTE CARLO [124–126]. Cette approche suit la position et le moment d'un ensemble de particules se déplaçant dans un composant sous l'influence d'un champ électrique et de forces de diffusion. Des nombres aléatoires sont choisis pour déterminer le temps entre collisions, le type de diffusion rencontré, et la direction des porteurs après une collision. La procédure est répétée typiquement de 10^4 à 10^6 fois, pour simuler le cheminement des porteurs dans le composant. Dans le cas de problèmes dépendants du temps, l'ensemble du système doit suffisamment grand pour représenter correctement un gaz d'électrons. Par la suite, la diminution de la taille des transistors a inspiré des améliorations à la méthode initiale en incluant notamment les effets de quantification d'énergie dans les dispositifs confinés [127, 128] et ceux d'impuretés discrètes [8, 129, 130]. L'inconvénient majeur de ce type de simulation est son besoin boulimique de puissance numérique et le bruit statistique associé qui diminue avec le nombre de porteurs de charges N_{ch} considérés en $1/\sqrt{N_{ch}}$.

A) Le modèle de dérive-diffusion

Le modèle de dérive-diffusion résulte de l'équation de transport de BOLTZMANN simplifiée, dans laquelle seuls les deux premiers moments sont retenus [131] :

- l'équation de continuité des particules :

$$\frac{\partial n}{\partial t} + \nabla \cdot (n\mathbf{v}) = \left(\frac{\partial n}{\partial t}\right)_{\text{coll}}, \quad (3.6)$$

- et l'équation de conservation du moment des électrons :

$$\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v} + \frac{e\mathbf{E}}{m^*} + \frac{1}{m^*n} \nabla (nk_B\theta_n) = \left(\frac{\partial \mathbf{v}}{\partial t}\right)_{\text{coll}}, \quad (3.7)$$

où m^* est la masse effective, k_B est la constante de BOLTZMANN, θ_n est la température de l'électron et \mathbf{E} est le champ électrique.

L'hypothèse sous-jacente est que les porteurs se déplacent dans des conditions d'équilibre. Exprimant la densité totale du courant \mathbf{J} comme la somme des densités de courant des électrons

et des trous \mathbf{J}_n et \mathbf{J}_p :

$$\mathbf{J} = \mathbf{J}_n + \mathbf{J}_p, \quad (3.8)$$

où $\mathbf{J}_n = -en\mathbf{v}_n$ et $\mathbf{J}_p = ep\mathbf{v}_p$ et à partir des équations (3.5) et (3.6), les équations de continuité des électrons et des trous s'écrivent :

$$\nabla \cdot \mathbf{J}_n - e \frac{\partial n}{\partial t} = eR_g, \quad (3.9)$$

et

$$\nabla \cdot \mathbf{J}_p + e \frac{\partial p}{\partial t} = -eR_g. \quad (3.10)$$

Le modèle de dérive-diffusion utilise ensuite une forme simplifiée de l'équation (3.7), dans laquelle la température des électrons est celle du réseau à l'équilibre et son gradient le long du composant est nulle. De plus le terme $\mathbf{v} \cdot \nabla \mathbf{v}$ est en général négligé et $\frac{\partial \mathbf{v}}{\partial t}$ est choisi égal à zéro car la majorité des simulations considèrent un temps de réponse du composant très inférieur à celui du circuit. A partir de ces hypothèses et en utilisant les équations (3.5) et (3.7), la vitesse des électrons s'exprime :

$$\mathbf{v}_n = -\mu_n \mathbf{E} - \frac{D_n}{n} \nabla n, \quad (3.11)$$

où la mobilité de l'électron μ_n et son coefficient de diffusion D_n sont définis comme :

$$\mu_n = \frac{e\tau_p}{m^*}, \quad (3.12)$$

et

$$D_n = \frac{k_B \theta \mu_n}{e}, \quad (3.13)$$

θ étant la température du réseau.

L'équation de la densité de courant des électrons est alors :

$$\mathbf{J}_n = -en\mathbf{v}_n = en\mu_n \mathbf{E} + eD_n \nabla n. \quad (3.14)$$

Une équation similaire est obtenue pour les trous :

$$\mathbf{J}_p = ep\mathbf{v}_p = ep\mu_p \mathbf{E} - eD_p \nabla p, \quad (3.15)$$

où μ_p et D_p représentent respectivement la mobilité et le coefficient de diffusion des trous.

Les équations de continuité (3.9) et (3.10), les équations de densité de courant (3.14) et (3.15) ainsi que l'équation de POISSON

$$\nabla^2 V = -\frac{N_D - n + p - N_A}{\epsilon_0 \epsilon_S} \times e, \quad (3.16)$$

où V est le potentiel électrostatique, N_D la densité de dopants donneurs, N_A la densité de dopants accepteurs, n la concentration en électrons et p celle en trous, constituent le système auto-

cohérent de base du transport classique. Actuellement, la majorité des logiciels de simulation de MOSFETs commercialisés à travers le monde sont inspirés de ces équations.

B) Le modèle hydrodynamique

En réalité, les porteurs ne répondent pas immédiatement aux variations du champ électrique. Dans les nanoMOSFETs, les champs électriques internes induisent de fortes densités de porteurs et donc un échauffement de ces derniers. Un modèle plus complet, appelé modèle hydrodynamique, basé sur un système d'équations dérivé de l'équation de transport de BOLTZMANN, prend en compte les effets de relaxation de l'énergie et du moment. Dans ce modèle la propagation des électrons dans un semiconducteur est traitée comme un flux de gaz chargé, thermiquement conducteur, et plongé dans un champ électrique. Les équations du formalisme hydrodynamique, obtenues en considérant cette fois les trois premiers moments de l'équation de BOLTZMANN, représentent plus justement les effets de transport de charges hors-équilibre.

Pour les électrons, l'équation de conservation des particules s'écrit [131] :

$$\frac{\partial n}{\partial t} + \nabla \cdot (n\mathbf{v}) = -R_g. \quad (3.17)$$

L'équation de conservation du moment est :

$$\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v} + \frac{e\mathbf{E}}{m^*} + \frac{1}{m^*n} \nabla (nk_B\theta_n) = -\frac{\mathbf{v}(w)}{\tau_p}, \quad (3.18)$$

où θ_n est la température de l'électron, et τ_p est le temps de relaxation du moment. w est la somme des composantes associées à l'énergie cinétique et thermique des électrons :

$$w = \frac{1}{2}m^*v^2 + \frac{3}{2}k_B\theta_n. \quad (3.19)$$

Le premier terme l'équation (3.19) est en général négligeable devant le second.

Le troisième moment est l'équation de conservation de l'énergie qui s'écrit :

$$\frac{\partial w}{\partial t} + e\mathbf{v} \cdot \mathbf{E} + \mathbf{v} \cdot \nabla w + \frac{1}{n} \nabla (n\mathbf{v}k_B\theta_n) = -\frac{w - w_0}{\tau_w}. \quad (3.20)$$

Il est plus pratique d'exprimer ces équations en fonction de la densité de courant $\mathbf{J}_n = -en\mathbf{v}_n$ et de la mobilité $\mu_n = (e/m^*)\tau_p$. Leur expression finale est [132] :

$$\frac{1}{e} \nabla \cdot \mathbf{J}_n - R_g = 0, \quad (3.21)$$

$$\mathbf{J}_n - \frac{\tau_{pn}}{e} (\mathbf{J}_n \cdot \nabla) \frac{\mathbf{J}_n}{n} = e\mu_n \left\{ \frac{k_B\theta_n}{e} \nabla n + n \nabla \left(\frac{k_B\theta_n}{e} - \varphi \right) \right\} \quad (3.22)$$

et

$$-\nabla \cdot \left\{ \kappa_n \nabla \theta_n + \frac{\mathbf{J}_n}{e} (w + k_B \theta_n) \right\} = \mathbf{E} \cdot \mathbf{J}_n - n \frac{w - w_0}{\tau_w} - R_g w, \quad (3.23)$$

où κ_n est la conductivité thermique reliée à la mobilité μ_n par la loi de WIEDMANN-FRANZ :

$$\kappa_n = \left(\frac{5}{2} + c \right) en \mu_n \left(\frac{k_B}{e} \right)^2 \theta_n, \quad (3.24)$$

avec c un paramètre ajustable souvent choisi égal à -1.

L'équation (3.23), appelée équation de conservation de l'énergie, est la clef de ce modèle. Son terme de gauche caractérise la variation du flux d'énergie dans l'espace. Le premier terme de droite définit l'énergie absorbée par les électrons à partir du champ électrique par unité de temps et de volume. Le second terme représente l'énergie fournie par les électrons au réseau cristallin via les collisions par unité de temps et de volume. Enfin, le troisième terme indique la perte en énergie à travers les processus de recombinaison. Cette équation implique que la variation spatiale du flux d'énergie soit égal à la somme totale du flux de chaleur et de l'énergie transportée. L'application du modèle hydrodynamique aux nanoMOSFETs est bien adaptée pour prendre en compte l'augmentation de la diffusion dans les régions à haute température (que le modèle de dérive-diffusion ne peut décrire).

C) Derniers remparts avant le régime balistique

Attardons-nous un instant sur l'évolution du transport lorsque les dimensions du transistor atteignent l'échelle nanométrique. Du fait de la réduction des dimensions, les champs électriques appliqués augmentent et avec eux, la vitesse de dérive des porteurs. Lorsque le champ électrique excède 1.5×10^3 V/cm, la vitesse de dérive n'est plus proportionnelle au champ appliqué. En augmentant encore le champ, la vitesse de dérive approche celle thermique et les processus de diffusion s'intensifient : la vitesse de dérive converge alors vers un maximum appelé vitesse de saturation ($v_{d(sat)}$). Dans le silicium, les électrons présentent (à 300K) une vitesse limite de 1×10^7 cm/s sous un champ électrique de 2×10^4 V/cm. Symétriquement, les trous ont une vitesse de saturation (à 300K) de 7×10^6 cm/s quand le champ atteint 5×10^4 V/cm. De plus, des variations rapides du champ électrique ont lieu sur des distances comparables aux longueurs caractéristiques du transport, c'est-à-dire les longueurs moyennes de relaxation du moment et de l'énergie. Débute alors un régime de transport dans lequel le champ électrique n'est plus homogène dans le temps et l'espace et où la vitesse des porteurs dépasse la vitesse de saturation d'équilibre : on parle de *velocity overshoot* [133–135]. Le concept de vitesse de dérive moyenne ou de mobilité devient sans fondement. Si l'on réduit encore les dimensions, la distance à traverser par les porteurs est inférieure à la distance moyenne entre deux événements diffusants (\equiv libre parcours moyen). Une grande proportion de porteurs est alors capable de traverser tout le canal (depuis le point d'injection jusqu'au point d'extraction) sans subir de collision. Les électrons peuvent se déplacer sans diffuser, comme dans un tube à vide : le transport est balistique. Le tableau (3.1) compare les dimensions des transistors produits ces dernières décennies avec quelques

longueurs physiques caractéristiques des semiconducteurs. Les nanoMOSFETs fabriqués actuellement présentent des dimensions bien inférieures au libre parcours moyen des semiconducteurs. De plus, leurs dimensions sont désormais comparables à la longueur d'onde de DE BROGLIE ($\lambda = 2\pi/k$ et $\varepsilon = (\hbar^2 k^2 / 2m)$), grandeur à partir de laquelle les effets quantiques et la nature ondulatoire des électrons ne sont plus négligeables.

| | |
|------------------|--|
| 10 μm | Longueur de canal des FETs en 1965 |
| 1 μm | Libre parcours moyen dans les semi-conducteurs à grande mobilité |
| 100 nm | Longueur de canal des FETs commerciaux en 2003 |
| | ↓ Nanostructures ↓ |
| | Longueur d'onde de DE BROGLIE dans les semi-conducteurs |
| 10 nm | Longueur de canal des FETs ultimes R&D en 2004 |
| 1 nm | |
| 0.1 nm | Distances atomiques |

TAB. 3.1. Longueurs caractéristiques des semi-conducteurs.

Les nanotransistors doivent donc être modélisés par une théorie qui puisse tenir compte des rapides variations spatiales et temporelles du champ électrique et qui inclue les effets quantiques. Plusieurs modèles ont été proposés. BACCARANI et REGGIANI ont par exemple présenté une approche qui couple le modèle de dérive-diffusion à celui hydrodynamique [136]. Le modèle de dérive-diffusion est utilisé pour décrire les effets quantiques tandis que l'approche hydrodynamique tient compte de la suritesse. En 1997 CHANG et FOSSUM [137] ont également développé un modèle compact basé sur le second moment de l'équation de transport de BOLTZMANN afin de traiter l'effet de suritesse dans les MOSFETs ultimes. Néanmoins, la combinaison du modèle de dérive-diffusion et de l'hydrodynamique est inadaptée à la limite du transport purement balistique [138]. NATORI [139] fut le premier à présenter un modèle qui s'applique à ce type de régime en supprimant le concept de mobilité. Entre ces deux régimes extrêmes (diffusif et balistique), il existe un régime de transition dit "quasi-balistique". LUNDSTROM *et al.* [140,141] ont donc proposé modèle plus général basé sur la théorie de la diffusion et capable de modéliser une large gamme de régimes. Dans cette approche, une partie du flux incident d'électrons provenant de la source traverse le canal et est récoltée par le drain. Le flux d'électrons restant est rétro-diffusé depuis le canal et retourne vers la source. La rétro-diffusion, définie par la fraction du flux d'électrons qui retourne vers la source, dépend de la diffusion des porteurs par le réseau et de la forme du potentiel électrostatique du canal. Les modèles de NATORI et LUNDSTROM *et al.* ne tiennent cependant pas compte de l'effet tunnel à travers la barrière de potentiel du canal, entre la source et le drain. Ce dernier s'intensifie pour des longueurs de canal inférieures à 15 nm et a été inclu par la suite par plusieurs groupes. Citons tout d'abord le modèle quasi-2D de PIROVANO *et al.* [142] qui couple le système auto-cohérent SCHRÖDINGER-POISSON aux équations semi-classiques de dérive-diffusion. PICUS et LIKHAREV [143] ont également proposé un modèle analytique pour décrire des MOSFETs de 10 nm. Plus récemment, NAVEH et LIKHAREV [110] ont présenté un modèle décrivant aussi l'effet tunnel à travers l'oxyde de grille à partir d'une approche WKB. En 2003, MUNTEANU et AUTRAN [144] développèrent le code BALMOS couplant

de façon auto-cohérente l'équation 2D de SCHRÖDINGER à celle de POISSON. Ce code peut tenir compte de la géométrie du canal et des conditions de polarisation. Une toute autre approche, la théorie du gradient de densité (DG theory pour l'anglais "Density Gradient"), initialement élaborée par ANCONA *et al.* [145], constitue un traitement généralisé du modèle de dérive-diffusion. Elle incorpore les effets quantiques en rendant les équations d'état des électrons et des trous dépendantes non seulement des densités des porteurs, mais également du gradient de ces densités. Par cette modification, la théorie est capable de modéliser le confinement quantique [146–150] et l'effet tunnel [151]. Récemment, LAUX *et al.* [152] ont proposé une méthode basée sur la résolution directe de l'équation de SCHRÖDINGER couplée à celle de POISSON permettant de modéliser quantiquement des composants 3D aux géométries variées dans lesquels les mécanismes de diffusion peuvent être inclus via la formulation des équations maîtresses [153, 154]. Enfin, SVIZHENKO *et al.* [155, 156] ont présenté une simulation 2D d'un MOSFET double-grille basée sur le formalisme des fonctions de GREEN hors-équilibre. Cette approche, numériquement améliorée par VENUGOPAL *et al.* [104, 157, 158], a été successivement appliquée pour traiter la diffusion via de la théorie des sondes de BÜTTIKER [159] et à également permis de vérifier l'influence de la forme des réservoirs (source et drain) sur le courant. Plus récemment encore, XIA *et al.* [160] ont couplé ce formalisme avec une approche atomique de liaisons fortes (type *sp3s**) pour décrire un transistor double grille. La méthode des fonctions de GREEN hors-équilibre présente également l'avantage de pouvoir s'étendre à toute sorte de composants non-conventionnels comme les nanotubes de carbone et autres transistors moléculaires. Dans la suite de ce chapitre, nous allons expliquer les capacités de ce formalisme en étudiant le transport d'un composant électronique très général de taille nanométrique.

3.1.2 Description physique d'un nanocomposant : cas général

Considérons un composant électronique connecté à deux électrodes dont les potentiels électrochimiques sont μ_G (gauche) et μ_D (droite). Lorsque μ_G et μ_D diffèrent sous l'effet d'une polarisation externe ($\mu_G - \mu_D = -eV_{DG}$), le nano-composant est dans un état hors-équilibre et un flux d'électrons traverse alors le système (voir figure(3.1)). Les deux électrodes sont des fils conducteurs macroscopiques qui peuvent être modélisées par des métaux ou des semi-conducteurs semi-infinis. Les réservoirs sont considérés suffisamment grands pour que les potentiels électrochimiques du matériau massif (μ_G et μ_D) ne soient pas modifiés par le courant I . Il est également envisageable de décrire le comportement d'une zone active (partie hors équilibre) connectées à deux contacts microscopiques, eux-mêmes couplés à des réservoirs macroscopiques agissant comme de parfaits conducteurs. Dans ce cas, il est nécessaire d'inclure une partie du réservoir dans la zone active. Plus précisément, la partie des réservoirs à inclure est définie par la distance sur laquelle la densité d'électrons n'atteint pas celle du matériau massif [104]. Si les conducteurs sont métalliques, ces régions diminuent du fait du fort écrantage des champs électriques et l'on retrouve le premier cas.

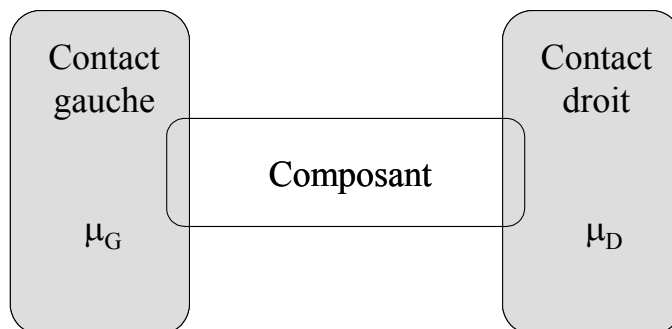


FIG. 3.1: Nano-dispositif couplé à deux électrodes semi-infinies (gauche et droite) avec des potentiels électrochimiques μ_G et μ_D .

3.2 Les limites du couplage faible

Soit un nano-composant faiblement couplé à deux électrodes auquel peut s'appliquer la théorie des perturbations. Dans de nombreux problèmes, cette approximation est bien justifiée et constitue la base d'importants développements théoriques qui décrivent la conduction au travers d'îlots métalliques ou semi-conducteurs [121].

3.2.1 La théorie des perturbations

Étudions le transfert de particules entre deux électrodes (figure(3.2)). L'hamiltonien du système peut s'écrire $H = H_0 + V$, où H_0 est l'hamiltonien des électrodes isolées et V représente leur couplage lorsqu'un nano-composant est disposé entre les deux. Les électrodes de gauche et de droite étant respectivement caractérisées par des quasi-continuum d'états $|i\rangle$ et $|j\rangle$, la règle d'or de FERMI décrit le taux de transfert (probabilité par unité de temps) d'un électron entre deux états :

$$W_{ij} = \frac{2\pi}{\hbar} |V_{ij}|^2 \delta(\varepsilon_i - \varepsilon_j), \quad (3.25)$$

où $V_{ij} = \langle i|V|j\rangle$ et ε_i (ε_j) est l'énergie de l'état $|i\rangle$ ($|j\rangle$). Le courant I est alors donné par la différence entre le flux d'électrons provenant de la gauche et celui provenant de la droite. En sommant sur tous les états et prenant en compte l'occupation des niveaux d'énergies, nous obtenons :

$$\begin{aligned} I = & -(e) \sum_{i \in G, j \in D} W_{ij} f(\varepsilon_i - \mu_G) [1 - f(\varepsilon_j - \mu_D)] \delta(\varepsilon_i - \varepsilon_j) \\ & - (-e) \sum_{i \in G, j \in D} W_{ij} f(\varepsilon_j - \mu_D) [1 - f(\varepsilon_i - \mu_G)] \delta(\varepsilon_i - \varepsilon_j), \end{aligned} \quad (3.26)$$

où f est la fonction de distribution de FERMI-DIRAC :

$$f(\varepsilon) = \left[1 + \exp\left(\frac{\varepsilon}{k_B \theta}\right) \right]^{-1}, \quad (3.27)$$

θ étant la température. L'équation (3.26) se simplifie en l'expression suivante :

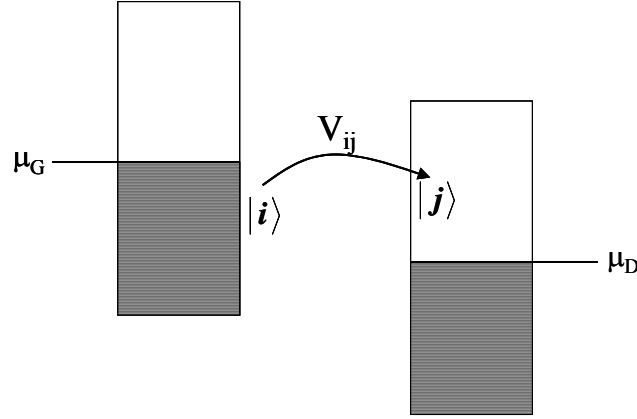


FIG. 3.2: Un électron a une faible probabilité d'être transféré via le couplage $V_{ij} = \langle i | V | j \rangle$, d'un état occupé $|i\rangle$ de l'électrode gauche vers un état vide $|j\rangle$ de l'électrode droite : la théorie des perturbations peut s'appliquer.

$$I = \frac{2\pi e}{\hbar} \sum_{i \in G, j \in D} |V_{ij}|^2 \{f(\varepsilon_j - \mu_D) - f(\varepsilon_i - \mu_G)\} \delta(\varepsilon_i - \varepsilon_j). \quad (3.28)$$

Introduisant la densité d'états des électrodes de gauche et de droite :

$$n_G(\varepsilon) = \sum_{i \in G} \delta(\varepsilon - \varepsilon_i), \quad n_D(\varepsilon) = \sum_{j \in D} \delta(\varepsilon - \varepsilon_j), \quad (3.29)$$

nous obtenons la formule usuelle du courant [164] :

$$I = \frac{2\pi e}{\hbar} \int |V(\varepsilon)|^2 \{f(\varepsilon - \mu_D) - f(\varepsilon - \mu_G)\} n_G(\varepsilon) n_D(\varepsilon) d\varepsilon, \quad (3.30)$$

où $V(\varepsilon)$ est le couplage des états à l'énergie ε .

Cette expression fut utilisée pour calculer le courant dans les jonctions métal-isolant-métal [163] ainsi qu'en microscopie tunnel [165,166]. En particulier, BARDEEN a établi des lois générales pour calculer les éléments de la matrice de couplage V_{ij} quand la région centrale est assimilable à une barrière isolante [164]. Dans la limite d'une barrière épaisse, l'approximation WENTZEL-KRAMERS-BRILLOUIN (W.K.B) donne :

$$V_{ij} \propto \exp(-kd), \quad \text{avec } k = \frac{\sqrt{2m_0 E_B}}{\hbar} \quad (3.31)$$

où E_B est la hauteur de barrière et d son épaisseur (figure(3.3)). On obtient donc que le courant dans la jonction varie linéairement selon la loi $\exp(-2kd)$ en fonction de l'épaisseur associée à l'isolant. Il s'agit d'une relation très importante, d'abord parce qu'elle propose une solution approchée aux problèmes quantiques qui peut s'avérer très utile dans certains cas notamment en physique des dispositifs, ensuite d'un point de vue fondamental, parce que cette théorie fait le lien entre la mécanique quantique et la mécanique classique (on parle d'approche *semi-classique*). De la mécanique quantique, l'approximation W.K.B permet de rendre compte des effets bien particuliers tels que l'effet tunnel ou le confinement. Mais certains des comportements quantiques

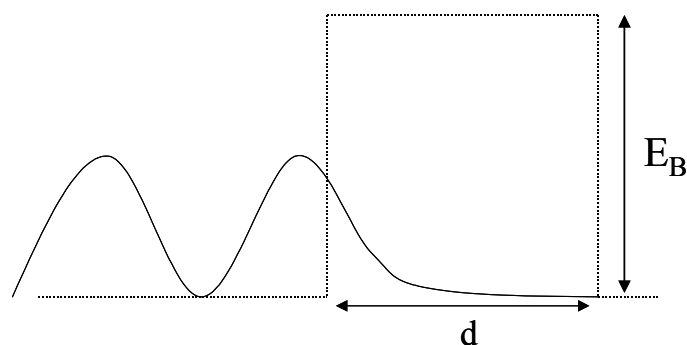


FIG. 3.3: *Barrière d'énergie dans une jonction métal-isolant-métal.*

essentiels ne peuvent être correctement décrits par le cas semi-classique. On peut montrer par exemple que les conditions de raccordement semi-classiques ne sont pas toujours réalisées quand elles sont appliquées à des potentiels réalistes. De plus, les cas où les nano-composants peuvent être réduits à une barrière isolante sont rares. Par exemple, les jonctions métal-molécule-métal, dans lesquelles les niveaux électroniques se situent entre les potentiels électrochimiques μ_G et μ_D , ne peuvent être décrites par cette approche.

3.2.2 La théorie orthodoxe du courant tunnel

Cette théorie décrit les nano-systèmes couplés à des électrodes via des jonctions dites "tunnel" dans lesquelles le faible couplage peut être traité au premier ordre en théorie des perturbations (règle d'or de FERMI). La jonction tunnel est une représentation "circuit" de l'effet tunnel existant entre les électrodes et le nano-dispositif. Les paramètres qui décrivent les caractéristiques électriques d'une telle jonction sont sa capacité, C , et sa résistance tunnel, R_t . En dehors de ses propriétés liées à l'effet tunnel, la jonction tunnel se comporte comme une capacité. Cette approche fut développée par KULIK ET SHEKHTER [167] afin de modéliser le comportement électrique d'îlots métalliques. Ce n'est que par la suite qu'elle fut adaptée aux boîtes quantiques de semiconducteurs et aux molécules, dont les niveaux d'énergie sont discrets [169]. La théorie orthodoxe suppose un couplage très faible, de sorte que la densité d'états ne soit pas modifiée¹ et que le courant puisse être par conséquent, calculé en utilisant les équations maîtresses [168]. Afin d'illustrer la physique de base de cette théorie, nous allons tout d'abord traiter le cas simple d'un îlot placé entre deux électrodes.

A) Cas de l'îlot isolé à un seul état ε_0

Considérons un îlot suffisamment petit pour ne contenir qu'un seul état d'énergie ε_0 situé entre μ_G et μ_D . Les taux de transfert à travers les jonctions de gauche et de droite sont alors

¹Si la hauteur des barrières entre les réservoirs et la zone active est grande (1 à 3 eV) la transparence et la densité d'états dans l'îlot deviennent quasi-indépendantes de l'énergie.

définis par W_G et W_D respectivement (figure (3.4)). En appliquant la règle d'or de FERMI, nous pouvons écrire :

$$\begin{aligned} W_G &= \frac{2\pi}{\hbar} \sum_{i \in G} |V_{i0}|^2 \delta(\varepsilon_i - \varepsilon_0) = \frac{\Gamma^G(\varepsilon_0)}{\hbar}, \\ W_D &= \frac{2\pi}{\hbar} \sum_{j \in D} |V_{j0}|^2 \delta(\varepsilon_j - \varepsilon_0) = \frac{\Gamma^D(\varepsilon_0)}{\hbar}, \end{aligned} \quad (3.32)$$

où Γ^G et Γ^D ont la dimension d'une énergie et décrivent le couplage du niveau ε_0 avec les électrodes². Notons que $1/W$ est le temps caractéristique, τ , que passe un électron sur ce niveau : pour de forts taux de transfert, l'électron passe "rapidement" d'un réservoir à l'autre : τ est très court. Lorsque au contraire les taux de transfert sont faibles, τ est élevé, et l'on peut définir exactement la charge présente dans l'îlot comme un multiple de e .

Le courant traversant les jonctions de gauche et de droite est exprimé par :

$$\begin{aligned} I_G &= (-e) [f(\varepsilon_0 - \mu_G) - f_0] W_G, \\ I_D &= (-e) [f_0 - f(\varepsilon_0 - \mu_D)] W_D, \end{aligned} \quad (3.33)$$

où f_0 est l'occupation moyenne de l'état ε_0 . En régime permanent, la conservation des charges suppose que $I = I_G = I_D$ et nous pouvons déduire :

$$f_0 = \frac{W_G \cdot f(\varepsilon_0 - \mu_G) + W_D \cdot f(\varepsilon_0 - \mu_D)}{W_G + W_D}. \quad (3.34)$$

Injectant (3.34) dans (3.33) et utilisant (3.32), le courant total peut se réécrire :

$$I = \frac{e}{\hbar} [f(\varepsilon_0 - \mu_D) - f(\varepsilon_0 - \mu_G)] \cdot \frac{\Gamma^G(\varepsilon_0) \cdot \Gamma^D(\varepsilon_0)}{\Gamma^G(\varepsilon_0) + \Gamma^D(\varepsilon_0)}. \quad (3.35)$$

L'expression précédente ainsi que l'équation (3.34) montrent que la différence entre les potentiels chimiques des 2 réservoirs crée un flux continu d'électrons à travers le niveau ε_0 dont l'occupation f_0 est un mélange de $f(\varepsilon_0 - \mu_D)$ et $f(\varepsilon_0 - \mu_G)$ [170, 173].

B) Cas général

Etudions le transfert tunnel d'électrons à l'énergie ε à travers les jonctions de gauche et de droite avec les taux respectifs $\Gamma^G(\varepsilon)/\hbar$ et $\Gamma^D(\varepsilon)/\hbar$. La théorie orthodoxe suppose que $\Gamma^G(\varepsilon)$, $\Gamma^D(\varepsilon) \ll U$ et $k_B\theta \ll U$, où U est l'énergie moyenne de la charge de COULOMB dans la nano-structure. U correspond à l'énergie moyenne des interactions répulsives entre les électrons d'états distincts de l'îlot. A chaque instant, la charge totale q dans l'îlot est bien définie et doit être un multiple entier de la charge de l'électron e . Pour une charge donnée q , la nano-structure peut être dans différentes configurations électroniques (indexées par n) caractérisées par l'énergie totale

²Le fait de ne considérer qu'un niveau d'énergie dans l'îlot sous-entend : $\Gamma^G, \Gamma^D \ll \Delta\varepsilon$, où $\Delta\varepsilon$ représente l'écart moyen entre deux niveaux d'énergie consécutifs.

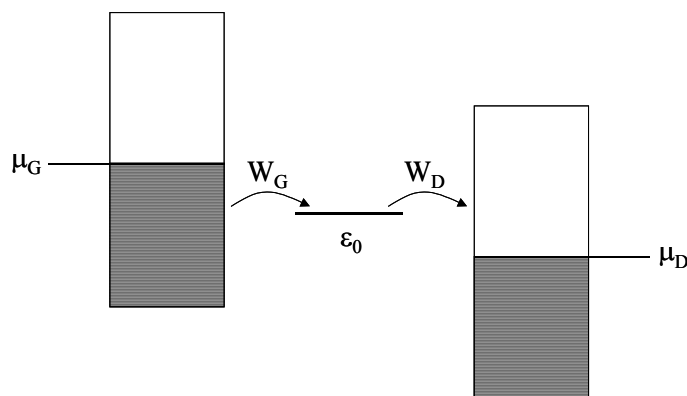


FIG. 3.4: *Transfert d'électrons à travers un îlot contenant un seul niveau ε_0 .*

$\varepsilon_n(q, V_{DG})$. La connaissance de ces énergies et des taux de transfert détermine complètement la courbe $I(V_{DG})$ [167, 169].

C) Îlots métalliques et effets du blocage de coulomb

L'application de la théorie orthodoxe aux îlots métalliques a été intensivement décrite dans la littérature [121, 175]. Si l'îlot n'est pas trop petit, les effets du confinement quantique peuvent être négligés et le transport est dominé par les effets de blocage de Coulomb. Contrairement à l'effet tunnel, le blocage de COULOMB est un phénomène classique, qui résulte des répulsions électron-électron entre les $(q - 1)$ électrons présents dans l'îlot et un électron extérieur. Ces interactions ne permettent à l'électron isolé de pénétrer dans la boîte, que lorsque la "dégénérescence d'états" est satisfaite, c'est-à-dire lorsque l'énergie totale de la boîte à $q - 1$ particules est égale à l'énergie de la boîte à q particules. Cette condition impose des niveaux de transition déterminés :

$$\varepsilon(q|q-1, V_{DG}) = \varepsilon_n(q-1) - \varepsilon_n(q) \quad (3.36)$$

$$= \left(-q + \frac{1}{2}\right)U - \eta e V_{DG}. \quad (3.37)$$

L'intervalle entre chaque pallier des courbes de $I(V_{DG})$ est donc constant et proportionnel à l'énergie de charge U . Ainsi, dès que le potentiel extérieur augmente de U , le nombre d'électrons dans le système augmente d'une unité.

Le transistor à un électron (ou **SET**) est le dispositif emblématique de ce phénomène (figure(3.5)). Il s'agit d'un composant à 3 électrodes : les 2 réservoirs et la grille. L'équivalent du canal d'un MOSFET conventionnel est ici l'îlot central qui est séparé des réservoirs par des jonctions tunnel. Le potentiel dans l'îlot est contrôlé de façon capacitive par la tension de grille. Il en résulte un nouveau terme dans l'équation (3.36) $-e\varphi_G$ où φ_G est une fonction linéaire de la tension de grille et de la capacité de couplage. Ainsi, en variant la tension de grille, les niveaux d'ionisation

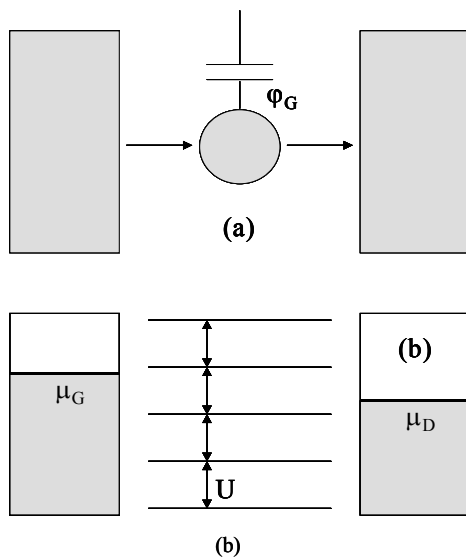


FIG. 3.5: (a) Transistor à un électron avec un îlot métallique. (b) Les niveaux ionisés de l'îlot sont régulièrement espacés par l'énergie de charge U .

sont décalés et le courant présente des oscillations de COULOMB, i.e. une dépendance périodique de la conductance avec φ_G .

3.3 Au delà de la théorie des perturbations

Jusqu'à présent, nous avons supposé que le couplage (faible) entre les contacts et la nanostructure pouvait être traité en théorie des perturbations. Dans le cas d'îlots métalliques, cette approche reste valide tant que $\Gamma^G, \Gamma^D \ll U$, où U est l'énergie moyenne de la charge de COULOMB. Dans les molécules ou les nanostructures semiconductrices, on doit imposer une condition supplémentaire, à savoir : $\Gamma^G, \Gamma^D \ll \Delta\varepsilon$, où $\Delta\varepsilon$ est l'écart moyen entre les niveaux d'énergie d'états quantiquement confinés. Cependant, nombreux sont les systèmes où les paramètres de couplage Γ^G, Γ^D sont du même ordre de grandeur et parfois plus grands que U et $\Delta\varepsilon$. Il apparaît donc nécessaire d'adopter une théorie applicable à toutes les forces de couplage. S'ajoute à ce contexte, le besoin de traiter, jusqu'à un certain degré, les effets de corrélation, en particulier de décrire les effets de charge quand $U \approx \Gamma^G, \Gamma^D$. La théorie de LANDAUER [122, 176, 177] fut la première à tenter de réaliser ce souhait en décrivant le transport électronique dans les systèmes mésoscopiques. Le point essentiel de la théorie de LANDAUER est que le conducteur est vu comme un centre diffuseur qui va transmettre une partie des électrons et réfléchir l'autre partie. Dans une première étape, nous allons expliciter la formule de LANDAUER avant de détailler l'approche équivalente implémentée en fonctions de GREEN [170]. Cette dernière, basée sur le formalisme des fonctions de GREEN hors-équilibre, est en principe susceptible d'inclure tous les effets souhaités. Le formalisme des fonctions de GREEN (aussi appelé formalisme de KELDYSH [178] ou de KADANOFF-BAYM [179]), offre de surcroît un cadre conceptuel propice à la description d'interactions électron-électron et électron-photon [182, 183].

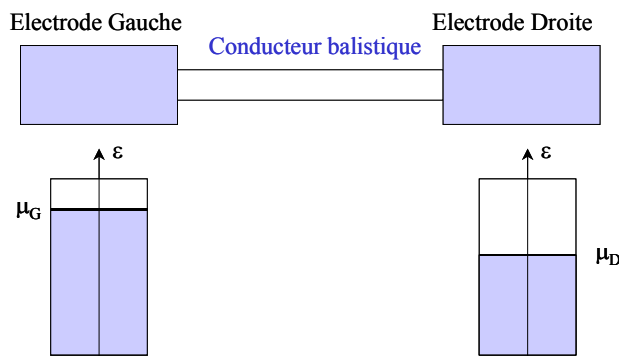


FIG. 3.6: Configuration d'un système mésoscopique balistique. Chaque réservoir, défini par son potentiel électrochimique, thermalise parfaitement tous les électrons incidents.

3.3.1 Formule de Landauer

A) Conducteur balistique

Considérons le cas fréquent d'un conducteur encadré par deux grands contacts électriques (figure(3.6)). Les réservoirs sont supposés présenter des propriétés équivalentes à celle du corps noir en optique : ils thermalisent parfaitement les électrons incidents en les absorbant sans leur laisser aucune chance de se réfléchir. La distribution d'électrons hors-équilibre disparaît par le biais de collisions inélastiques qui ont lieu au sein des réservoirs. En revanche, le transport dans le conducteur est balistique, c'est-à-dire cohérent et élastique. Nous supposons de plus dans un premier temps, qu'il n'y a pas de réflexion interne : le conducteur est parfait et la probabilité qu'un électron soit transmis d'une extrémité à une autre vaut 1. Le courant peut être alors calculé en considérant un mode particulier du système à une dimension $|i(k)\rangle$ se propageant vers la droite, occupé selon la fonction de distribution du réservoir gauche $f(\varepsilon - \mu_G)$. Sachant qu'un gaz d'électrons libres, avec n électrons par unité de longueur, se déplaçant à une vitesse v , transporte un courant égal à $-nev$ et que la densité électronique de cet état $|i(k)\rangle$, pour un conducteur de longueur L , est $1/L$, le courant transporté par cet état est :

$$I^+ = -\frac{e}{L} \sum_i v_i f(\varepsilon_i - \mu_G) = -\frac{e}{L} \sum_i \frac{1}{\hbar} \left(\frac{\partial \varepsilon}{\partial k} \right)_i f(\varepsilon_i - \mu_G). \quad (3.38)$$

Si nous transformons la somme sur les états $|i\rangle$ par une intégrale dans l'espace des modes³ :

$$\sum_k \rightarrow \frac{L}{2\pi} \int dk, \quad (3.39)$$

nous obtenons :

$$I^+ = -\frac{e}{h} \int_{-\infty}^{+\infty} f(\varepsilon - \mu_G) d\varepsilon. \quad (3.40)$$

³Les états d'un système 1D de longueur L sont régulièrement espacés tous les $2\pi/L$ dans l'espace des modes [187].

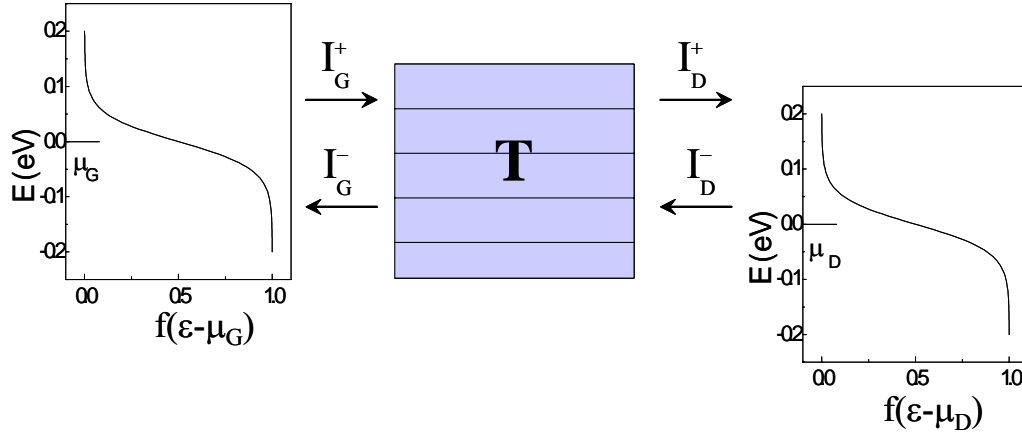


FIG. 3.7: *Bilan des courants traversant un conducteur non-transparent.*

Enfin, considérant un nombre de modes N [188], le courant partiel provenant du réservoir gauche devient :

$$I^+ = -\frac{e}{h} \int_{-\infty}^{+\infty} N f(\varepsilon - \mu_G) d\varepsilon. \quad (3.41)$$

Une définition similaire des états se propageant vers la gauche (émis par le réservoir de droite) donne le courant partiel I^- :

$$I^- = +\frac{e}{h} \int_{-\infty}^{+\infty} N f(\varepsilon - \mu_D) d\varepsilon. \quad (3.42)$$

Finalement, le courant global est égal à la somme des 2 composantes :

$$I = I^+ + I^- = -\frac{e}{h} \int_{-\infty}^{+\infty} N (f(\varepsilon - \mu_G) - f(\varepsilon - \mu_D)) d\varepsilon. \quad (3.43)$$

B) Conducteur diffuseur

Si le conducteur n'est maintenant plus parfait, la probabilité T pour que l'électron injecté à une extrémité du conducteur soit transmis jusqu'à l'autre extrémité n'est plus égale à 1. Par conséquent, nous pouvons définir la probabilité $(1 - T)$ pour que l'électron soit réfléchi par le conducteur. Pour obtenir la nouvelle formule du courant, il nous faut effectuer un bilan des différents courants dans le système (figure(3.7)). Au niveau de l'électrode gauche, le courant par unité d'énergie circulant vers la droite s'écrit :

$$I_G^+ = -\frac{e}{h} N f(\varepsilon - \mu_G). \quad (3.44)$$

De même, le courant, par unité d'énergie, du contact droit circulant vers la gauche est :

$$I_D^- = \frac{e}{h} N f(\varepsilon - \mu_D). \quad (3.45)$$

Ces deux composantes vont ensuite interagir avec le conducteur. Une partie de chaque courant partiel sera transmise et la partie restante sera réfléchiée. Il en résulte :

$$I_G^- = (1 - T) I_G^+ + T I_D^- \quad (3.46)$$

$$I_D^+ = T I_G^+ + (1 - T) I_D^- \quad (3.47)$$

Le courant global traversant le système vaut :

$$I = I_G^+ - I_G^- = I_D^+ - I_D^- \quad (3.48)$$

En intégrant sur l'énergie nous obtenons :

$$I = -\frac{e}{h} \int_{-\infty}^{+\infty} NT(\varepsilon) (f(\varepsilon - \mu_G) - f(\varepsilon - \mu_D)) d\varepsilon. \quad (3.49)$$

L'équation (3.49) est la formule généralisée de LANDAUER qui est utilisée pour décrire le transport dans de nombreux systèmes mésoscopiques.

C) Le quantum de conductance

Considérons de nouveau un fil quantique 1D parfaitement conducteur. Chaque extrémité du conducteur est connectée à un réservoir d'électrons dont la résistance de contact est nulle ($T = 1$ dans la figure(3.7)). Ce système ne possède a priori aucune résistance et sa conductance devrait être infinie. Utilisant la formule de LANDAUER (3.49), le courant s'exprime de la façon suivante :

$$I = -\frac{e}{h} \int_{-\infty}^{+\infty} [f_G(\varepsilon - \mu_G) - f_D(\varepsilon - \mu_D)] d\varepsilon. \quad (3.50)$$

Dans une nano-structure idéalement balistique dont les résistances de contact sont nulles, I dépend essentiellement de la fonction de distribution de FERMI-DIRAC de chaque réservoir. Cette fonction impose que chaque état d'énergie possède au maximum une seule particule : il s'agit d'une propriété fondamentale des particules fermioniques dont font partie les électrons. Appliquons maintenant une différence de potentiel entre les deux réservoirs $-eV_{DG} = \mu_G - \mu_D$. Un flux d'électrons traverse le système et la conductance D peut s'exprimer de la façon suivante ($D = I/V_{DG}$) :

$$D = -\frac{e}{h} \int_{-\infty}^{+\infty} \frac{1}{1 + \exp[(\varepsilon - \mu_D + eV_{DG})/k\theta]} - \frac{1}{1 + \exp[(\varepsilon - \mu_D)/k\theta]} d\varepsilon. \quad (3.51)$$

Lorsque $V_{DG} \rightarrow 0$, l'équation (3.51) devient :

$$D = -\frac{e^2}{h} \int_{-\infty}^{+\infty} \frac{\partial f(\varepsilon)}{\partial \varepsilon} d\varepsilon, \quad (3.52)$$

pour finalement se réduire à une constante :

$$D_0 = \frac{e^2}{h}. \quad (3.53)$$

L'équation (3.53) s'applique à toute température et définit le quantum de conductance. D'où vient la résistance du système? La résistance est tout d'abord la conséquence de la constriction des réservoirs infinis en un système 1D dans lequel seule une quantité finie de courant peut circuler. Elle est définie ensuite par la fonction de distribution des réservoirs qui fixe la valeur maximale du courant que peut transmettre un canal de conduction (ou mode). La valeur du quantum de conductance résulte en effet de la statistique de FERMÍ-DIRAC. La quantité d'électrons injectés dans un système balistique 1D est limitée par la vitesse des particules à l'équilibre dans les réservoirs (\approx vitesse de FERMÍ, lorsque l'intervalle d'énergie des électrons participants à la conduction est étroit) et par le nombre moyen d'électrons par niveau d'énergie et unité de temps. Une relation similaire est vérifiée pour des particules bosoniques. Dans ce cas, la distribution de BOSE-EINSTEIN remplace celle de FERMÍ-DIRAC et le quantum de conductance de vient :

$$G_b = \frac{N_b e^2}{h}, \quad (3.54)$$

où N_b n'est en général pas limité à 1 et peut prendre des valeurs bien supérieures.

Dans les systèmes macroscopiques où le transport est diffusif, le potentiel varie linéairement le long du matériau conducteur, en suivant la loi d'OHM. Dans les nano-composants ou les fils moléculaires, le matériau adopte un comportement qui ne peut être interprété sans quantifier la conductance [184, 185] : le courant n'augmente plus linéairement avec la tension mais évolue par paliers successifs. La frontière entre ces deux limites est la longueur de cohérence L_φ qui définit l'échelle mésoscopique séparant le transport classique du transport quantique. Un conducteur macroscopique est en fait une collection de conducteurs mésoscopiques de taille L_φ . La résistance d'un échantillon macroscopique peut être obtenue en additionnant classiquement la résistance de chaque domaine mésoscopique, tandis qu'à l'intérieur d'un domaine mésoscopique, la loi d'OHM n'est plus valable car les propriétés ne sont plus locales.

3.3.2 Intérêt de la matrice densité

Quelle est la densité d'électrons, n d'un composant connecté à deux contacts avec deux potentiels électrochimiques différents? Nous pourrions supposer dans un premier temps que les niveaux d'énergie soient occupés avec une probabilité f_α de valeur intermédiaire entre la fonction de FERMÍ-DIRAC de chaque réservoir (comme dans la théorie Orthodoxe [Eq.(3.34)]) :

$$n(\mathbf{r}) = \sum_{\alpha} \Psi_{\alpha}(\mathbf{r}) \Psi_{\alpha}^*(\mathbf{r}) f_{\alpha}. \quad (3.55)$$

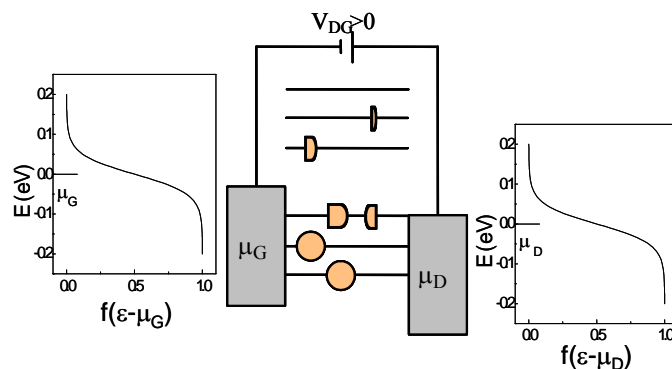


FIG. 3.8: Application d'une tension positive sur le réservoir de droite : les contacts gauche et droite veulent imposer leur distribution de FERMÍ-DIRAC aux électrons de la zone active.

Malheureusement, la réponse est en générale plus compliquée. Plusieurs états peuvent être occupés de façon corrélée via une matrice appelée "matrice densité", $\rho_{\alpha\beta}$:

$$n(\mathbf{r}) = \sum_{\alpha,\beta} \Psi_{\alpha}(\mathbf{r}) \Psi_{\beta}^*(\mathbf{r}) \rho_{\alpha\beta} \quad (3.56)$$

Le point central de la mécanique statistique hors-équilibre est de déterminer la matrice densité $\rho_{\alpha\beta}$. Une fois trouvée, toutes les quantités du système (charge, courant, etc...) peuvent être calculées.

A) Matrice densité et opérateur de courant

La différence entre la probabilité d'occupation f_{α} et la matrice densité $\rho_{\alpha\beta}$ ressort en considérant un conducteur à une dimension parfaitement conducteur ($T = 1$) connecté à deux contacts dont les potentiels électrochimiques sont μ_G et μ_D (figure (3.8)). Si une tension positive est appliquée entre la source et le drain, le drain aura un potentiel électrochimique inférieur à celui de la source : $\mu_D = \mu_G - eV_{DG}$. Les réservoirs ont des fonctions de FERMÍ-DIRAC différentes et chacun cherche à imposer son équilibre au composant. La zone active est donc "coincée" entre deux contacts aux volontés antagonistes qui le maintiennent dans un état intermédiaire hors-équilibre : il s'agit du régime balistique. Comme nous l'avons montré dans le paragraphe (3.3.1), un des célèbres résultats de la physique mésoscopique est la quantification de la conductance d'un conducteur balistique 1D :

$$g = \frac{I}{(\mu_G - \mu_D) / (-e)} = \frac{2e^2}{h}. \quad (3.57)$$

Ici $g = 2D_0$ (équation (3.53)) en raison des contributions identiques d'électrons de spin opposé dans chaque canal de conduction. Afin d'illustrer l'intérêt de la matrice densité, nous allons redémontrer cette équation. Les vecteurs propres d'un conducteur 1D peuvent s'écrire sous la

forme (L : longueur de la zone active) :

$$\Psi_{+k}(x) = e^{+ikx}/\sqrt{L} \quad \Psi_{-k}(x) = e^{-ikx}/\sqrt{L}, \quad (3.58)$$

et les énergies associées sont dans l'approximation de la bande parabolique :

$$\varepsilon_{+k} = \varepsilon_{-k} = E_c + (\hbar^2 k^2 / 2m), \quad (3.59)$$

où E_c est le bas de la bande de conduction. Les états $+k$ sont occupés par les électrons venant du contact gauche, alors que les états $-k$ sont occupés par les électrons venant du contact de droite. Le transport étant cohérent, les états $+k$ et $-k$ restent à l'équilibre thermodynamique avec leur réservoir d'origine, c'est-à-dire qu'ils sont distribués en énergie selon la fonction de distribution de FERMI-DIRAC des contacts de gauche et de droite respectivement :

$$f_{+k} = f(\varepsilon_{+k} - \mu_G) \quad f_{-k} = f(\varepsilon_{-k} - \mu_D) \quad (3.60)$$

En notant que la densité de courant apportée par une onde plane [Eq. (3.58)] est donnée par⁴ :

$$J_{+k} = (\hbar k / mL) \quad J_{-k} = (-\hbar k / mL), \quad (3.61)$$

nous obtenons l'expression du courant :

$$I = 2 \text{ (spin)} \times (-e) \sum_{k>0} J_{+k} f(\varepsilon - \mu_G) + J_{-k} f(\varepsilon - \mu_D) \quad (3.62)$$

$$= (-2e) \sum_k \frac{\hbar k}{mL} [f(\varepsilon - \mu_G) - f(\varepsilon - \mu_D)]. \quad (3.63)$$

Passant de la somme discrète à l'intégrale ($\sum_k \rightarrow \int dk L/2\pi$) nous pouvons écrire (à 0K) :

$$I = \frac{-2e}{h} \int_0^\infty d\varepsilon_k [f(\varepsilon_k - \mu_G) - f(\varepsilon_k - \mu_D)] \quad (3.64)$$

$$= \frac{-2e}{h} (\mu_G - \mu_D), \quad (3.65)$$

d'où la relation (3.57).

Cependant, la forme des vecteurs propres n'est pas unique. Les états $+k$ et $-k$ ont la même énergie et toute combinaison linéaire des deux reste un vecteur propre de la région active. Nous pouvons donc choisir les états suivants :

$$\Psi_{c,k} = \sqrt{2/L} \cos kx \quad \text{et} \quad \Psi_{s,k} = \sqrt{2/L} \sin kx. \quad (3.66)$$

Quelle que soit l'occupation de ces états le courant total sera nul, puisque le sinus et le cosinus sont des superpositions d'états $+k$ et $-k$: ils ne transportent individuellement pas de courant

⁴L'expression de la vitesse de groupe : $v_G = (1/\hbar) \cdot (\partial\varepsilon/\partial k)$

($J_{c,k} = J_{s,k} = 0$) :

$$I = 2 \text{ (spin)} \sum_{k>0} J_{c,k} f_{c,k} + J_{s,k} f_{s,k} = 0. \quad (3.67)$$

Où est la faille? En fait, si l'on choisit d'utiliser les états en cosinus et sinus, nous ne pouvons pas calculer le courant à partir des probabilités d'occupation puisque l'opérateur de courant n'est pas diagonal dans cette représentation. Ecrivons tout d'abord l'opérateur de courant dans la base des ondes planes ($J_{op} = (1/\hbar)(\partial\varepsilon/\partial k)$) :

$$[J_{op}]_{o.p.} = \begin{bmatrix} +k & -k \\ \hbar k/mL & 0 \\ 0 & -\hbar k/mL \end{bmatrix}. \quad (3.68)$$

Représentons ensuite l'opérateur de courant dans la base des sinus et cosinus à l'aide de la matrice unitaire⁵ de changement de base $[V]$ dont les colonnes représentent l'ancienne base (ondes planes, "o.p.") en fonction de la nouvelle base (cosinus-sinus, "cs") :

$$[V] = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ +i & -i \end{bmatrix}. \quad (3.69)$$

Nous obtenons :

$$[J_{op}]_{cs} = [V] [J_{op}]_{o.p.} [V]^\dagger = \begin{bmatrix} 0 & -i\hbar k/mL \\ +i\hbar k/mL & 0 \end{bmatrix}. \quad (3.70)$$

Le éléments diagonaux dans cette représentation sont nuls, indiquant qu'aucun état n'engendre du courant par lui-même. En revanche, s'ils sont corrélés par les termes non-diagonaux, un courant pourra exister. Dans la base des ondes planes, la matrice densité est diagonale et les éléments représentent les probabilités d'occupation f_α ,

$$[\rho]_{o.p.} = \begin{bmatrix} f_{+k} & 0 \\ 0 & f_{-k} \end{bmatrix} \quad (3.71)$$

de telle sorte que la représentation en sinus et cosinus s'écrive :

$$[\rho]_{cs} = [V] [\rho]_{o.p.} [V]^\dagger = \frac{1}{2} \begin{bmatrix} (f_{+k} + f_{-k}) & -i(f_{+k} - f_{-k}) \\ i(f_{+k} - f_{-k}) & (f_{+k} + f_{-k}) \end{bmatrix}. \quad (3.72)$$

Le courant est alors donné par⁶ : $I = 2 \text{ (spin)} \times (-q) \times \text{Trace}(\rho J_{op})$. Dans la représentation "cs" le courant est dû aux éléments non-diagonaux de la matrice densité et de l'opérateur du courant. Ces éléments moyennent les effets d'interférence entre les états "c" et "s", c'est pourquoi ils sont souvent désignés sous le nom de "cohérences". Dans les composants nanométriques, il est important d'avoir une flexibilité dans le choix de la base de représentation puisqu'on ne sait pas a priori laquelle diagonalisera la matrice densité. Le problème est ensuite d'exprimer la matrice densité sous sa forme la plus avantageuse.

⁵Matrice dont la trace vaut 1.

⁶Rappel : $\text{Trace}(\rho J_{op})$ représente la valeur moyenne de l'opérateur J_{op} .

Nous avons mentionné précédemment que la densité électronique pourrait être calculée à partir de la matrice densité en utilisant l'équation (3.56). Cette équation peut s'interpréter comme la transformation unitaire vers la représentation dans l'espace réel :

$$[\rho]_{\text{espace réel}} = [V] [\rho] [V]^\dagger. \quad (3.73)$$

La matrice de transformation $[V]$ est obtenue à partir des amplitudes des fonctions d'ondes Ψ_α aux points "r" dans l'espace réel :

$$[V]_{\vec{r},\alpha} = \Psi_\alpha(\mathbf{r}) \sqrt{\Omega}, \quad (3.74)$$

où Ω est le volume unitaire entourant chaque point⁷. Substituant (3.74) dans (3.73) :

$$\rho(\mathbf{r}, \mathbf{r}') = \Omega \sum_{\alpha,\beta} \Psi_\alpha(\mathbf{r}) \Psi_\beta(\mathbf{r}') \rho_{\alpha\beta} \rightarrow \Omega n(\mathbf{r}) = [\rho(\mathbf{r}, \mathbf{r}')]_{\mathbf{r}'=\mathbf{r}}. \quad (3.75)$$

En conséquence la densité électronique $n(\mathbf{r})$ est simplement définie par les éléments diagonaux de la matrice densité représentée dans l'espace réel (à une constante près).

B) Résumé

Dans la théorie du transport quantique, la matrice densité est la quantité centrale à partir de laquelle toutes les grandeurs physiques sont obtenues. La difficulté est de trouver la matrice densité dans une représentation choisie. Pour cela, il ne suffit pas de connaître le composant à travers l'hamiltonien H et le potentiel auto-cohérent U (comme dans tout SCHRÖDINGER-POISSON), mais nous devons également savoir comment le composant est couplé aux deux contacts. Cette information est contenue dans les fonctions de self-énergie qui, associées au formalisme des fonctions de GREEN hors-équilibre, conduisent à des relations claires.

3.3.3 Formalisme de la diffusion

La théorie de LANDAUER [122, 176, 177] est basée sur l'approximation que les électrons n'interagissent pas entre eux dans la zone active conductrice. Dans cette limite, le formalisme des fonctions de GREEN hors-équilibre conduit à des expressions du courant simples qui peuvent être établies en utilisant une approche assez "transparente" inspirée du formalisme de la diffusion élastique. Deux formules équivalentes seront démontrées car la plus utilisée dans le milieu du transport moléculaire ne nous permettra pas de traiter l'influence de défauts ponctuels dans les transistors MOSFET à trois dimensions (cf chapitre 6).

Considérons le système divisé en deux parties de la figure (3.9), représentant un nano-dispositif dans lequel le transport est quasi-cohérent, les processus de diffusion inélastique dans la zone

⁷Le facteur Ω est dû au procédé de discrétisation des coordonnées de l'espace réel. Ce procédé est conceptuellement confortable puisqu'il donne à la matrice de transformation une dimension finie.

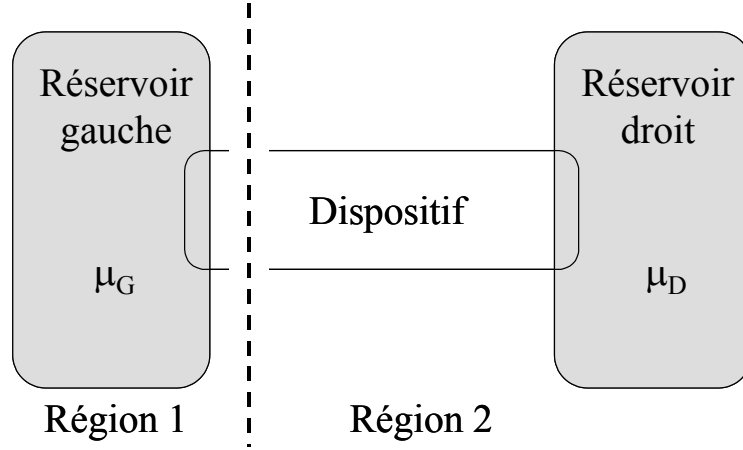


FIG. 3.9: Nano-structure virtuellement clivée en deux régions.

active pouvant être négligés. Le formalisme de la diffusion élastique permet de calculer le courant traversant la structure en utilisant les états propres de l'hamiltonien total $H = H_0 + V$ où H_0 est l'hamiltonien des deux régions non couplées et où V représente leur couplage :

$$H = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix}, \quad H_0 = \begin{bmatrix} H_{11}^0 & 0 \\ 0 & H_{22}^0 \end{bmatrix} \quad \text{et} \quad V = \begin{bmatrix} 0 & V_{12} \\ V_{21} & 0 \end{bmatrix} \quad (3.76)$$

Parmi les états propres de H , nous distinguons deux catégories :

- ceux d'énergie ε_i ($|i_+\rangle$) provenant de la région 1, dont une composante est partiellement réfléchiée et la seconde est partiellement transmise à la région 2.
- les états symétriques ($|j_-\rangle$) d'énergie ε_j , provenant de la région 2 (voir figure (3.10)). Les états $|i_+\rangle$ sont à l'équilibre thermodynamique avec le réservoir gauche, c'est-à-dire qu'ils sont distribués en énergie selon la fonction de FERMI-DIRAC $f(\varepsilon_i - \mu_G)$. Symétriquement, les états $|j_-\rangle$ sont occupés selon la fonction de distribution de FERMI-DIRAC $f(\varepsilon_j - \mu_D)$.

Les états $|i_+\rangle$ sont solution de l'équation de SCHRÖDINGER :

$$\lim_{\eta \rightarrow 0^+} (\varepsilon_i - H_0 + i\eta) |i_+\rangle = V |i_+\rangle. \quad (3.77)$$

Si $V = 0$, les solutions sont alors les états propres de H_0 dans la région de gauche :

$$\lim_{\eta \rightarrow 0^+} (\varepsilon_i - H_0 + i\eta) |i\rangle = 0. \quad (3.78)$$

Les solutions formelles de l'équation (3.77) s'écrivent donc ($\eta \rightarrow 0^+$) :

$$|i_+\rangle = |i\rangle + (\varepsilon_i - H_0 + i\eta)^{-1} V |i_+\rangle. \quad (3.79)$$

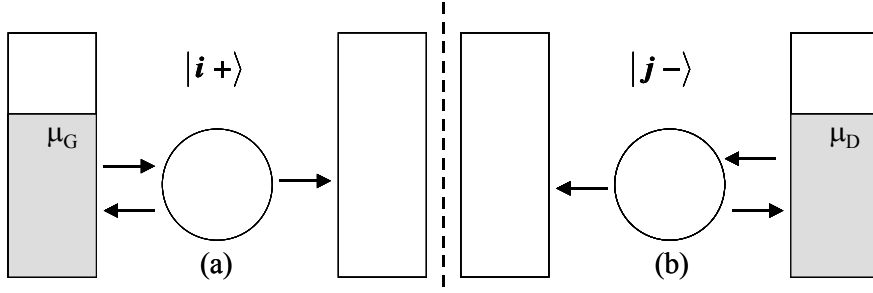


FIG. 3.10: Dans un système hors-équilibre, les états propres de l'hamiltonien total sont divisés en deux catégories : (a) ceux injectés depuis le contact gauche et occupés jusqu'au potentiel électrochimique μ_G et (b) ceux injectés depuis le contact droit et occupés jusqu'au potentiel électrochimique μ_D .

En remarquant que les équations (3.77) et (3.78) peuvent s'écrire respectivement :

$$\lim_{\eta \rightarrow 0^+} (\varepsilon_i - H_0 - V + i\eta) |i_+\rangle = 0, \quad (3.80)$$

$$\lim_{\eta \rightarrow 0^+} (\varepsilon_i - H_0 - V + i\eta) |i\rangle = -V |i\rangle, \quad (3.81)$$

les états $|i\rangle$ s'expriment sous la forme ($\eta \rightarrow 0^+$) :

$$|i\rangle = |i_+\rangle - (\varepsilon - H_0 - V) |i\rangle. \quad (3.82)$$

Pour déterminer la structure électronique du système perturbé, il est utile d'introduire les opérateurs G et G_0 . G (G_0) correspond à la fonction de GREEN retardée [190] du système (non) couplé :

$$G_0(\varepsilon) = \lim_{\eta \rightarrow 0^+} (\varepsilon - H_0 + i\eta)^{-1},$$

$$G(\varepsilon) = \lim_{\eta \rightarrow 0^+} (\varepsilon - H + i\eta)^{-1}. \quad (3.83)$$

La partie imaginaire est introduite de façon à éviter les problèmes de division par zéro quand ε est égal à une valeur propre de H (ou H_0). Une forme plus explicite de G (similairement G_0) est :

$$G(\varepsilon) = \lim_{\eta \rightarrow 0^+} \sum_k \frac{|k\rangle \langle k|}{\varepsilon - \varepsilon_k + i\eta}, \quad (3.84)$$

où les vecteurs $|k\rangle$ sont les vecteurs propres de H . La forme matricielle équivalente est :

$$G = \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix} \text{ et } G_0 = \begin{bmatrix} (G_0)_{11} & 0 \\ 0 & (G_0)_{22} \end{bmatrix}. \quad (3.85)$$

L'injection de (3.83) dans (3.79) conduit à l'équation de LIPPMANN-SCHWINGER :

$$|i_+\rangle = |i\rangle + G_0(\varepsilon_i) V |i_+\rangle, \quad (3.86)$$

qui peut également s'écrire à l'aide de l'équation (3.82) :

$$|i\rangle = |i_+\rangle - G(\varepsilon_i) V |i\rangle, \quad (3.87)$$

ou

$$|i_+\rangle = |i\rangle + G(\varepsilon_i) V |i\rangle. \quad (3.88)$$

Les fonctions de GREEN G et G_0 sont reliées par l'équation de DYSON :

$$G = G_0 + G_0 V G, \quad (3.89)$$

qui est directement déduite des équations (3.83). En effet :

$$\begin{aligned} G^{-1}(\varepsilon) &= \lim_{\eta \rightarrow 0^+} (\varepsilon - H_0 - V + i\eta) = G_0^{-1}(\varepsilon) - V \\ &\Rightarrow G(\varepsilon) = G_0(\varepsilon) + G_0(\varepsilon) V G(\varepsilon). \end{aligned} \quad (3.90)$$

Considérons maintenant le courant total I . Il peut être calculé en utilisant une généralisation du théorème d'ERHNFEST [191]. Dans la suite, nous allons utiliser une approche équivalente basée sur le formalisme de la matrice densité. Partons de la définition macroscopique du courant :

$$I = \left(\frac{dQ_D}{dt} \right)_0, \quad (3.91)$$

où Q_D est la charge totale de la région 2, donnée par :

$$Q_D = \sum_{j \in 2} (-e) \langle j | \rho | j \rangle, \quad (3.92)$$

où ρ est l'opérateur densité hors-équilibre qui sera défini par la suite. Les kets $|j\rangle$ sont les états propres de H_0 dans la région 2. L'indice 0 de l'équation (3.91) signifie que seules les charges circulant à travers l'interface séparant les régions 1 et 2 doivent être considérées. L'évolution temporelle de l'opérateur densité est régie par l'équation de LIOUVILLE:

$$\frac{\partial \rho}{\partial t} = \frac{1}{i\hbar} [H_0, \rho] + \frac{1}{i\hbar} [V, \rho], \quad (3.93)$$

Seul le second terme de droite de l'équation (3.93) (résultant du couplage V), doit être retenu, puisque pour $V = 0$, l'intensité du courant circulant dans le nano-composant est nulle⁸.

Ecrivons maintenant le courant total sous la forme $I = I^+ + I^-$, en séparant les contributions de chaque catégorie d'états. Le courant I^+ résulte des électrons se propageant dans les états $|i_+\rangle$ injectés depuis le côté gauche et diffusés par le potentiel V .

⁸Approche valable si ρ_{11} (partie de l'opérateur densité décrivant la région 1) et ρ_{22} (partie de l'opérateur densité décrivant la région 2) sont indépendantes de la tension appliquée V . Généralement vrai dans le cas de réservoirs à l'équilibre thermo-dynamique connectés à des régions balistiques.

Des équations (3.91), (3.92), (3.93), résulte l'expression du courant partiel I^+ :

$$I^+ = \frac{(-e)}{i\hbar} \sum_{j \in 2} \langle j | [V, \rho_+] | j \rangle, \quad (3.94)$$

où ρ_+ est la contribution des états $|i_+\rangle$ dans ρ^9 :

$$\rho_+ = \sum_{i \in 1} |i_+\rangle f(\varepsilon_i - \mu_G) \langle i_+|. \quad (3.95)$$

Injectant (3.95) dans (3.94), nous obtenons :

$$I^+ = \frac{(-e)}{i\hbar} \sum_{j \in 2, i \in 1} \{ \langle j | V | i_+\rangle \langle i_+ | j \rangle - c.c. \} f(\varepsilon_i - \mu_G), \quad (3.96)$$

où *c.c.* dénote le complexe conjugué du terme qui le précède.

De la même manière, nous pouvons exprimer le courant I^- qui correspond à un électron provenant du réservoir droit diffusé par la zone active et absorbé par le réservoir gauche.

$$I^- = -\frac{(-e)}{i\hbar} \sum_{j \in 2, i \in 1} \{ \langle i | V | j_-\rangle \langle j_- | i \rangle - c.c. \} f(\varepsilon_j - \mu_D). \quad (3.97)$$

(Le signe $-$ est dû à la direction opposée du transfert de charges). Dans le cas particulier où le système est à l'équilibre ($\mu_G = \mu_D$), le courant total I doit être nul. Ceci implique que les contributions de I^+ et I^- à énergie égales ($\varepsilon_i = \varepsilon_j$) s'annulent, i.e. :

$$\sum_{i \in 1} \{ \langle i | V | j_-\rangle \langle j_- | i \rangle - c.c. \} = \sum_{j \in 2} \{ \langle j | V | i_+\rangle \langle i_+ | j \rangle - c.c. \}. \quad (3.98)$$

On en déduit l'expression du courant total I :

$$I = \frac{(-e)}{i\hbar} \sum_{j \in 2, i \in 1} \{ \langle j | V | i_+\rangle \langle i_+ | j \rangle - c.c. \} [f(\varepsilon_i - \mu_G) - f(\varepsilon_i - \mu_D)] \quad (3.99)$$

Exprimant les fonctions propres du système couplé ($|i_+\rangle$) comme la somme des projections sur les deux sous-systèmes 1 et 2, nous pouvons écrire :

$$|i_+\rangle = |i_{1+}\rangle + |i_{2+}\rangle,$$

La nouvelle expression du courant est alors :

$$I = \frac{(-e)}{i\hbar} \sum_{j \in 2, i \in 1} \{ \langle j | V_{21} | i_{1+}\rangle \langle i_{2+} | j \rangle - c.c. \} [f(\varepsilon_i - \mu_G) - f(\varepsilon_i - \mu_D)], \quad (3.100)$$

⁹Consulter [89] pour une étude approfondie de l'opérateur densité.

qui peut également se réécrire :

$$I = \frac{(-e)}{i\hbar} \sum_{i \in 1} \{ \langle i_{2+} | V_{21} | i_{1+} \rangle - c.c. \} [f(\varepsilon_i - \mu_G) - f(\varepsilon_i - \mu_D)]. \quad (3.101)$$

Il reste à calculer les différents termes intervenants dans les équations du courant à l'aide du formalisme de diffusion. Deux expressions seront démontrées suivant la relation de LIPPMANN-SCHWINGER choisie.

A) Première expression du courant

L'équation de LIPPMANN-SCHWINGER (3.86) permet d'exprimer les vecteurs propres du système couplé en fonction des vecteurs propres du système découplé :

$$\begin{aligned} |i_{1+}\rangle &= |i\rangle + (G_0 V)_{12} |i_{2+}\rangle, \\ |i_{2+}\rangle &= (G_0 V)_{21} |i_{1+}\rangle. \end{aligned} \quad (3.102)$$

Ces relations peuvent se réécrire sous la forme :

$$\begin{aligned} |i_{1+}\rangle &= \Lambda_{11} |i\rangle, \\ |i_{2+}\rangle &= (G_0 V)_{21} \Lambda_{11} |i\rangle, \end{aligned} \quad (3.103)$$

où $\Lambda = (I - G_0 V G_0 V)^{-1}$. En injectant l'équation (3.103) dans l'équation du courant (3.101), nous obtenons :

$$I = \frac{(-e)}{i\hbar} \sum_{i \in 1} \langle i | \Lambda^\dagger V G_0 V \Lambda - \Lambda^\dagger V G_0^* V \Lambda | i \rangle \{ f(\varepsilon_i - \mu_G) - f(\varepsilon_i - \mu_D) \}, \quad (3.104)$$

où Λ^\dagger est l'adjoint de l'opérateur Λ .

De la définition (3.84), il résulte¹⁰ :

$$G_0 - G_0^* = -2\pi i n(\varepsilon), \quad (3.105)$$

où $n(\varepsilon)$ désigne l'opérateur $\delta(\varepsilon - H_0)$. Ceci permet d'écrire :

$$I = -\frac{2\pi e}{\hbar} \sum_{i \in 1} \langle i | \Lambda^\dagger V n V \Lambda | i \rangle \{ f(\varepsilon_i - \mu_G) - f(\varepsilon_i - \mu_D) \}. \quad (3.106)$$

Le couplage des deux réservoirs élargit les niveaux de la nano-structure et crée un continuum

¹⁰Voir appendice sur les fonctions de GREEN.

d'énergie. La somme discrète de l'équation (3.106) se transforme ainsi en intégrale sur l'énergie.

$$I = -\frac{2\pi e}{\hbar} \int \{f(\varepsilon - \mu_G) - f(\varepsilon - \mu_D)\} \sum_{i \in 1} \delta(\varepsilon - \varepsilon_i) \langle i | \Lambda^+ V n V \Lambda | i \rangle d\varepsilon. \quad (3.107)$$

Utilisant le fait que :

$$\delta(\varepsilon - \varepsilon_i) \langle i | = \langle i | \delta(\varepsilon - H_0) = \langle i | n, \quad (3.108)$$

nous pouvons réécrire (3.107) sous la forme :

$$I = -\frac{2\pi e}{\hbar} \int \{f(\varepsilon - \mu_G) - f(\varepsilon - \mu_D)\} \sum_{i \in 1} \langle i | n \Lambda^\dagger V n V \Lambda | i \rangle d\varepsilon. \quad (3.109)$$

Enfin, les fonctions $|i\rangle$ forment une base du sous-espace de la région 1, et permettent de simplifier l'expression matricielle du courant en :

$$I = -\frac{2\pi e}{\hbar} \int \{f(\varepsilon - \mu_G) - f(\varepsilon - \mu_D)\} Tr_1(n_{11} \Lambda_{11}^\dagger V_{12} n_{22} V_{21} \Lambda_{11}) d\varepsilon, \quad (3.110)$$

où Tr_1 représente la trace sur le sous-espace de la région 1. Cette expression peut encore être synthétisée :

$$I = -\frac{e}{h} \int T(\varepsilon) \{f(\varepsilon - \mu_G) - f(\varepsilon - \mu_D)\} d\varepsilon, \quad (3.111)$$

et constitue la formule généralisée de l'équation de LANDAUER qui relie le courant au coefficient de transmission $T(\varepsilon) = 4\pi^2 Tr(n_{11} \Lambda_{11}^\dagger V_{12} n_{22} V_{21} \Lambda_{11})$ [177] à travers la région active.

Cette expression, peu courante, nous permettra néanmoins de modéliser les défauts ponctuels dans les MOSFETs 3D (cf chapitre 6). Dans la partie suivante, nous allons obtenir une autre équation du courant en tout point équivalente à celle ci-dessous¹¹, en utilisant une expression différente de l'équation de LIPPMANN-SCHWINGER.

B) Equation équivalente du courant

Au lieu de diviser le système en deux parties, nous pouvons nous inspirer de l'architecture du nano-composant en choisissant un découpage en 3 parties : les deux réservoirs et la zone active (figure (3.11)). Dans cette configuration, les électrons participants à la conduction ne proviennent plus de la région 1 ou 2 comme précédemment, mais des réservoirs (région G ou D). La matrice de l'hamiltonien du système couplé H et la fonction de GREEN associée ont désormais la forme suivante :

$$H = \begin{bmatrix} H_{GG} & H_{GM} & 0 \\ H_{MG} & H_{MM} & H_{MD} \\ 0 & H_{DM} & H_{DD} \end{bmatrix}, \quad G = \begin{bmatrix} G_{GG} & G_{GM} & G_{GD} \\ G_{MG} & G_{MM} & G_{MD} \\ G_{DG} & G_{DM} & G_{DD} \end{bmatrix}. \quad (3.112)$$

¹¹Voir appendice.

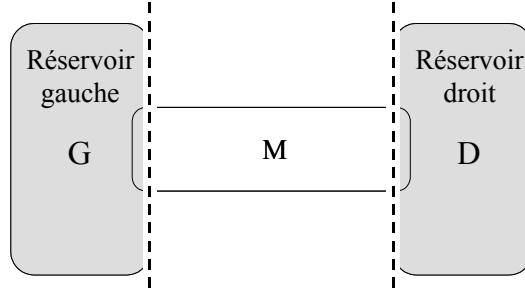


FIG. 3.11: Afin de tenir compte de l'architecture du dispositif, le système est divisé en trois parties : les deux réservoirs et la partie centrale M .

L'équation (3.99) devient :

$$I = \frac{(-e)}{i\hbar} \sum_{j \in D, i \in G} \{ \langle j | V | i_+ \rangle \langle i_+ | j \rangle - c.c. \} [f(\varepsilon_i - \mu_G) - f(\varepsilon_i - \mu_D)], \quad (3.113)$$

où

$$V = \begin{bmatrix} 0 & V_{GM} & 0 \\ V_{MG} & 0 & V_{MD} \\ 0 & V_{DM} & 0 \end{bmatrix}. \quad (3.114)$$

Nous injectons l'équation (3.88) de LIPPMANN-SCHWINGER dans (3.113) et utilisons le fait que $\langle j | i \rangle = 0$, pour obtenir l'équation suivante :

$$I = \frac{(-e)}{i\hbar} \sum_{j \in D, i \in G} \left\{ \langle j | t(\varepsilon_i) | i \rangle \langle i | V G^\dagger(\varepsilon_i) | j \rangle - c.c. \right\} [f(\varepsilon_i - \mu_G) - f(\varepsilon_i - \mu_D)], \quad (3.115)$$

où t est l'opérateur de diffusion donné par :

$$t(\varepsilon_i) = V + V G(\varepsilon_i) V. \quad (3.116)$$

D'après l'équation (3.89), nous pouvons remplacer G^\dagger dans (3.115) par $G_0^\dagger + G^\dagger V G_0^\dagger$. Puisque les vecteurs $|j\rangle$ sont vecteurs propres de H_0 (donc de G_0), nous avons :

$$G_0^\dagger(\varepsilon_i) |j\rangle = \frac{|j\rangle}{\varepsilon_i - \varepsilon_j - i\eta}, \quad (3.117)$$

et déduisons que :

$$\langle i | V G^\dagger(\varepsilon_i) | j \rangle = \frac{\langle i | t^\dagger(\varepsilon_i) | j \rangle}{\varepsilon_i - \varepsilon_j - i\eta}. \quad (3.118)$$

La nouvelle expression du courant est alors :

$$I = \frac{(-e)}{i\hbar} \sum_{j \in D, i \in G} \left\{ \frac{|\langle i | t^\dagger(\varepsilon_i) | j \rangle|^2}{\varepsilon_i - \varepsilon_j - i\eta} - c.c. \right\} [f(\varepsilon_i - \mu_G) - f(\varepsilon_i - \mu_D)], \quad (3.119)$$

qui dans la limite $\eta \rightarrow 0^+$, devient¹² :

$$I = \frac{2\pi(-e)}{\hbar} \sum_{j \in D, i \in G} |\langle i | t^\dagger(\varepsilon_i) | j \rangle|^2 \{f(\varepsilon_i - \mu_G) - f(\varepsilon_i - \mu_D)\} \delta(\varepsilon_i - \varepsilon_j). \quad (3.120)$$

Nous retrouvons ici l'expression (3.28) obtenue en théorie des perturbations à l'exception près que V a été remplacé par l'opérateur de diffusion $t(\varepsilon)$. Plus précisément, V correspond au développement au premier ordre de $t(\varepsilon)$ en puissance de V . Pour vérifier cela, il suffit d'injecter l'équation de DYSON (3.89) dans (3.116) :

$$t = V + VG_0V + VG_0VG_0V + \dots \quad (3.121)$$

Considérons maintenant le cas familier dans lequel les 2 contacts sont uniquement connectés par la nano-structure ($\langle i | V | j \rangle = 0$). Nous avons :

$$t_{ij} = \langle i | t | j \rangle = \sum_{n,m} V_{in} G_{nm} V_{mj}, \quad (3.122)$$

où les n et m désignent les états propres ($|n\rangle$ et $|m\rangle$) de H_0 du nano-composant découplé (la dépendance en énergie de G et t est implicite). Utilisant le fait que $G_{n,m}^* = G_{m,n}^\dagger$, il vient :

$$|t_{ij}|^2 = \sum_{n,m,n',m'} V_{in} G_{nm} V_{mj} V_{n'i} G_{m'n'}^\dagger V_{jm'}. \quad (3.123)$$

Par analogie avec l'équation (3.32), nous pouvons définir les matrices de couplage $\Gamma^G(\varepsilon)$ et $\Gamma^D(\varepsilon)$ par :

$$\begin{aligned} \Gamma_{n,m}^G(\varepsilon) &= 2\pi \sum_{i \in G} V_{ni} V_{im} \delta(\varepsilon - \varepsilon_i), \\ \Gamma_{n,m}^D(\varepsilon) &= 2\pi \sum_{j \in D} V_{nj} V_{jm} \delta(\varepsilon - \varepsilon_j). \end{aligned} \quad (3.124)$$

L'introduction de (3.124) et (3.123) dans (3.120), conduit à une expression compacte du courant :

$$I = -\frac{e}{h} \int \sum_{n,n',m,m'} \Gamma_{n,n'}^G(\varepsilon) \Gamma_{m,m'}^D(\varepsilon) G_{nm}(\varepsilon) G_{m'n'}^\dagger(\varepsilon) \quad (3.125)$$

$$\times \{f(\varepsilon - \mu_G) - f(\varepsilon - \mu_D)\} d\varepsilon, \quad (3.126)$$

$$= -\frac{e}{h} \int Tr \left[\Gamma^G \Gamma^D G^\dagger \right] \{f(\varepsilon - \mu_G) - f(\varepsilon - \mu_D)\} d\varepsilon,$$

où la trace (Tr) de la matrice est calculée en sommant sur tous les états du nano-composant découplé. Cette expression peut se réécrire sous la forme :

$$I = -\frac{e}{h} \int T(\varepsilon) \{f(\varepsilon - \mu_G) - f(\varepsilon - \mu_D)\} d\varepsilon, \quad (3.127)$$

et nous retrouvons la formule de LANDAUER généralisée dans laquelle le coefficient de transmis-

¹²Voir appendice

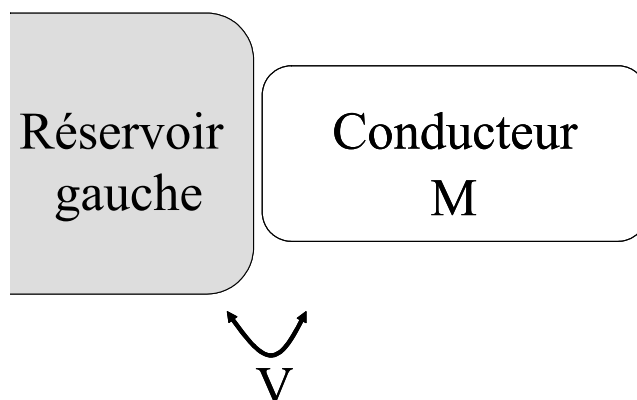


FIG. 3.12: *Conducteur connecté à un réservoir via la matrice de couplage V .*

sion s'écrit désormais : $T(\varepsilon) = Tr [\Gamma^G G \Gamma^D G^\dagger]$.

L'expression (3.125), équivalente à (3.110) (voir appendice), est la plus couramment utilisée en transport moléculaire. Dans la suite de cette étude nous avons choisi d'utiliser l'expression (3.110) afin de traiter l'influence des défauts ponctuels sur les caractéristiques de nanotransistors 3D (cf chapitre 6).

3.3.4 Calcul des fonctions de Green et Self-énergie

Dans la section précédente, nous avons introduit les fonctions de GREEN dans le cadre du formalisme de la diffusion. Il est à ce stade important de s'interroger sur les avantages concrets du formalisme des fonctions de GREEN par rapport à une approche plus "classique" qui résoudrait directement l'équation de SCHRÖDINGER. Nous allons montrer que le formalisme des fonctions de GREEN est particulièrement pertinent pour décrire des systèmes ouverts dont font partie les nano-structures connectées à des contacts semi-infinis. En effet, les fonctions de GREEN permettent de se concentrer sur la partie active du composant et de remplacer l'influence des contacts externes par des self-énergies. Le concept de self-énergie peut également servir à décrire les interactions électron-électron et électron-phonon [157]. Dans le cas présent, nous appliquons cette notion à la description de l'effet sur une nano-structure d'un contact semi-infini.

Considérant un conducteur couplé à un réservoir semi-infini, la matrice de l'hamiltonien à la forme suivante (figure (3.12)) :

$$\begin{bmatrix} [H_0]_{GG} & [V]_{GM} \\ [V]_{MG} & [H_0]_{MM} \end{bmatrix}, \quad (3.128)$$

où les indices M et G réfèrent au nano-composant et au réservoir de gauche respectivement. A partir de la définition de la fonction de GREEN (3.83), nous obtenons ($\eta \rightarrow 0^+$) :

$$G = \begin{bmatrix} G_{GG} & G_{GM} \\ G_{MG} & G_{MM} \end{bmatrix} = \begin{bmatrix} [(\varepsilon + i\eta) I - H_0]_{GG} & -[V]_{GM} \\ -[V]_{MG} & [(\varepsilon + i\eta) I - H_0]_{MM} \end{bmatrix}^{-1}. \quad (3.129)$$

Il peut paraître *a priori* simple d'obtenir la matrice (3.129) et de calculer son inverse. Cependant la dimension de la matrice du réservoir est en toute rigueur infinie! Nous traitons un système ouvert, c'est-à-dire connecté à des électrodes en théorie semi-infinies. Nous pourrions dans une première approximation tronquer la matrice du réservoir au delà d'une certaine distance pour retrouver un système fermé. Néanmoins, cette solution imposerait des conditions aux limites réfléchissantes, qui sont en totale contradiction avec les conditions d'un système ouvert dans lequel une onde incidente est parfaitement absorbée par le réservoir et ne se réfléchit pas. Il nous faut donc procéder plus délicatement dans le couplage du nano-système. L'égalité (3.129) se traduit par le système à deux équations suivant :

$$[(\varepsilon + i\eta)I - H_0]_{GG} G_{GM} - [V]_{GM} G_{MM} = 0, \quad (3.130)$$

$$- [V]_{MG} G_{GM} + [(\varepsilon + i\eta)I - H_0]_{MM} G_{MM} = I. \quad (3.131)$$

De l'équation (3.130) il vient :

$$G_{GM} = [G_0]_{GG} [V]_{GM} G_{MM}, \quad (3.132)$$

où

$$[G_0]_{GG} = [(\varepsilon + i\eta)I - H_0]_{GG}^{-1}, \quad (3.133)$$

est la fonction de GREEN du réservoir semi-infini isolé. Substituant l'équation (3.132) dans (3.131), nous obtenons :

$$G_{MM} = [(\varepsilon + i\eta)I - H_0]_{MM} - [V]_{MG} [G_0]_{GG} [V]_{GM}^{-1} \quad (3.134)$$

Les effets du couplage à un réservoir semi-infini sont ainsi pris en compte à travers le terme $[V]_{MG} [G_0]_{GG} [V]_{GM}$ sans augmenter la taille de la matrice à étudier (G_{MM}). Cette démarche peut se généraliser au cas d'un conducteur couplé à deux électrodes et donne (si le couplage entre électrodes est négligeable) :

$$G_{MM} = \{[(\varepsilon + i\eta)I - H_0]_{MM} - [\Sigma]\}^{-1}, \quad (3.135)$$

où

$$[\Sigma] = [V]_{MG} [G_0]_{GG} [V]_{GM} + [V]_{MD} [G_0]_{DD} [V]_{DM}, \quad (3.136)$$

$$= [\Sigma^G] + [\Sigma^D], \quad (3.137)$$

et

$$[G_0]_{DD} = [(\varepsilon + i\eta)I - H_0]_{DD}^{-1}, \quad (3.138)$$

est la fonction de GREEN du réservoir droit isolé.

Ceci montre que l'effet du couplage à un réservoir peut être simplement décrit en ajoutant une self-énergie Σ à l'hamiltonien H_0 . La self-énergie renormalise l'hamiltonien du nano-composant et

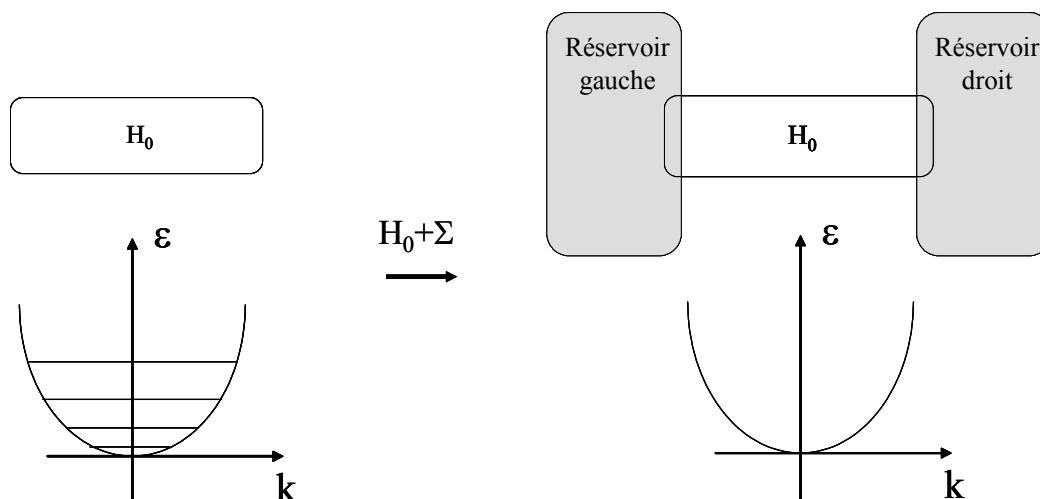


FIG. 3.13: L'interaction du nano-système avec des contacts entraîne une renormalisation de son hamiltonien par une self-énergie Σ et la création d'un continuum d'énergie.

remplace son spectre discret par un continuum (figure (3.13)). Il s'agit d'un concept très général permettant d'éliminer les grands réservoirs et de travailler exclusivement dans le sous-espace de la zone active dont les dimensions sont nettement plus petites. Cependant, il est toujours nécessaire de calculer les grandes (en théorie infinies) matrices $[G_0]_{GG}$ et $[G_0]_{DD}$ des réservoirs isolés. Une première solution est de constater que seuls les éléments de matrice décrivant la surface des réservoirs sont connectés à la nano-structure (le couplage V caractérise les interactions à courtes distances). Nous pouvons alors considérer une partie finie du réservoir, simuler une périodicité et utiliser le théorème de BLOCH afin d'obtenir l'hamiltonien du réservoir. Une seconde approche consiste à traiter un réservoir semi-infini, dont la fonction de GREEN de surface est calculée à l'aide des méthodes de récursion ou de décimation [192] (cf chapitre4).

3.4 Interactions électron-électron et auto-cohérence

Les interactions électron-électron peuvent avoir une influence majeure sur les propriétés de transport des nano-composants. Nous avons d'abord vu dans le paragraphe (3.2.2) qu'il était possible de traiter convenablement les effets du blocage de COULOMB lorsque le couplage avec les contacts était faible ($\Gamma^G, \Gamma^D \ll U$). D'autre part, les approches plus générales comme la théorie de la diffusion considèrent des particules indépendantes.

Dans cette section, nous traitons les théories de champ moyen et montrons comment les appliquer aux problèmes hors-équilibre en soulignant l'importance de l'auto-cohérence lorsque le système est balistique.

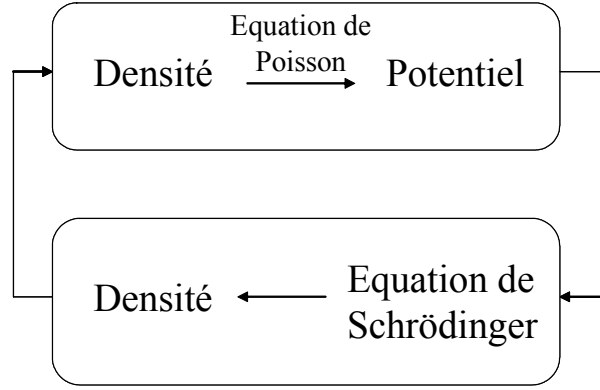


FIG. 3.14: La densité électronique et le potentiel doivent être calculés de manière auto-cohérente afin de tenir compte de l'occupation hors-équilibre des niveaux d'énergies.

3.4.1 Calculs auto-cohérents en théorie du champ moyen

Nous supposons que le système peut être décrit par une équation mono-électronique dans laquelle le potentiel dépend de la densité d'électrons $n(\mathbf{r})$. C'est le cas de méthodes basées sur l'approximation de HARTREE. La résolution auto-cohérente des équations de SCHRÖDINGER et de POISSON, utilisée dans la simulation de dispositifs électroniques [186], se positionne dans cette thématique. Lorsque le nano-composant est connecté à deux contacts dont les potentiels électrochimiques diffèrent, il en résulte une variation de la densité électronique et du potentiel, qui doivent être calculés de façon auto-cohérente (figure (3.14)). La partie auto-cohérente du potentiel peut se décliner en deux contributions : le potentiel de HARTREE et, en théorie de la densité fonctionnelle, le potentiel d'échange-corrélation. Le potentiel de HARTREE peut être déterminé à partir de $n(\mathbf{r})$ avec l'équation de POISSON, exactement comme dans le cas à l'équilibre. En ce qui concerne le potentiel d'échange-corrélation, il est admis que les fonctionnelles d'échange-corrélation usuelles sont capables de décrire les électrons dans des situations hors-équilibre.

Nous devons donc évaluer la densité d'électrons $n(\mathbf{r})$ dans la zone active (partie hors-équilibre) en calculant les quantités $\langle m | \rho | m \rangle$, où ρ est l'opérateur densité et les kets $|m\rangle$ représentent les orbitales atomiques dans le nano-dispositif. Nous avons considéré précédemment, que les états $|i_+\rangle$ et $|j_-\rangle$ sont à l'équilibre avec les réservoirs gauche et droit respectivement. Le système est à nouveau clivé en trois régions : les deux réservoirs et la zone active. Nous avons donc :

$$\rho = \rho_+ + \rho_- \quad (3.139)$$

$$= \sum_{i \in G} |i_+\rangle \langle i_+| f(\varepsilon_i - \mu_G) + \sum_{j \in D} |j_-\rangle \langle j_-| f(\varepsilon_j - \mu_D). \quad (3.140)$$

L'équation de LIPPMANN-SCHWINGER (3.88) donne :

$$\rho_+ = \int f(\varepsilon - \mu_G) (I + GV) \left[\sum_{i \in G} |i\rangle \langle i| \delta(\varepsilon - \varepsilon_i) \right] (I + VG^\dagger) d\varepsilon. \quad (3.141)$$

La fonction delta peut s'exprimer en terme de fonctions de GREEN :

$$\delta(\varepsilon - \varepsilon_i) = \frac{i}{2\pi} \langle i | G_0 - G_0^\dagger | i \rangle, \quad (3.142)$$

et négligeant le recouvrement entre orbitales, nous obtenons la matrice ρ_+ dans la région active :

$$[\rho_+]_{MM} = \frac{i}{2\pi} \int f(\varepsilon - \mu_G) [G]_{MM} [V]_{MG} \left([G_0]_{GG} - [G_0^\dagger]_{GG} \right) [V]_{GM} [G^\dagger]_{MM} d\varepsilon, \quad (3.143)$$

où nous négligeons le couplage direct entre réservoirs¹³. A l'aide des équations (3.124), nous déduisons finalement :

$$[\rho_+]_{MM} = \frac{1}{2\pi} \int f(\varepsilon - \mu_G) [G]_{MM} [\Gamma^G] [G^\dagger]_{MM} d\varepsilon. \quad (3.144)$$

Une équation similaire est obtenue pour ρ_- , et l'opérateur densité global s'écrit :

$$[\rho]_{MM} = \frac{1}{2\pi} \int f(\varepsilon - \mu_G) [G]_{MM} [\Gamma^G] [G^\dagger]_{MM} + f(\varepsilon - \mu_D) [G]_{MM} [\Gamma^D] [G^\dagger]_{MM} d\varepsilon. \quad (3.145)$$

3.4.2 Profil du potentiel auto-cohérent

Le potentiel mono-électronique (HARTREE, échange-corrélation) dans le nano-composant est calculé de façon auto-cohérente en couplant l'équation de POISSON à celle de SCHRÖDINGER. L'évolution du profil de potentiel peut alors être étudiée en fonction de la tension appliquée. Considérons une constriction conductrice connectée à deux réservoirs métalliques (figure (3.15)a). La partie centrale présente un centre diffuseur alors que les contacts sont parfaits ($R_{contacts} = 0$). A l'équilibre les potentiels électrochimiques sont alignés et pour un composant symétrique, l'énergie potentielle électrostatique V_{el} est la même dans les régions gauche et droite de la constriction. $\Delta_0 = \mu_G \text{ ou } D - V_{el}$ est fixée de façon à obtenir la neutralité (figure (3.15)b).

En appliquant une tension V_{DG} , la différence entre les potentiels électrochimiques devient : $\mu_G - \mu_D = eV_{DG}$. Dans l'intervalle d'énergie $[\mu_G, \mu_D]$, les électrons sont injectés dans la constriction depuis le contact gauche, mais à cause de la partie centrale diffusive, une fraction R est réfléchié et T est transmise (figure (3.15)c). Donc pour chaque état se propageant, il y a $1 + R$ électrons dans la partie gauche (au lieu de 2 à l'équilibre) et T dans la partie droite (au lieu de 0 à l'équilibre). Face à cette perturbation, le système réagit via l'auto-cohérence afin d'assurer l'électro-neutralité des régions proches des contacts : l'écart $\mu_G - V_{elG} = \Delta_G$ augmente alors que son symétrique Δ_D diminue. En effet, considérant une densité d'états constante (valable pour de faibles tensions appliquées), la neutralité électrostatique pour la partie gauche s'écrit :

$$2\Delta_0 = 2\Delta_G - TeV_{DG} \quad (\text{si } eV_{DG} \ll \Delta_0) \quad (3.146)$$

i.e.

¹³Les termes en I disparaissent car les fonctions $|i\rangle$ définissent les états du réservoir de gauche et ne pénètrent pas dans la zone active (où la charge est calculée).

$$\Delta_G = \Delta_0 + \frac{T}{2}eV_{DG} > \Delta_0. \quad (3.147)$$

Similairement, pour la région droite :

$$2\Delta_0 = 2\Delta_D + TeV_{DG} \quad (3.148)$$

i.e.

$$\Delta_D = \Delta_0 - \frac{T}{2}eV_{DG} < \Delta_0. \quad (3.149)$$

Finalement :

$$V_{\acute{e}l_G} - V_{\acute{e}l_D} = (\mu_G - \Delta_G) - (\mu_D - \Delta_D) \quad (3.150)$$

$$= (1 - T)eV_{DG}, \quad (3.151)$$

qui tend vers zéro quand $T \rightarrow 1$. Nous retrouvons tout d'abord l'argumentation de LANDAUER-BÜTTIKER [177]. La conductance doit être mesurée à partir des deux réservoirs macroscopiques à l'équilibre thermodynamique ($\mu_G - \mu_D = eV_{DG}$) et non à partir des extrémités de la zone active ($V_{\acute{e}l_G} - V_{\acute{e}l_D}$). Dans le premier cas on trouve la conductance proportionnelle à T alors que dans le second cas elle perd tout sens physique en divergeant pour $T \rightarrow 1$.

La diminution de l'écart $V_{\acute{e}l_G} - V_{\acute{e}l_D}$ quand $T \rightarrow 1$ s'explique en raisonnant sur la densité électronique et les résistances de contacts. Dans un système balistique où les barrières à l'interface entre la région active et les contacts sont négligeables, la résistance de la partie centrale est proche (égale si $T = 1$) de celle des réservoirs métalliques (inverse du quantum de conductance G_0) : une partie de la polarisation est alors appliquée à l'intérieur des réservoirs et non dans la nano-structure. En effet, un flux incident d'électrons provenant du contact gauche, aura un coefficient de réflexion R quasi-nul. La situation est différente à l'intérieur des électrodes, où les états électroniques peuvent être diffusés par des phonons ou des défauts. Les états provenant de la gauche et de la droite sont alors équitablement représentés et le matériau massif se trouve à l'équilibre thermodynamique, caractérisé par un potentiel électrochimique μ_G . Par conséquent, la densité électronique n'est pas la même à l'intérieur de l'électrode et près du contact et l'auto-cohérence engendre une chute de potentiel dans les réservoirs.

3.5 Conclusion

Dans ce chapitre, nous avons montré les limites des modèles classiques ou semi-classiques pour simuler le transport électronique de nano-structures. Parmi les différentes approches quantiques apportées par les calculs de structure électronique, le formalisme des fonctions de GREEN hors-équilibre constitue une méthode très efficace capable d'étudier un large éventail de systèmes : depuis ceux faiblement couplés (SETs) jusqu'à ceux présentant un continuum d'énergie (MOSFETs). Dans la suite de cette étude, nous utiliserons donc ce formalisme pour décrire la

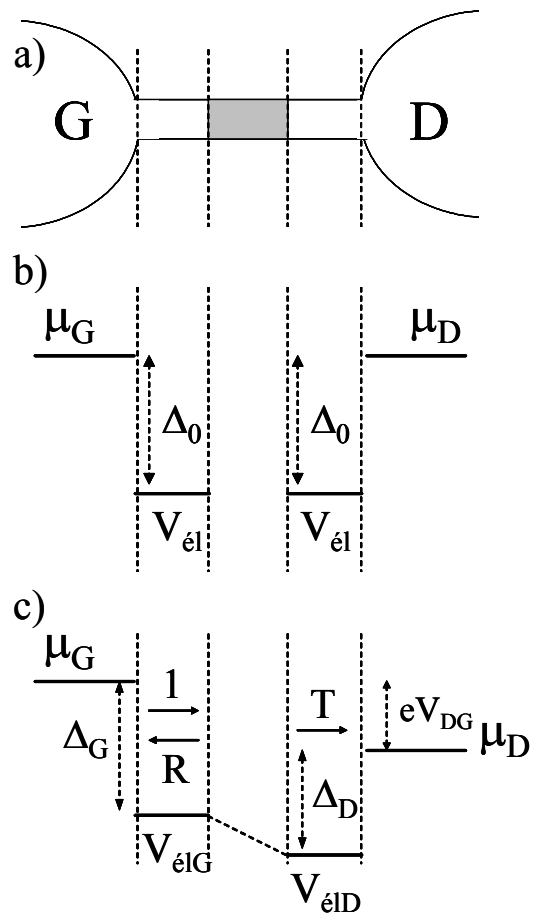


FIG. 3.15: a) Schéma d'une constriction (contenant un centre diffusif représenté en gris) connecté à deux réservoirs métalliques et leur diagramme énergétique associé (b) lorsque le système est à l'équilibre, (c) lorsqu'une tension V_{DG} est appliquée.

conduction de transistors MOSFETs et traiter l'influence de défauts ponctuels.

Chapitre 4

Modélisation atomique de transistors MOSFETs double-grille à canal de conduction confiné

Les progrès de la technologie métal-oxyde-semiconducteur (MOS) conduisent à des transistors de taille nanométrique. A ce niveau de miniaturisation, les effets quantiques ne sont plus négligeables et modifient ostensiblement les propriétés de transport des matériaux. Dans ce contexte, le formalisme des fonctions de GREEN hors-équilibre constitue une méthode pertinente pour décrire le comportement quantique des nano-transistors. Cette approche permet en effet de modéliser les systèmes ouverts (cf § (3.3.4)), ainsi que les défauts ponctuels atomiques (cf chapitre 6).

Parmi les différentes architectures émergentes, le transistor MOSFET double-grille apparaît comme un des meilleurs candidats pour relever le défi de réduction des transistors. L'une des particularités de ce composant est sa très bonne immunité face aux effets canaux courts. La double-grille permet de mieux contrôler le potentiel du canal et de résister au courant tunnel source-drain qui, comme nous le verrons, dégrade la pente sous le seuil des composants. L'objectif de ce chapitre est de simuler un MOSFET double-grille aux dimensions ultimes en utilisant la théorie des liaisons fortes couplée au formalisme des fonctions de GREEN. La section (4.1) compare plusieurs approches susceptibles de modéliser le transport à travers un système mésoscopique. La section (4.2) présente le modèle du transistor double-grille à canal unique de conduction, dans lequel l'axe source-canal-drain est modélisé en liaisons fortes par une chaîne atomique. Nous décrirons alors les résultats clés des simulations, avant de discuter dans la section (4.3) l'influence de la longueur de grille et de son travail de sortie sur le courant de drain. L'ensemble des résultats montreront que la majorité des propriétés électriques d'un tel dispositif sont comparables à celles des transistors MOSFETs conventionnels, confirmant la capacité de miniaturisation des double-grilles. La dernière section conclura.

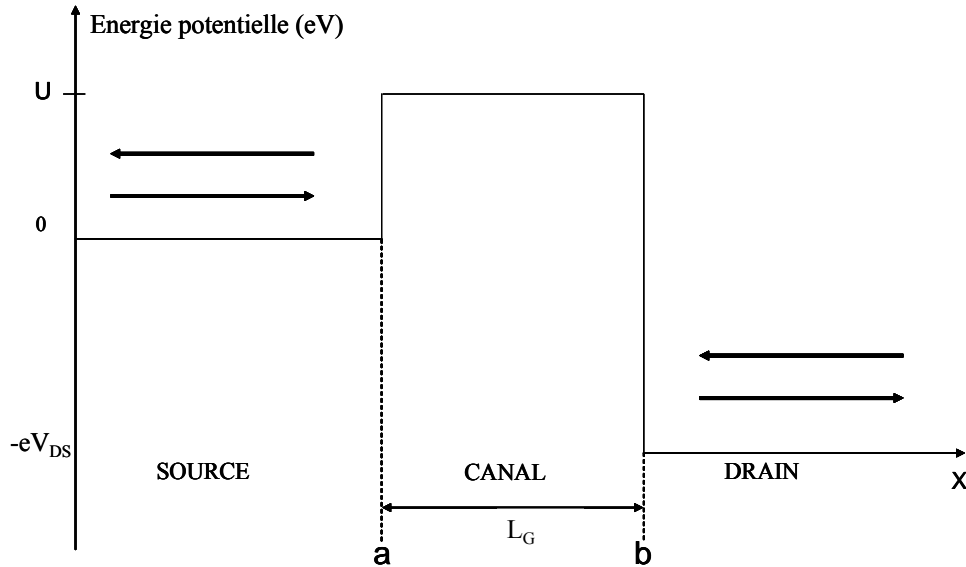


FIG. 4.1: Allure de l'énergie potentielle 1D le long de l'axe source-drain d'un transistor bloqué ($V_G \simeq 0$ V). Les solutions de l'équation de SCHRÖDINGER sont dans l'approximation de la masse effective des combinaisons linéaires d'ondes planes.

4.1 Les différentes approches

La description quantique d'une particule s'effectue à partir de fonctions d'ondes qui sont soit délocalisées dans l'espace telles les ondes planes, soit localisées sous forme d'orbitales atomiques (cf § (2.2.5)). Dans la partie suivante, nous comparons deux approches utilisant chacune une base différente de fonctions d'onde et discutons des limites de leur équivalence dans le cas du calcul du coefficient de transmission à travers un système unidimensionnel.

4.1.1 L'approximation de la masse effective : fonctions d'ondes délocalisées

L'approche quantique la plus fréquemment utilisée en microélectronique est l'approximation de la masse effective. La validité de cette méthode macroscopique repose sur la parabolicité des bas de bandes d'énergie. Considérant un système unidimensionnel, l'approximation de la masse effective représente la propagation d'électrons dans un matériau homogène via des ondes planes. La figure (4.1) décrit schématiquement l'énergie potentielle le long de l'axe source-drain d'un transistor MOSFET dans son état bloqué ($V_G \simeq 0$ V). Dans la source, l'équation de SCHRÖDINGER indépendante du temps s'écrit :

$$-\frac{\hbar^2}{2m_t} \frac{\partial^2}{\partial x^2} \Psi_S(x) = \varepsilon \Psi_S(x), \quad (4.1)$$

où m_t est la masse effective transverse du silicium ($=0.19 \times m_e$)¹. Notant $k_S = \sqrt{\frac{2m_t^* \varepsilon}{\hbar^2}}$, les solutions de l'équation caractéristique de l'équation (4.1) sont :

$$r_{12} = \pm ik_S. \quad (4.2)$$

La solution de l'équation de SCHRÖDINGER est une combinaison linéaire des exponentielles des deux solutions (4.2) :

$$\Psi_S(x) = A_S e^{ik_S x} + B_S e^{-ik_S x}, \quad (4.3)$$

où A_S et B_S sont des constantes complexes. Ψ_S est la superposition de deux ondes : la première (terme en A_S) correspond à une particule incidente, d'impulsion $p_S = \hbar k_S$, se propageant de gauche à droite ; la seconde (terme en B_S) à une particule réfléchie, d'impulsion $-\hbar k_S$, se propageant en sens opposée de la précédente. La fonction d'onde étant définie à une phase près, l'expression suivante reste correcte :

$$\Psi_S(x) = A_S e^{ik_S(x-a)} + B_S e^{-ik_S(x-a)}. \quad (4.4)$$

Une démarche identique dans les deux autres régions conduit aux solutions suivantes (si $0 < \varepsilon < U$) :

$$\begin{aligned} \Psi_C(x) &= A_C e^{K_C(x-a)} + B_C e^{-K_C(x-a)}, \text{ où } K_C = \sqrt{\frac{2m_t(U-\varepsilon)}{\hbar^2}}, \\ \Psi_D(x) &= A_D e^{ik_D(x-b)} + B_D e^{-ik_D(x-b)}, \text{ où } k_D = \sqrt{\frac{2m_t(\varepsilon + eV_{DS})}{\hbar^2}}. \end{aligned} \quad (4.5)$$

La fonction d'onde d'un état quantique ainsi que sa dérivée doivent être continues en tout point. Tenant compte du fait que $B_D = 0$ ², les conditions de raccordement en a et b sont :

$$\Psi_S(a) = \Psi_C(a) \Rightarrow A_S + B_S = A_C + B_C, \quad (4.6)$$

$$\Psi'_S(a) = \Psi'_C(a) \Rightarrow ik_S(A_S - B_S) = K_C(A_C - B_C), \quad (4.7)$$

$$\Psi_C(b) = \Psi_D(b) \Rightarrow A_C e^{K_C(b-a)} + B_C e^{-K_C(b-a)} = A_D, \quad (4.8)$$

$$\Psi'_C(b) = \Psi'_D(b) \Rightarrow K_C(A_C e^{K_C(b-a)} - B_C e^{-K_C(b-a)}) = ik_D A_D. \quad (4.9)$$

La résolution des 4 équations de raccordement ne permet pas de déterminer les 5 inconnues $A_{S,C,D}$ et $B_{S,C}$, mais nous pouvons déterminer les rapports des flux transmis et réfléchi au flux incident :

$$T = \frac{k_D}{k_S} \left| \frac{A_D}{A_S} \right|^2, \quad (4.10)$$

$$R = \left| \frac{B_S}{A_S} \right|^2. \quad (4.11)$$

¹ Si l'on considère qu'un système 1D résulte du confinement ultime d'un fil 3D de silicium suivant les directions transverses, la vallée présentant une masse effective m_t le long de l'axe du transport et des masses m_l (5 fois plus importantes) suivant la section transverse sera énergétiquement la plus basse.

² Une particule incidente provenant de $x = -\infty$ ne subit pas de réflexion dans un réservoir semi-infini, parfait.

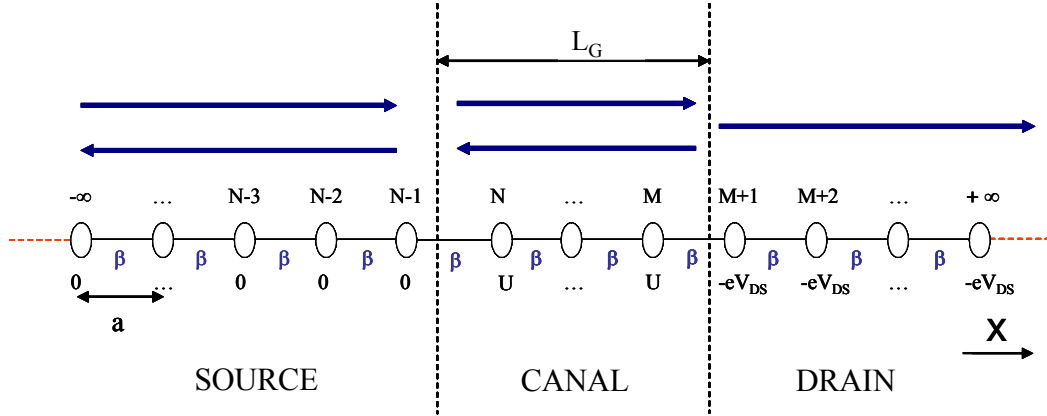


FIG. 4.2: Chaîne linéaire atomique représentant l'axe source-drain d'un transistor MOSFET. Les atomes sont numérotés de $-\infty$ à $+\infty$ et les indices sous chaque atome indiquent l'énergie potentielle intra-atome. a est le paramètre de maille et β l'énergie de couplage.

Ces rapports, respectivement appelés coefficients de transmission³ et de réflexion, donnent les probabilités pour qu'une particule, arrivant de $x = -\infty$, franchisse la marche de potentiel ou retourne vers son contact d'origine. Finalement l'expression analytique du coefficient de transmission est :

$$T = \frac{k_D}{k_S} \times \frac{16 e^{-2K_C(b-a)}}{\left| \left(1 + \frac{K_C}{ik_S}\right) \left(1 + \frac{ik_D}{K_C}\right) e^{-2K_C(b-a)} + \left(1 - \frac{K_C}{ik_S}\right) \left(1 - \frac{ik_D}{K_C}\right) \right|^2}. \quad (4.12)$$

4.1.2 Les liaisons fortes : fonctions d'ondes localisées

La chaîne linéaire est l'équivalent atomique 1D du système précédemment décrit dans l'approximation de la masse effective. Nous utilisons la version la plus simple de la théorie des liaisons fortes, dans laquelle une orbitale est placée sur chaque atome et où seules les interactions entre premiers voisins sont considérées (couplage $\beta < 0$). Numérotant les atomes de $-\infty$ à $+\infty$, la solution de l'équation de SCHRÖDINGER est une combinaison linéaire des orbitales atomiques et s'écrit :

$$\Psi(x) = \sum_n a_n \phi_n, \quad (4.13)$$

où ϕ_n est l'orbitale atomique du $n^{\text{ième}}$ atome. Nous appliquons à nouveau sur la chaîne atomique un potentiel créneau comme illustré par la figure (4.2). Dans les régions de source et de canal, les solutions sont des combinaisons linéaires d'ondes incidentes et réfléchies :

- Si $n < N$: $a_n = A_S e^{ik_S(n-N)a} + B_S e^{-ik_S(n-N)a}$,
- Si $N \leq n \leq M$: $a_n = A_C e^{K_C(n-N)a} + B_C e^{-K_C(n-N)a}$.

³Le rapport des 2 vecteurs d'ondes k_D et k_S dans T prend en compte la différence de vitesse de propagation entre la source et le drain.

Les réservoirs étant parfaits, aucun électron n'est réfléchi dans la région du drain où les pré-facteurs de la combinaison linéaire s'écrivent :

$$a_n = A_D e^{ik_D(n-M)a}, \quad n > M. \quad (4.14)$$

La méthode des variations, qui consiste à minimiser l'énergie du système en ajustant les pré-facteurs a_n , conduit à l'équation :

$$\langle \phi_n | H - \varepsilon | \Psi \rangle = 0 \Rightarrow (\langle \phi_n | H | \phi_n \rangle - \varepsilon) a_n = (a_{n-1} + a_{n+1}) \beta, \quad (4.15)$$

d'où sont dérivées trois relations de dispersion :

$$\cdot n < N - 1 : \varepsilon = 2\beta \cos(k_S a), \quad (4.16)$$

$$\cdot N + 1 < n < N + M - 1 : \varepsilon = U + 2\beta \cosh(K_C a), \quad (4.17)$$

$$\cdot M + 1 < n : \varepsilon = -eV_{DS} + 2\beta \cos(k_D a). \quad (4.18)$$

L'énergie d'interaction β entre deux atomes consécutifs est ajustée de manière à obtenir le comportement d'un fil nanométrique de silicium confiné : la courbure du bas de la bande de conduction est choisie égale à celle obtenue avec la masse effective transverse du silicium ($m_t = 0.19 \times m_e$). En effet, d'après l'équation (4.16), l'énergie d'une chaîne atomique à l'équilibre est :

$$\varepsilon = 2\beta \cos(ka), \quad (4.19)$$

qui, pour de faibles valeurs de vecteurs d'ondes s'écrit :

$$\varepsilon = 2\beta \left(1 - \frac{(ka)^2}{2} + o(ka)^2 \right). \quad (4.20)$$

L'approximation de la masse effective considère un bas de bande de conduction parabolique de la forme :

$$\varepsilon = \frac{\hbar^2 k^2}{2m_t}. \quad (4.21)$$

Par identification des termes en k^2 des équations (4.20) et (4.21), nous obtenons :

$$\beta = -\frac{\hbar^2}{2m_t a^2} = -1.25 \text{ eV}, \quad (4.22)$$

où a ($= 4 \text{ \AA}$) est le paramètre de maille suivant la direction \mathbf{x} .

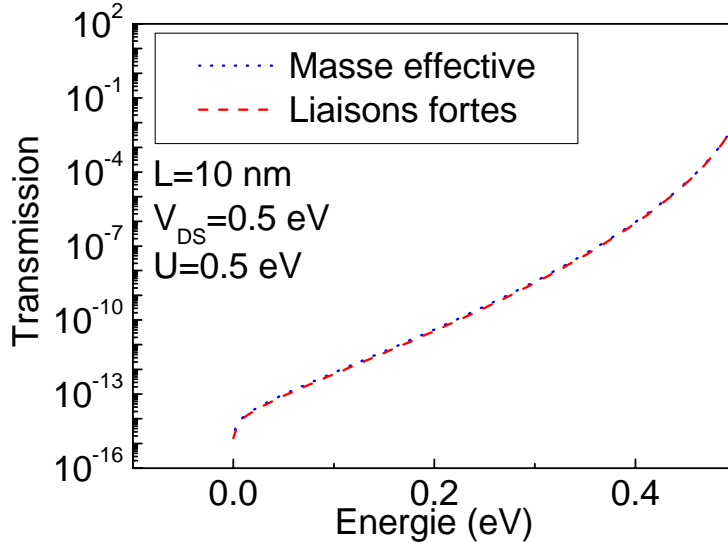


FIG. 4.3: Comparaison des coefficients de transmissions obtenus successivement dans l'approximation de la masse effective et la théorie des liaisons fortes.

Les conditions de raccordement s'expriment maintenant en fonction des pré-facteurs :

$$\begin{aligned}
 a_N^{(S)} &= a_N^{(C)} \Rightarrow A_S + B_S = A_C + B_C \\
 a_{N-1}^{(S)} &= a_{N-1}^{(C)} \Rightarrow A_S e^{-k_S a} + B_S e^{i k_S a} = A_C e^{-K_C a} + B_C e^{K_C a} \\
 a_M^{(C)} &= a_M^{(D)} \Rightarrow A_C e^{K_C (M-N)a} + B_C e^{-K_C (M-N)a} = A_D \\
 a_{M+1}^{(C)} &= a_{M+1}^{(D)} \Rightarrow A_C e^{K_C (M-N+1)a} + B_C e^{-K_C (M-N+1)a} = A_D e^{i k_D a}.
 \end{aligned} \tag{4.23}$$

Après résolution des équations (4.23), l'expression du coefficient de transmission est :

$$\begin{aligned}
 T &= \frac{k_D}{k_S} \times \left| \frac{A_D}{A_S} \right|^2 \\
 &= \frac{k_D}{k_S} \times \left| \frac{2i \sin(k_S a) \cdot \sinh(K_C a) \cdot e^{-K_C (M-N)a}}{(e^{K_C a} - e^{i k_D a}) (e^{i k_S a} - e^{K_C a}) + (e^{i k_D a} - e^{-K_C a}) (e^{i k_S a} - e^{-K_C a}) e^{-2K_C (M-N)a}} \right|^2.
 \end{aligned} \tag{4.24}$$

4.1.3 Comparaison et limites des deux approches

La figure (4.3) représente les coefficients de transmission obtenus avec les deux théories. Nous constatons que les transmissions sont identiques. Néanmoins, les deux théories ne sont pas tout à fait équivalentes. Tout d'abord la relation de dispersion des électrons dans les matériaux n'est parabolique qu'en bas des bandes d'énergie. Lorsque l'énergie des électrons augmente sous l'effet d'un apport extérieur (température, champ électrique), l'approximation de la masse effective n'est rapidement plus vérifiée (figure (4.4)). Dans le cas présent, l'approche en liaisons fortes à une orbitale par atome fournit néanmoins une relation dispersion parabolique jusqu'à environ

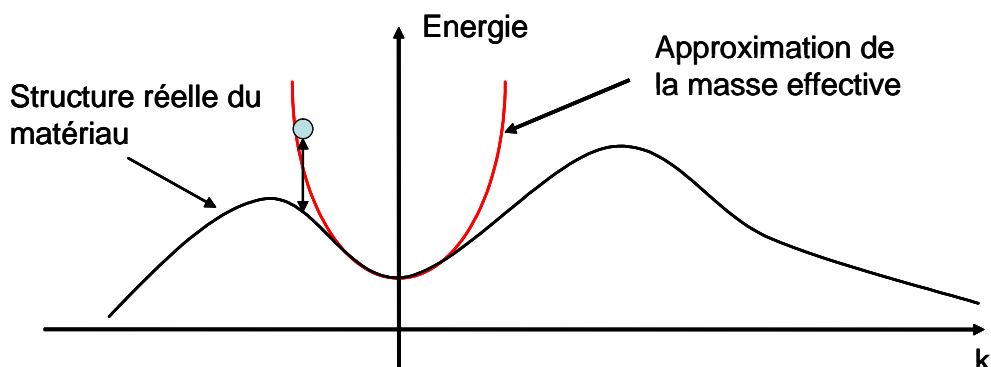


FIG. 4.4: Représentation schématique des limites de l'approximation de la masse effective pour des électrons hautement énergétiques.

1 eV au-dessus du bas de bande de conduction : la différence avec l'approximation de la masse effective restera donc négligeable dans l'intervalle de tensions des MOSFETs (≈ 0.5 V). Il en est différemment avec une description en liaisons fortes plus sophistiquée à plusieurs orbitales par atome, considérant les interactions aux seconds, voir troisièmes voisins [99]. De plus, la méthode des liaisons fortes présente l'avantage de faire apparaître, bien que de façon simpliste dans la version à une orbitale par atome, la structure atomique du matériau. Nous pouvons par conséquent mesurer l'influence de la position des défauts cristallins dans le réseau atomique (dislocations, défauts ponctuels...). Nous aborderons ce domaine aux chapitres 6 et 7.

Dans la suite de ce chapitre, nous insérerons la chaîne linéaire atomique à la place du canal d'un MOSFET double-grille confiné. On adoptera le terme de MOSFET double-grille à canal de conduction unique, en référence à la description de LANDAUER du quantum de conductance [177]. La théorie des liaisons fortes sera exprimée dans le formalisme des fonctions de GREEN.

4.2 Modélisation du transistor double-grille à canal unique de conduction

4.2.1 Présentation du modèle et aspects numériques

Le Modèle

Le transistor MOSFET double-grille est représenté schématiquement sur la figure (4.5). La structure, symétrique, est caractérisée par deux grilles métalliques identiques (de longueur L_G et de travail de sortie ϕ_M supposé fixe = 4.323 V). Les électrodes de grille sont séparées du canal par une couche de diélectrique high- κ à haute permittivité, ($\epsilon_R=19$) dont l'épaisseur équivalente d'oxyde (EOT pour "Equivalent Oxide Thickness") vaut 0.6 nm. Enfin, l'axe source-canal-drain est modélisé par une chaîne atomique linéaire de silicium traitée en liaisons fortes(cf § (4.1.2)).

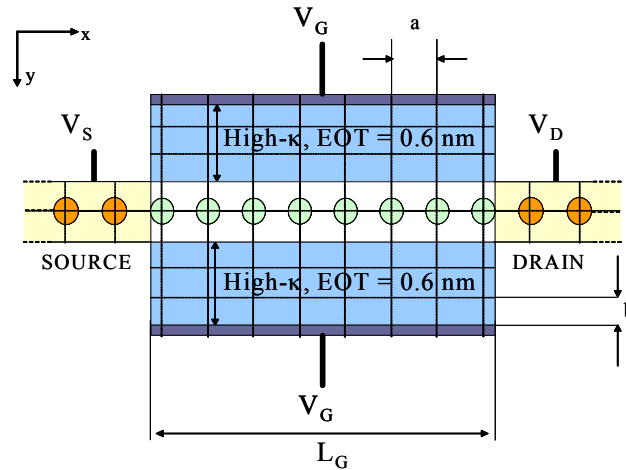


FIG. 4.5: MOSFET double-grille à canal de conduction confiné. L'équation de POISSON 2D est résolue en utilisant la méthode des différences finies : les deux directions \mathbf{x} et \mathbf{y} sont décrites par un maillage uniforme de pas respectivement égal à a et b .

Nous utilisons une version simple de la théorie des liaisons fortes, dans laquelle chaque atome ne possède qu'une orbitale et où seules les interactions aux premiers voisins sont envisagées. Bien qu'équivalente à l'approximation de la masse effective, cette approche simplifiée fait apparaître explicitement la structure atomique de la région active. En conséquence, elle autorise la modélisation de défauts ponctuels tels que les lacunes ou les impuretés cristallines et constitue un modèle efficace pour décrire l'influence de ces défauts sur les caractéristiques électriques des MOSFETs (cf chapitre 6). Le formalisme des fonctions de GREEN hors-équilibre décrit les propriétés électriques et physiques de la chaîne atomique (plus particulièrement le courant et les charges) alors que le reste du système (oxydes et grilles) est traité classiquement d'un point de vue électrostatique.

De par la diminution constante de leur taille, les régions de la source et du drain sont assimilées à des contacts semi-infinis parfaitement conducteurs. La réduction des dimensions des réservoirs induit des matériaux très fortement dopés et soulève plusieurs problèmes technologiques tels que le seuil limite de solubilité des atomes dopants dans le semiconducteur ou la diffusion latérale de ces mêmes impuretés dans le canal au cours des différents recuits [11]. Dans ce contexte, source et drain sont modélisés par des conducteurs parfaits à l'équilibre thermodynamique, dont les niveaux de FERMİ respectifs diffèrent de l'énergie potentielle drain-source ($-eV_{DS}$) appliquée. Ils sont simulés numériquement par une chaîne atomique semi-infinie pour laquelle la fonction de GREEN est obtenue par une méthode de récursion⁴. La caractéristique première d'un réservoir est son électroneutralité. Nous avons montré au § (3.4.2) que l'électroneutralité des contacts s'obtenait via la boucle d'auto-cohérence en décalant les bandes d'énergie des bords de la zone active par rapport aux niveaux de FERMİ. Dans un système unidimensionnel couplé à des réservoirs idéalement conducteurs, la chute de potentiel dans la zone active s'annule lorsque la probabilité pour un électron d'être transmis à travers la zone active augmente

⁴Méthode détaillée dans la suite du paragraphe.

(c'est-à-dire lorsque le transistor est dans l'état passant). La dépendance en énergie de la densité d'états 1D ($\propto 1/\sqrt{\varepsilon}$) est à l'origine de ce comportement. Si une polarisation est appliquée entre deux réservoirs parfaitement conducteurs, seule une variation nulle du potentiel dans la zone active permet d'équilibrer les charges des réservoirs. En revanche, si l'interface réservoir/zone active présente suffisamment de réflexions, l'électroneutralité des contacts est assurée, et la différence de potentiel V_{DS} entre la source et le drain s'applique en totalité sur la zone active. A l'équilibre et sans barrière de potentiel dans le canal, la probabilité de transmettre un électron d'un réservoir à un autre vaut 1, mais le courant total reste nul puisque les charges circulant en sens opposé s'annulent exactement (figure (4.6)). Dans un état hors-équilibre, le champ électrique interdit aux électrons du drain de traverser le canal pour rejoindre la source, alors que le phénomène opposé est très fréquent. Il en résulte un déficit de charges côté source et un excédent côté drain. Dans notre modèle, l'électro-neutralité des contacts est obtenue en ajustant le bas de bande de conduction des réservoirs. Les densités d'états 1D variant en $1/\sqrt{\varepsilon}$, une très faible variation énergétique du bas de bande de conduction suffit à équilibrer les charges aux interfaces. A l'équilibre, les bandes de conduction (BC) des réservoirs sont alignées avec celle de la zone active. Le niveau de FERMI de chaque réservoir est ajusté afin d'obtenir les charges aux interfaces souhaitées (figure (4.7)a). L'équilibrage des charges des réservoirs dans le régime hors-équilibre s'effectue ensuite en décalant la bande de conduction des réservoirs par rapport à celle de la zone active pour chaque tension V_{DS} . Avec cette configuration, l'allure de la densité d'états en $1/\sqrt{\varepsilon}$ facilite désormais l'obtention de l'électroneutralité par un très faible ajustement en énergie (≈ 0.01 eV) de la bande de conduction des réservoirs. L'énergie d'interaction β des réservoirs est supposée identique à celle du canal (figure (4.7)b). Ce choix est conforté par les résultats de GUO et LUNDSTROM [105] qui ont montré qu'une variation de la masse effective du métal de 0.1 à $1 \times m_e$ ne conduisait qu'à une variation des propriétés de transport inférieure à 5%. Afin d'atteindre les densités électroniques rencontrées dans les MOSFETs (10^{26} m^{-3} dans les réservoirs), la bande de conduction du canal est fixée aux interfaces source/canal et canal/drain à 0.11 eV sous chaque niveau de FERMI des contacts. Cette approche, qui sacrifie un traitement rigoureux des contacts, offre néanmoins une méthode efficace capable d'étudier le transport électronique de la zone active. Dans la suite, nous nous proposons de détailler l'application du formalisme des fonctions de GREEN au cas particulier du MOSFET double-grille confiné. Définissons tout d'abord la matrice G de la fonction de GREEN retardée de l'axe source-drain à une énergie ε :

$$G(\varepsilon) = \lim_{\eta \rightarrow 0^+} \frac{1}{(\varepsilon + i\eta)I - H_{\text{inf}}}, \quad (4.25)$$

où $i^2 = -1$ et H_{inf} représente la matrice de l'hamiltonien de la chaîne atomique infinie exprimé

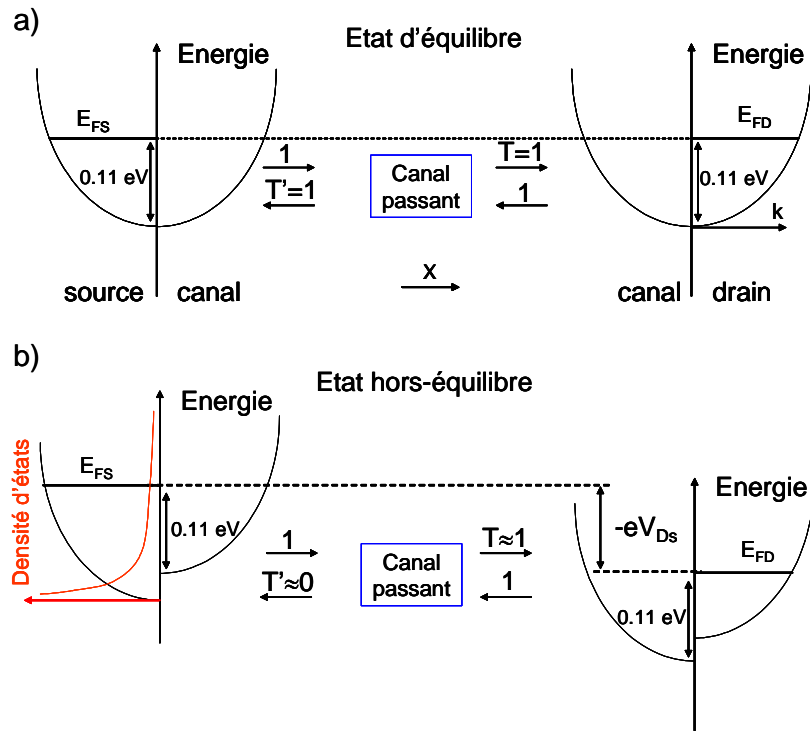


FIG. 4.6: Représentation du transport de charge dans un système balistique 1D a) à l'équilibre, b) dans un régime hors-équilibre. L'électro-neutralité des contacts est obtenue en grâce aux réflexions des interfaces.

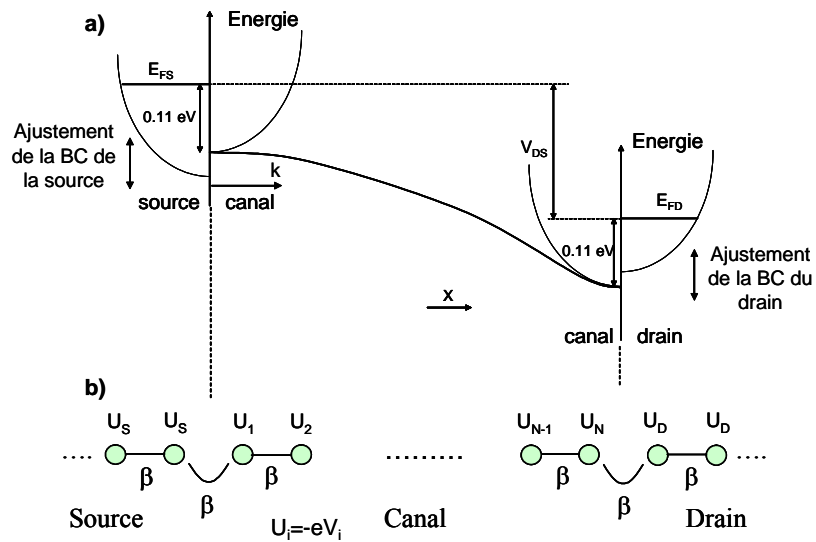


FIG. 4.7: a) Modélisation des bandes de conduction aux interfaces source/canal et canal/drain. b) Equivalence en terme de potentiel électrostatique et énergies de couplage. V_i représente le potentiel électrostatique le long de la chaîne.

en liaisons fortes :

$$H_{\text{inf}} = \begin{bmatrix} \ddots & & & & & \dots & \dots \\ \ddots & -eV_{i-2}(x) & \beta & 0 & 0 & 0 & \dots \\ \ddots & \beta & -eV_{i-1}(x) & \beta & 0 & 0 & \ddots \\ \ddots & 0 & \beta & -eV_i(x) & \beta & 0 & \ddots \\ \ddots & 0 & 0 & \beta & -eV_{i+1}(x) & \beta & \ddots \\ \dots & 0 & 0 & 0 & \beta & -eV_{i+2}(x) & \ddots \\ \dots & \dots & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}, \quad (4.26)$$

où V est le potentiel électrostatique supposé diagonal le long de la chaîne d'atomes numérotés de $-\infty$ à $+\infty$.

Les réservoirs (source et drain) étant des systèmes ouverts, les dimensions de la matrice H_{inf} sont en toute rigueur infinies. Ce problème numérique est contourné en introduisant le concept de self-énergie (cf § (3.3.4)) qui consiste à coupler chaque extrémité du canal à une self-énergie afin d'assurer l'élargissement des niveaux discrets de la partie étudiée et d'éliminer le grand, en fait infini, hamiltonien des contacts.

Nous travaillons désormais dans l'espace associé à la région active dont les dimensions sont numériquement tractables. Les expressions des fonctions de GREEN de surface des réservoirs isolés de la source et du drain, respectivement g_S et g_D , sont calculées suivant la méthode de récursion⁵ :

Considérons une chaîne semi-infinie représentant un réservoir R dont le potentiel électrostatique est constant et les atomes sont numérotés de 1 à $+\infty$ (figure (4.8)). g_R est la fonction de GREEN retardée associée à l'hamiltonien de la chaîne semi-infinie H_R . De façon équivalente, nous définissons H_R^0 et g_R^0 comme les opérateurs correspondant au même système, mais dans lequel les interactions entre les atomes 0 et 1 ont été supprimées. L'expression matricielle basée sur l'équation de DYSON s'écrit :

$$\begin{aligned} g_{r00} &= g_{r00}^0 + g_{r00}^0 K_{01} g_{r10}, \\ g_{r10} &= g_{r11}^0 K_{10} g_{r00}, \end{aligned} \quad (4.27)$$

où K est la matrice de couplage connectant l'atome isolé 0 à la chaîne semi-infinie. Dans une chaîne atomique, tous les termes de K sont nuls sauf K_{10} et K_{01} qui valent chacun β . Le système d'équations (4.27) peut se réécrire en fonction de g_{r00} uniquement :

$$g_{r00} = (1 - g_{r00}^0 K_{01} g_{r11}^0 K_{10})^{-1} g_{r00}^0. \quad (4.28)$$

Puisque l'atome 0 n'est pas connecté à la chaîne semi-infinie dans H_R^0 , sa fonction de GREEN

⁵ g_S et g_D représentent l'élément de matrice de la fonction de GREEN du bord de la chaîne semi-infinie.

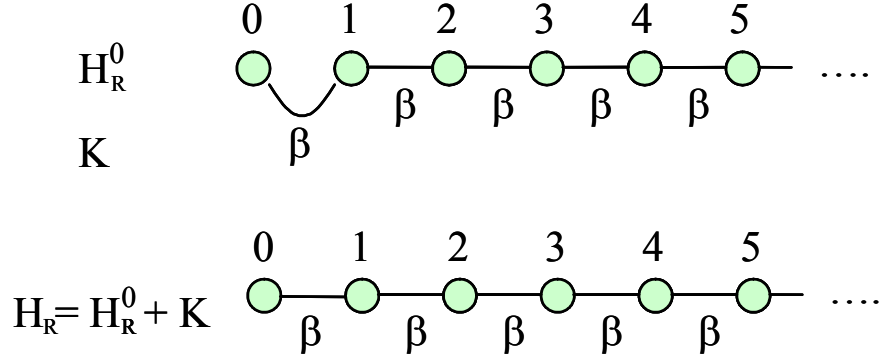


FIG. 4.8: Calcul de la fonction de GREEN de surface d'une chaîne linéaire semi-infinie. L'hamiltonien de départ H_R^0 décrit une chaîne linéaire plus un atome isolé. La perturbation K représente le couplage entre l'atome 0 et la chaîne semi-infinie commençant à l'atome 1.

locale est par définition :

$$g_{r00}^0(\varepsilon) = \lim_{\eta \rightarrow 0^+} (\varepsilon - H_{00}^0 + i\eta)^{-1}. \quad (4.29)$$

De plus, les deux chaînes semi-infinies sont strictement équivalentes que l'atome 0 soit connecté ou non :

$$g_{r11}^0 = g_{r00}. \quad (4.30)$$

Substituant les équations (4.29) et (4.30) dans (4.28), nous obtenons :

$$g_{r00} = \lim_{\eta \rightarrow 0^+} \frac{1}{\varepsilon - H_{r00}^0 - \beta^2 g_{r00} + i\eta}, \quad (4.31)$$

qui constitue une équation quadratique de l'inconnue g_{r00} dont les solutions sont données par :

$$g_{r00}(\varepsilon) = \frac{\varepsilon - H_{r00}^0}{2\beta^2} \left(1 - \sqrt{1 - \frac{4\beta^2}{(\varepsilon - H_{r00}^0)^2}} \right), \quad \text{si } |\varepsilon - H_{r00}^0| > 2|\beta| \quad (4.32)$$

$$g_{r00}(\varepsilon) = \frac{\varepsilon - H_{r00}^0}{2\beta^2} - \frac{i}{|\beta|} \sqrt{1 - \left| \frac{\varepsilon - H_{r00}^0}{2\beta} \right|^2}, \quad \text{si } |\varepsilon - H_{r00}^0| < 2|\beta|. \quad (4.33)$$

Dans le cas de la source et du drain, les fonctions de GREEN retardées de surface s'écrivent donc :

$$g_{S,D}(\varepsilon) = \frac{\varepsilon + eV_{S,D}}{2\beta^2} \left(1 - \sqrt{1 - \frac{4\beta^2}{(\varepsilon + eV_{S,D})^2}} \right) \quad \text{si } |\varepsilon + eV_{S,D}| > 2|\beta|, \quad (4.34)$$

$$g_{S,D}(\varepsilon) = \frac{\varepsilon + eV_{S,D}}{2\beta^2} - \frac{i}{|\beta|} \sqrt{1 - \left| \frac{\varepsilon + eV_{S,D}}{2\beta} \right|^2} \quad \text{si } |\varepsilon + eV_{S,D}| < 2|\beta|, \quad (4.35)$$

où $V_{S(D)}$ est le potentiel électrostatique de la source (du drain). Les self-énergies se présentent

matriciellement sous la forme suivante (cf § (3.3.4)) :

$$\Sigma_S = V_{SC}^\dagger \times g_S \times V_{SC} = \begin{bmatrix} \beta \cdot g_S \cdot \beta & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 \end{bmatrix} \quad (4.36)$$

$$\Sigma_D = V_{CD} \times g_D \times V_{CD}^\dagger = \begin{bmatrix} 0 & \cdots & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & 0 & 0 \\ 0 & \cdots & 0 & \beta \cdot g_D \cdot \beta \end{bmatrix}, \quad (4.37)$$

où Σ_S (resp. Σ_D) est la self-énergie de la source (resp. du drain) et V_{SC} (resp. V_{CD}) représente la matrice de couplage source-canal (resp. canal-drain). Tous les termes de la matrice de couplage sont nuls, sauf ceux associés à la fonction de GREEN de surface des réservoirs. La figure (4.9) illustre l'influence des self-énergies et de leur couplage à la zone active en calculant le coefficient de transmission à travers la chaîne linéaire à l'équilibre. La disparition des pics de transmission avec le renforcement du couplage témoigne de l'élargissement des niveaux discrets du canal isolé vers un continuum d'énergie. La taille de la self-énergie est égale à celle de la région active et la forme finale de la matrice de la fonction de GREEN retardée devient :

$$G(\varepsilon) = \lim_{\eta \rightarrow 0^+} \frac{1}{(\varepsilon + i\eta)I - H - \Sigma_S - \Sigma_D}, \quad (4.38)$$

où H est à présent la matrice de l'hamiltonien du canal à N_X atomes :

$$H = \begin{bmatrix} -eV_1(x) & \beta & 0 & \cdots & \cdots \\ \beta & -eV_2(x) & \beta & \ddots & \cdots \\ 0 & \beta & \ddots & \ddots & 0 \\ \cdots & \ddots & \ddots & -eV_{N-1}(x) & \beta \\ \cdots & \cdots & 0 & \beta & -eV_{N_X}(x) \end{bmatrix}. \quad (4.39)$$

Une fois la fonction de GREEN retardée calculée, la densité d'électrons et le courant de drain sont déterminés. Définissons pour cela deux nouvelles quantités, fonctions des self-énergies :

$$\Gamma_S = i \left(\Sigma_S - \Sigma_S^\dagger \right) \quad \Gamma_D = i \left(\Sigma_D - \Sigma_D^\dagger \right). \quad (4.40)$$

Physiquement, les fonctions Γ fixent la vitesse à laquelle un électron initialement dans un état particulier du canal se propage vers le réservoir de droite ou celui de gauche [173]. Les fonctions spectrales associées à la source et au drain sont alors exprimées par :

$$A_S = G\Gamma_S G^\dagger \quad \text{et} \quad A_D = G\Gamma_D G^\dagger. \quad (4.41)$$

La fonction spectrale de la source étant remplie selon la distribution de FERMI-DIRAC de la

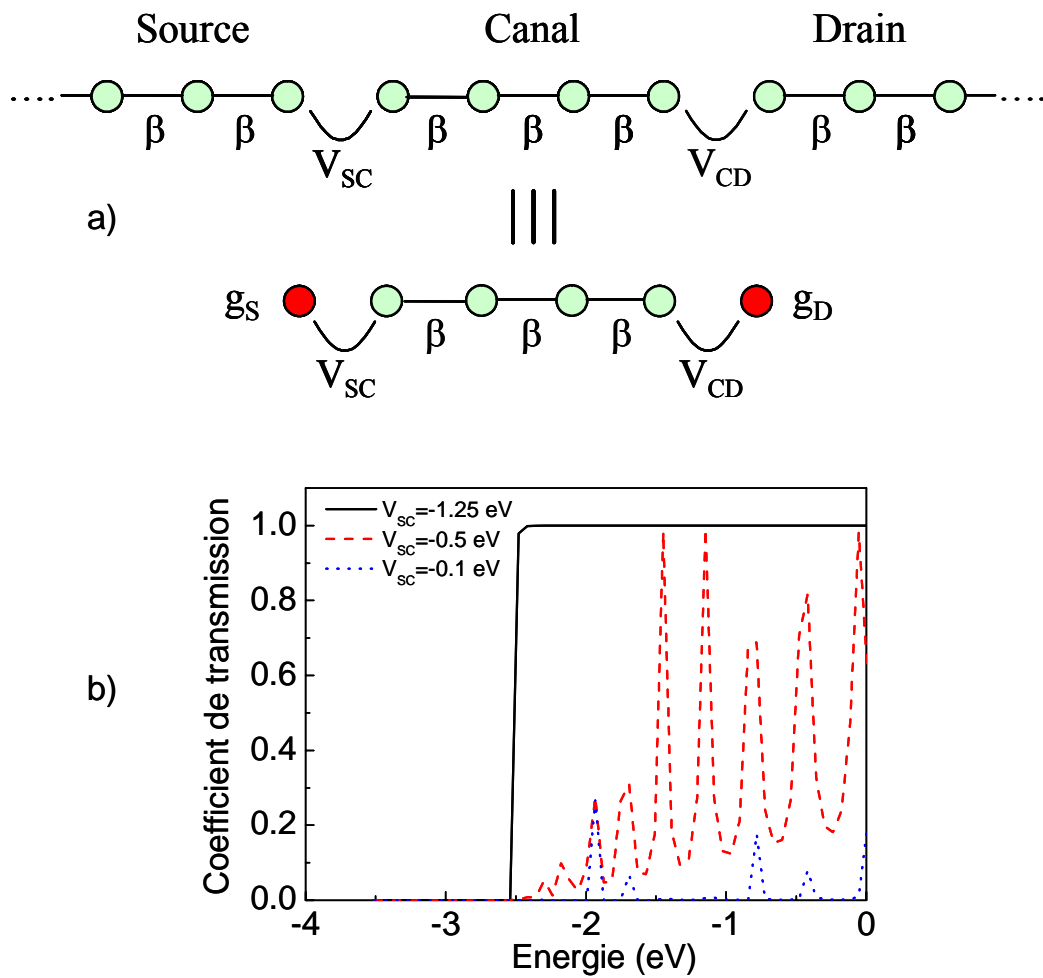


FIG. 4.9: (a) Le concept de la self-énergie permet de modéliser les réservoirs d'électrons en élargissant les niveaux discrets de la zone active isolée. (b) Evolution du coefficient de transmission en fonction du couplage de la zone active aux réservoirs. $\beta = -1.25$ eV, $V_{SC} = V_{CD}$ et $g_{S(D)}$ est la fonction de GREEN retardée de surface de la source (du drain).

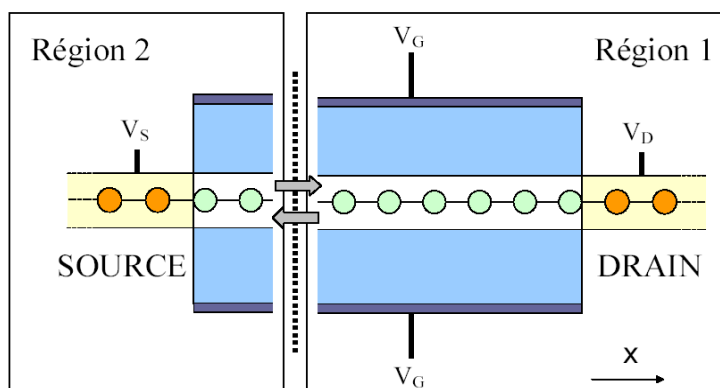


FIG. 4.10: *Calcul du courant : le système est virtuellement clivé en deux régions. Cette approche, dite théorie de la diffusion, exprime la diffusion des états électroniques d'une première région en états de la seconde.*

source $f_S(\varepsilon)$, et la fonction spectrale du drain étant remplie selon la distribution de FERMI-DIRAC du drain $f_D(\varepsilon)$, la matrice de la densité électronique s'écrit (cf § (3.3.2)) :

$$n = \frac{1}{ab^2} \int_{-\infty}^{+\infty} \frac{d\varepsilon}{\pi} [f_S(\varepsilon) A_S + f_D(\varepsilon) A_D], \quad (4.42)$$

où ab^2 constitue le volume de la maille unitaire de la chaîne ($a = b = 0.4 \text{ nm}$) dans lequel les orbitales atomiques se concentrent. L'équation (4.42) rétro-agit alors sur l'équation de POISSON à deux dimensions⁶ afin d'obtenir une solution auto-cohérente. Le courant du canal est ensuite calculé. Pour ce faire, le composant est virtuellement clivé en deux parties avant de déterminer la densité électronique transmise entre les deux régions (figure (4.10)). De par la conservation des charges (transport balistique), le courant total doit être indépendant du plan de clivage le long de l'axe \mathbf{x} . La fonction de GREEN retardée du système non-couplé se présente la sous forme matricielle suivante :

$$G_0(\varepsilon) = \lim_{\eta \rightarrow 0^+} \begin{bmatrix} (\varepsilon + i\eta)I - H_{11} - \Sigma_S & 0 \\ 0 & (\varepsilon + i\eta)I - H_{22} - \Sigma_D \end{bmatrix}, \quad (4.43)$$

où H_{11} (respectivement H_{22}) est la matrice de l'hamiltonien de la région 1 (respectivement de la région 2). Soit C la matrice de couplage assurant la connexion entre les régions 1 et 2 :

$$C = \begin{bmatrix} 0 & H_{12} \\ H_{21} & 0 \end{bmatrix}, \quad (4.44)$$

où H_{12} et H_{21} sont les matrices des blocs non-diagonaux de l'hamiltonien de la chaîne linéaire.

⁶L'équation de POISSON 2D considère nécessairement le dispositif infini suivant la direction \mathbf{z} perpendiculaire au plan de la feuille. En conséquence, les charges de la chaîne linéaire sont équivalentes pour cette équation à celles d'un plan atomique. L'influence sur le potentiel final auto-cohérent reste cependant négligeable, puisque les charges d'une chaîne atomique sont contrôlées par les parties des oxydes de grille se trouvant dans leur plus proche voisinage, c'est-à-dire directement au-dessus et au-dessous.

En se reportant au § (3.3.3), l'expression finale du courant s'écrit⁷ :

$$I = -\frac{4\pi e}{\hbar} \int_{-\infty}^{+\infty} d\varepsilon \operatorname{Tr}_1 \left[N_0 \Lambda^\dagger C N_0 C \Lambda \right] (f_S(\varepsilon) - f_D(\varepsilon)), \quad (4.45)$$

où Tr_1 désigne la trace restreinte à la base des orbitales de la région 1, $N_0 (= -(1/\pi) \operatorname{Im}(G_0))$ est la matrice densité à l'équilibre du système non-couplé, et $\Lambda = (I - G_0 C G_0 C)^{-1}$.

Aspects numériques

Résolution de l'équation de Poisson à deux dimensions (2D)

La résolution numérique de l'équation de POISSON est obtenue en utilisant la loi de GAUSS :

$$\iint \left[\epsilon \vec{E}(x, y) \right] \cdot d\vec{s} = \int_{\Omega} e [p - n + N_D - N_A] d\Omega, \quad (4.46)$$

où \vec{E} est le champ électrique, p est la concentration en trous, n est la concentration en électrons, N_D et N_A sont les concentrations en donneurs et accepteurs, e est la charge élémentaire et ϵ est la constante diélectrique dépendant de la position spatiale. A partir du théorème d'OSTROGRADSKI⁸ et de la relation $\vec{E} = -\overrightarrow{\operatorname{grad}}V$ où V est le potentiel électrostatique, l'équation de POISSON (ou forme locale de la loi de GAUSS) s'écrit :

$$\Delta V = -\frac{e [p - n + N_D - N_A]}{\epsilon} = -\frac{\rho}{\epsilon}, \quad (4.47)$$

avec ρ la densité de charges. La solution d'un tel système peut s'obtenir en maillant le domaine étudié en $N_X \times N_Y$ noeuds, où N_X et N_Y représentent le nombre de noeuds suivant les directions \mathbf{x} et \mathbf{y} respectivement. La solution 2D de l'équation de POISSON est ainsi composée de $N_X \times N_Y$ valeurs de potentiels, initialement inconnus, correspondant à chaque noeud du réseau. Le canal du double-grille, intrinsèque, est connecté à deux réservoirs dopés n^+ . Néanmoins, la résolution de l'équation de POISSON se limite à la région du canal et de ses oxydes, le potentiel étant considéré constant au delà de ces régions. On conserve ainsi à l'équilibre ($V_G = V_{DS} = 0$ V) une barrière de potentiel entre la source et le drain dûe aux jonctions n^+ -int, tout en négligeant la pénétration du potentiel du canal dans les contacts (hypothèse des réservoirs parfaitement conducteurs présentant des réflexions aux interfaces). Afin d'obtenir les équations susceptibles de résoudre le système d'inconnues, nous devons appliquer l'équation (4.47) (dans laquelle les termes N_A , N_D et p sont maintenant nuls) aux noeuds internes et utiliser des conditions particulières aux limites pour les noeuds frontaliers. Etudions tout d'abord le cas d'un noeud interne quelconque $[m, n]$ (ligne m et colonne n) de la figure (4.11). L'approximation des différences finies aux dérivées

⁷L'équation (4.45) est la généralisation de la formule de LANDAUER (cf § (3.3.1)) qui relie le courant au coefficient de transmission $T(\varepsilon) = 4\pi^2 \operatorname{Tr}_1 [N_0 \Lambda^\dagger C N_0 C \Lambda]$.

⁸Le théorème d'OSTROGRADSKI indique que le flux d'un champ vectoriel sortant de la surface fermée et orientable \vec{s} , est égal à l'intégrale étendue au volume Ω intérieur à \vec{s} de la divergence du champ \vec{V} .

spatiales exprime l'équation (4.47) sous la forme :

$$\frac{1}{b^2}V_{m-1,n} + \frac{1}{a^2}V_{m,n-1} - 2\left(\frac{1}{b^2} + \frac{1}{a^2}\right)V_{m,n} + \frac{1}{a^2}V_{m,n+1} + \frac{1}{b^2}V_{m+1,n} = -\frac{1}{\epsilon}e(N_D - N_A - n + p)_{m,n}, \quad (4.48)$$

où a et b sont les pas de réseau dans les directions \mathbf{x} et \mathbf{y} respectivement. Suivant que le noeud $[m, n]$ se situe dans les oxydes ou le silicium, la constante diélectrique ϵ est ϵ_{SiO_2} ou ϵ_{Si} . Dans le cas où le noeud est positionné sur une interface Si/SiO₂, la continuité de la composante perpendiculaire $\epsilon \vec{E}$ s'écrit :

$$\epsilon_{\text{sup}}E_{\text{sup}}^{\perp} = \epsilon_{\text{inf}}E_{\text{inf}}^{\perp}, \quad (4.49)$$

et également sous la forme :

$$\epsilon_{\text{sup}}\left(\frac{\partial V}{\partial y}\right)_{\text{sup}} = \epsilon_{\text{inf}}\left(\frac{\partial V}{\partial y}\right)_{\text{inf}}, \quad (4.50)$$

où ϵ_{sup} et ϵ_{inf} sont les constantes diélectriques du matériau respectivement au-dessus et au-dessous de l'interface. En utilisant les notations de la figure (4.11.b), nous obtenons :

$$(4V_{m-1,n} - 3V_{m,n} - V_{m-2,n})\epsilon_{\text{sup}} = (4V_{m+1,n} - 3V_{m,n} - V_{m+2,n})\epsilon_{\text{inf}}. \quad (4.51)$$

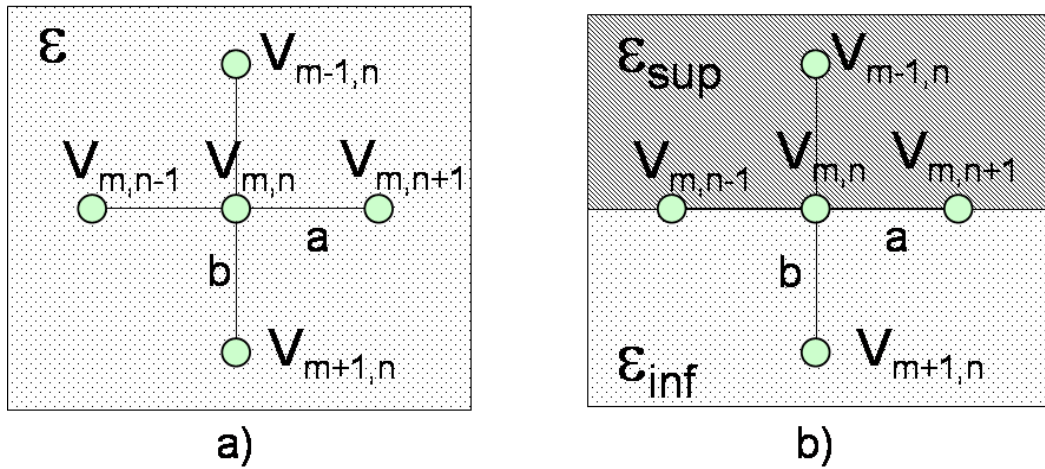


FIG. 4.11: *Discretisation de l'équation de POISSON en différences finies. a) Dans un matériau homogène de constante diélectrique ϵ . b) A l'interface entre deux matériaux aux constantes diélectriques distinctes ϵ_{sup} et ϵ_{inf} .*

Les conditions aux limites de DIRICHLET sont imposées sur les noeuds appartenant au contour de la structure. L'équation à laquelle doit satisfaire le potentiel de grille est donc :

$$V_{m,n} = V_G - V_{FB}, \quad (4.52)$$

où V_{FB} est la tension de bande plate qui traduit la différence des travaux de sortie du métal de grille et du silicium.

Les potentiels des extrémités latérales du canal (qui traduisent le début des réservoirs), respectent les égalités suivantes :

$$V_{m,n} = V_S = 0 \text{ V}, \quad (4.53)$$

$$V_{m,n} = V_D = V_{DS}. \quad (4.54)$$

Connaissant la charge électronique n , les équations (4.47), (4.48), (4.51), (4.52) et (4.53) constituent un système linéaire dont la résolution peut être directement effectuée.

Nous ne considérons pas dans cette solution les conditions aux limites de NEUMANN, qui permettent en général d'ajuster la valeur du potentiel des contacts source et drain afin d'assurer l'électro-neutralité des charges dans ces régions (cf chapitre 6). Les réservoirs sont ici supposés parfaitement conducteurs et l'électro-neutralité est assurée par les réflexions d'interfaces.

L'auto-cohérence

La méthode précédemment décrite est une approche mono-électronique auto-cohérente dans laquelle chaque électron, traité séparément, est soumis à un potentiel électrostatique résultant de son interaction coulombienne avec la densité électronique du système (\equiv potentiel de HARTREE, cf § (2.2)). La boucle d'auto-cohérence consiste à résoudre l'équation de SCHRÖDINGER, exprimée dans le formalisme des fonctions de GREEN, et à la coupler avec l'équation de POISSON. A partir d'un potentiel électrostatique d'essai, l'équation de SCHRÖDINGER fournit une densité de charges (équation (4.42)) qui est injectée dans l'équation de POISSON. Par une double intégration, nous obtenons un nouveau potentiel qui servira d'entrée à l'équation de SCHRÖDINGER. L'opération est ainsi réitérée jusqu'à convergence du potentiel électrostatique et des charges. A l'intérieur de la boucle SCHRÖDINGER-POISSON couramment rencontrée, une seconde boucle assure l'électro-neutralité des contacts (figure (4.12)). La charge des contacts est maintenue constante grâce aux réflexions d'interface, en décalant le bas de la bande de conduction de chaque réservoir. Concrètement, il s'agit d'ajuster les valeurs V_S et V_D des équations (4.34) et (4.35).

Temps de Calcul

La figure (4.13) montre l'évolution du temps de calcul CPU d'un point de courant en fonction de la longueur de grille et pour différents V_G . L'allure générale traduit une augmentation en puissance du temps de calcul avec la longueur de grille. De plus, à L_G constante le temps de calcul augmente avec la tension de grille. Comme nous le verrons dans le § (4.2.2), le transistor commutant de l'état "off" à "on" voit sa quantité de charges transmises croître de façon exponentielle. Le système entre alors dans un état hors-équilibre extrême et l'électro-neutralité des réservoirs est plus difficile à atteindre.

Elargissement des niveaux d'énergie : terme complexe des fonctions de Green

Nous avons montré que des self-énergies connectées aux extrémités de la zone active pouvaient remplacer les réservoirs d'électrons et redonner un continuum d'énergie. L'autre méthode permettant d'élargir les niveaux sans décrire les réservoirs est d'augmenter "artificiellement"

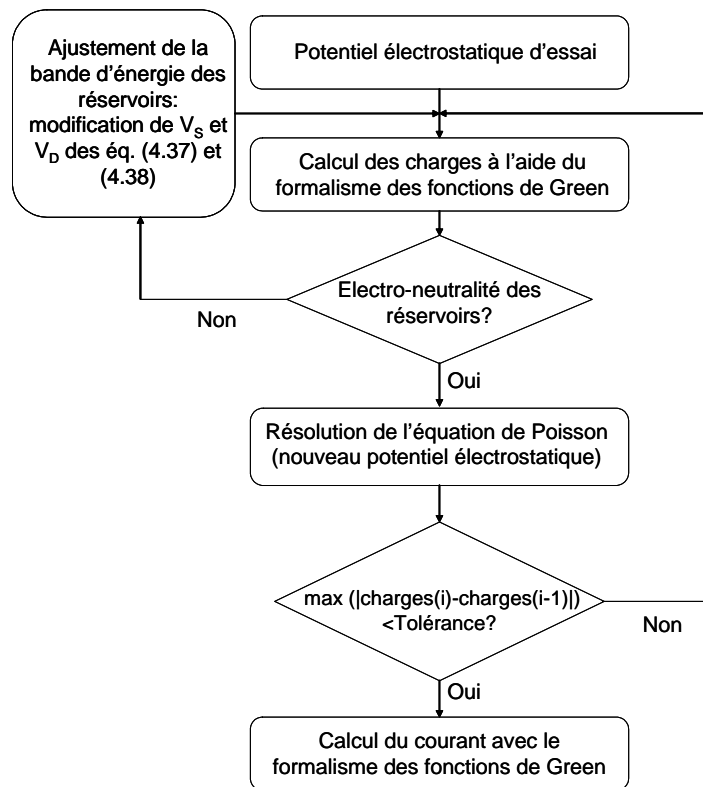


FIG. 4.12: *Etapes numériques de la simulation. L'indice i représente le nombre d'itérations de la boucle SCHRÖDINGER-POISSON.*

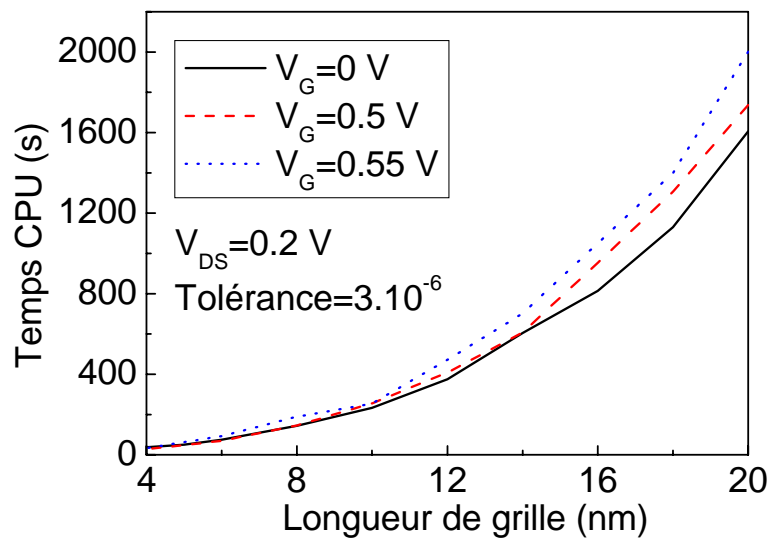


FIG. 4.13: *Temps de calcul CPU d'un point de courant en fonction de la longueur de grille, pour différentes tensions V_G .*

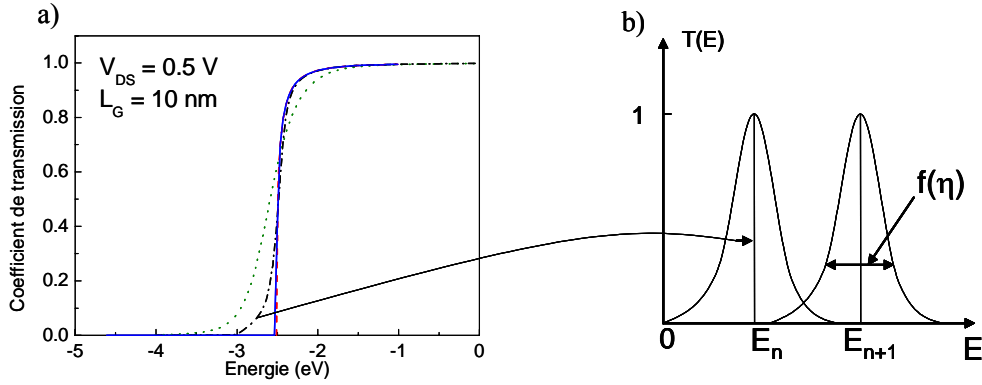


FIG. 4.14: a) Evolution du coefficient de transmission d'une chaîne linéaire en fonction de l'élargissement complexe η . Ligne discontinue : chaîne infinie avec un élargissement nul. Ligne continue : chaîne finie avec un élargissement artificiel $\eta = 0.1$. Ligne tiret-pointillés : chaîne finie avec un élargissement $\eta = 0.5$. Ligne en pointillés : chaîne finie avec un élargissement $\eta = 1$. b) La partie complexe en η élargit les niveaux discrets en lorentziennes.

la partie complexe η de la fonction de GREEN du système isolé. La figure (4.14) montre qu'un ajustement judicieux de η conduit au même coefficient de transmission qu'un système réellement connecté à des réservoirs. Cependant dans le reste de cette étude, le paramètre η sera choisi de façon à ne pas modifier les propriétés de transport des électrons.

4.2.2 Résultats et discussion

L'évolution de l'énergie potentielle de la structure double-grille pour différentes tensions de grille V_G est représentée figure (4.15). Nous observons la diminution usuelle de la barrière de potentiel avec l'augmentation de la tension de grille : le transistor évolue de l'état bloqué vers l'état passant. La densité électronique le long du canal pour des tensions de grille variant de 0 à 0.3 V (figure (4.16)) confirme ce processus. A faible V_G la barrière de potentiel du canal s'oppose à la transmission d'électrons entre la source et le drain. La variation exponentielle de la densité électronique le long du canal témoigne d'un transport par effet tunnel à travers la barrière de potentiel, et permet de clairement différencier les électrons provenant de la source de ceux provenant du drain. Lorsque la tension de grille augmente, le transistor devient progressivement passant et la concentration de charges dans le canal augmente.

Dans une approche classique, n'incluant pas l'effet tunnel, les électrons ne peuvent pas traverser la barrière de potentiel du canal et leur concentration chute en tout début de canal (figure (4.17)). A tension et dimensions équivalentes, le modèle classique prévoit 10000 fois moins d'électrons en régime sous le seuil que le modèle quantique. Les caractéristiques simulées du courant de drain (I_D) en fonction de V_G de la structure double-grille de longueur $L_G = 10 \text{ nm}$ soumise à 2 températures sont représentées sur la figure (4.18). A température ambiante, le composant opère avec une pente sous le seuil S de 74.1 mV/décade. Lorsque la température diminue ($T = 100 \text{ K}$), le régime sous le seuil ne présente plus une pente constante en fonction de V_G . Ces deux courbes

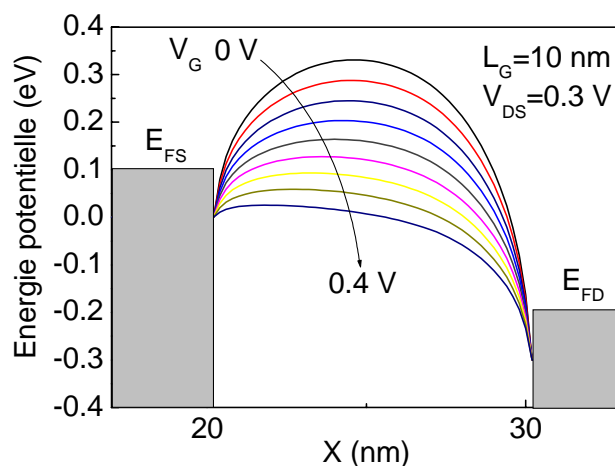


FIG. 4.15: Profils de l'énergie potentielle le long du canal sous une tension de drain $V_{DS} = 0.3\text{V}$. Les régions grisées représentent la source et le drain dont les niveaux de FERMÍ sont respectivement E_{FS} et E_{FD} . V_G varie de 0 à 0.4 V par pas de 0.05.

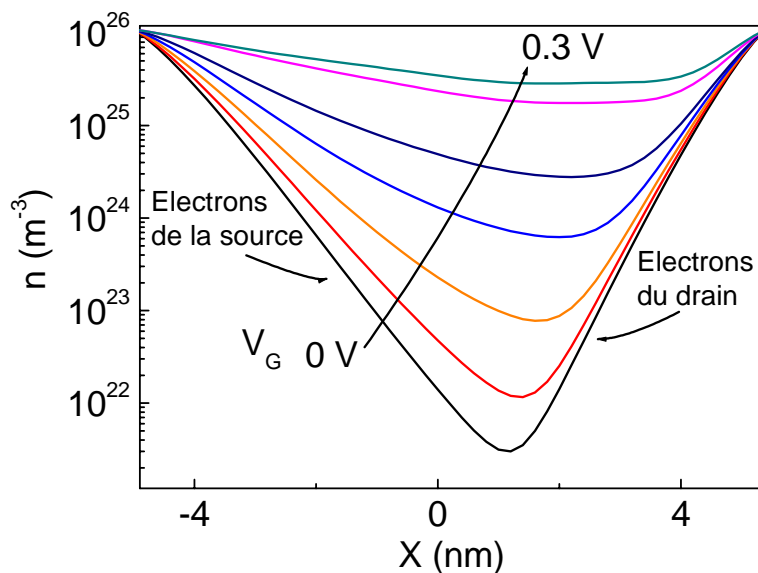


FIG. 4.16: Densité électronique le long du canal obtenue par le formalisme des fonctions de GREEN. A faible V_G l'évolution exponentielle du nombre de porteurs dans le canal traduit une domination de l'effet tunnel dans le transport. L'augmentation de V_G fait commuter le transistor et accroît la concentration d'électrons dans le canal. $V_{DS} = 0.2\text{V}$ et V_G varie de 0 à 0.3 V par pas de 0.05. $L_G = 10\text{ nm}$.

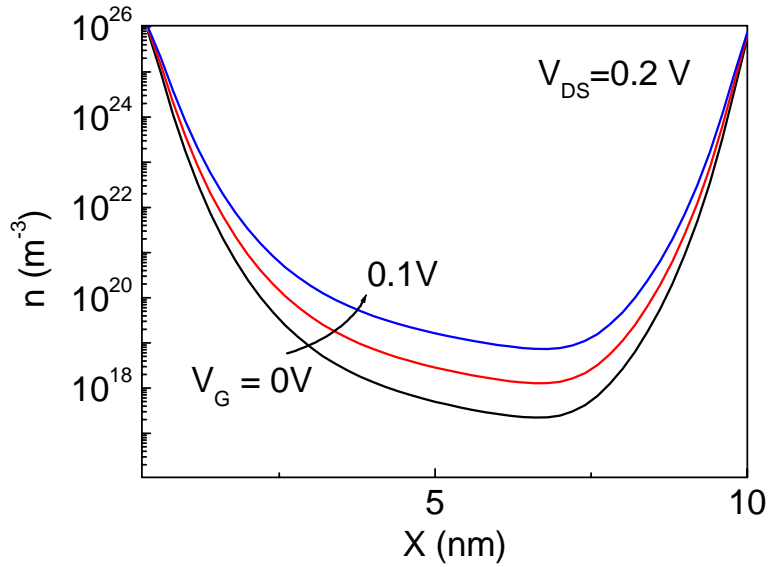


FIG. 4.17: Densité électronique le long du canal proposée par une approche classique de dérivation-diffusion (cf § (3.1.1)). A tension et dimensions équivalentes, la quantité d'électrons en régime sous le seuil chute d'un facteur 10^4 par rapport à l'approche quantique des fonctions de GREEN (figure (4.16)). $V_{DS} = 0.2$ V et V_G varie de 0 à 0.1 V par pas de 0.05. $L_G = 10$ nm.

séparent distinctement la contribution du courant tunnel de celle du courant thermique : la valeur de I_{OFF} (courant à $V_G = 0$ V) augmente d'un facteur 10^3 entre 100 K et 300 K. La diminution de la température conduit à une suppression du courant thermique par la barrière de potentiel du canal (figure (4.15)) et le courant tunnel devient alors la principale composante du courant total. L'observation de l'effet tunnel est encore plus manifeste si l'on divise le courant de drain en deux composantes : thermo-ionique et tunnel (figure (4.19)). Pour chaque valeur de tension de grille, nous définissons la composante tunnel par le courant résultant des électrons qui traversent la barrière de potentiel du canal et la composante thermo-ionique par le courant résultant des électrons d'énergie supérieure à la sus-dite barrière. La figure (4.20) représente les deux composantes du courant de drain d'un dispositif de longueur $L_G = 10$ nm à température ambiante. La pente sous le seuil S du courant total s'éloigne de la valeur idéale de 60 mV/décade alors que la composante thermo-ionique reste plus modérée avec une pente de 69 mV/décade : pour des longueurs de grille de 10 nm, la composante tunnel n'est plus négligeable à faible V_G et dégrade la pente sous le seuil. Les courbes $I_D - V_G$ pour six différentes valeurs de L_G sont représentées figure (4.21). Un composant de 16 nm offre une pente sous le seuil quasi-parfaite de 64 mV/décade. Lorsque la longueur de grille diminue, S augmente en suivant une loi surlinéaire (figure (4.22))⁹. En spécifiant un ratio minimum I_{ON}/I_{OFF} de 100 et une pente sous le seuil de 100 mV/décade, nous trouvons une limite d'intégration de 7 nm pour ce type de MOSFET double-grille ultime. Les caractéristiques $I_D - V_{DS}$ du MOSFET double-grille ultime sont proches de celles fournies par les transistors MOSFETs conventionnels (figure (4.23)). Dans les dispositifs balistiques, la saturation apparaît lorsque la barrière de potentiel du canal est

⁹La dépendance en puissance est symptomatique de l'évolution de la transmission par effet tunnel en fonction de l'épaisseur de la barrière.

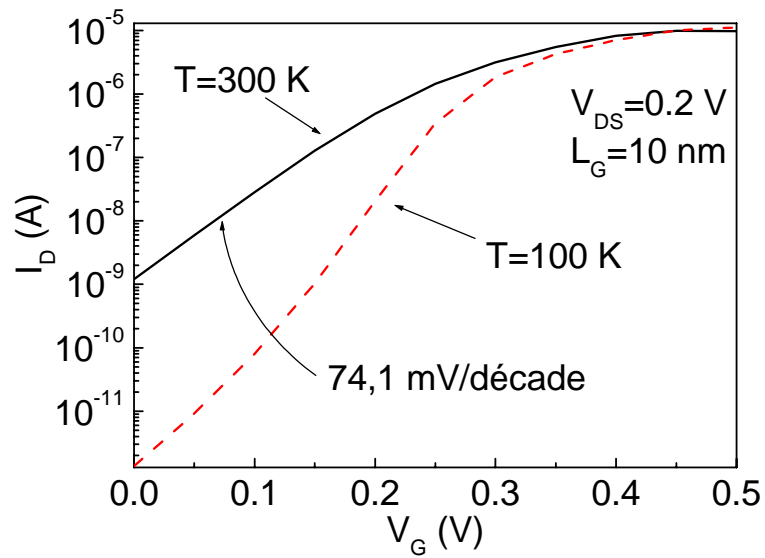


FIG. 4.18: Caractéristiques I_D en fonction de V_G simulées à deux températures, 100 et 300 K ($V_{DS} = 0.2$ V et $L_G = 10$ nm).

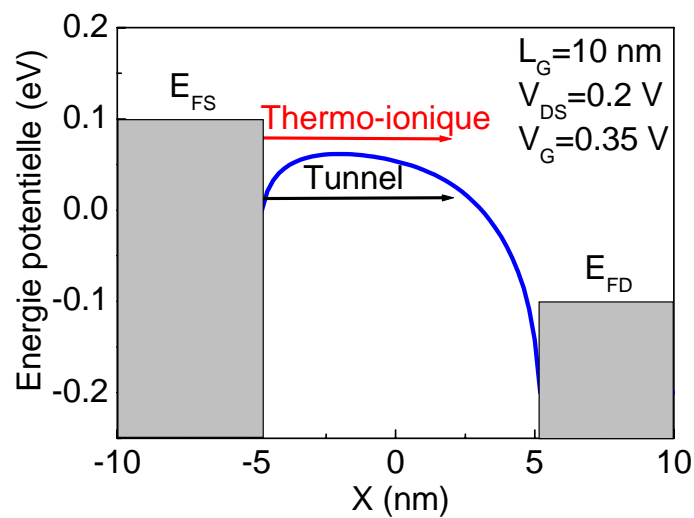


FIG. 4.19: Le courant total d'un nano-transistor peut se décomposer en une composante tunnel et une composante thermo-ionique.

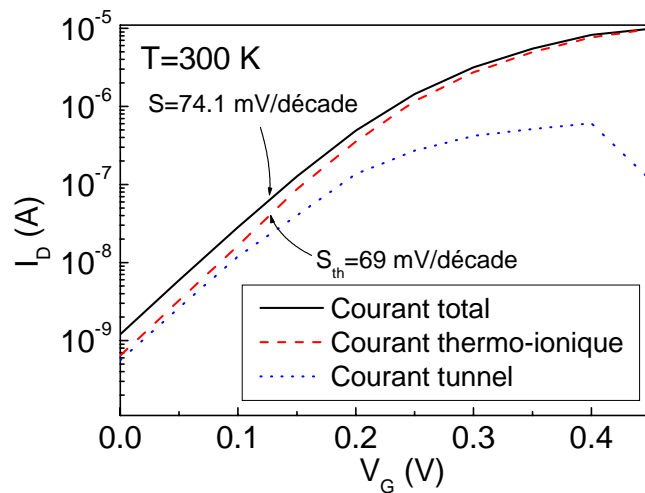


FIG. 4.20: Caractéristiques I_D-V_G calculées pour une longueur de grille $L_G = 10$ nm. Tous les autres paramètres sont identiques à ceux de la figure (4.18). Le courant total est la somme des composantes thermo-ionique et tunnel ($T = 300$ K).

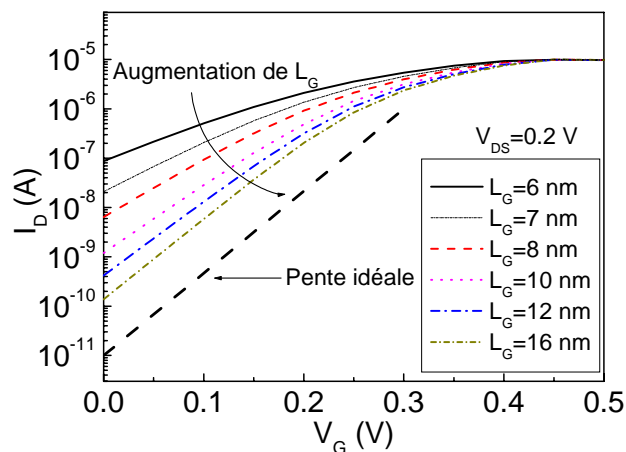


FIG. 4.21: Courbes I_D-V_G pour des composants aux dimensions comprises entre 6 à 16 nm. Le courant tunnel dégrade fortement la pente sous le seuil des petits composants.

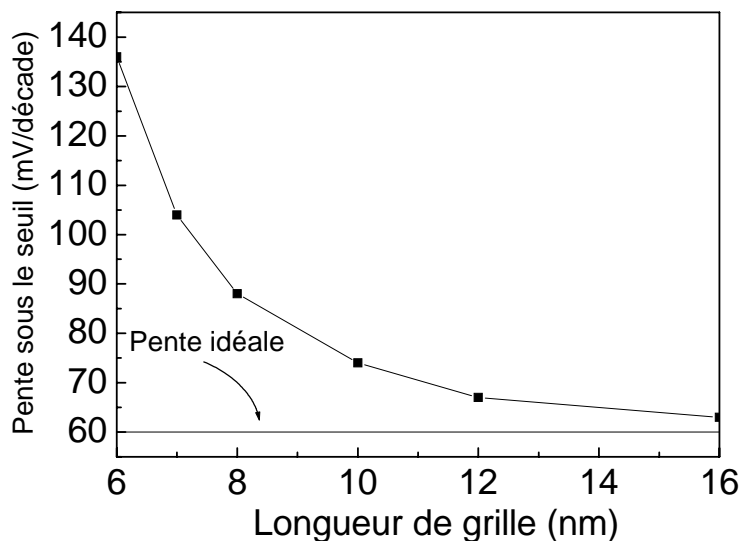


FIG. 4.22: *Pente sous le seuil en fonction de la longueur de grille.*

supprimée par la tension de grille positive et est due à "l'épuisement" de la source qui limite la quantité d'électrons injectés dans le canal. En effet, la distribution de FERMİ-DİRAC, qui tolère dans le meilleur des cas deux électrons par niveau d'énergie (spin), "freine" le transport des charges à travers le composant balistique (cf § (3.3.2)). L'intervalle énergétique d'émission des électrons étant étroit, une majoration du courant de saturation est obtenue en multipliant la densité électronique de la source ($\approx 10^{26} \text{ m}^{-3}$) par la vitesse de FERMİ des électrons de la source ($v_F = (1/\hbar) \times (\partial\varepsilon/\partial k)_{E_{FS}}$).

A l'état passant ($V_G = 0.45 \text{ V}$), la légère diminution du courant pour de fortes valeurs de V_{DS} , résulte de réflexions associées à la chute du potentiel entre la source et le drain le long d'un canal court. Sans barrière de potentiel, les particules balistiques transitent librement depuis la source vers le drain, mais des réflexions dues à la variation abrupte du potentiel sur une distance très courte (quelques nanomètres) demeurent : le transport de charges est alors d'autant plus perturbé que la tension V_{DS} augmente. Ce comportement est illustré par la figure (4.24), qui représente les coefficients de transmission à $V_G = 0.45 \text{ V}$ et V_{DS} variant de 0.25 à 0.4 V. Elle montre clairement que jusqu'à 0.1 eV au-dessus la bande de conduction du canal, la transparence diminue avec une hausse de la polarisation. La figure (4.23) montre également que la conductance du transistor dans l'état passant ($V_G = 0.45 \text{ V}$), est quantifiée en unité de G_0 , le quantum de conductance prévu en physique mésoscopique (cf § (3.3.1)) [170]. Dans les transistors à canal très confiné les énergies des modes transverses (le long de la direction \mathbf{y}), sont fortement dégénérées et la conductance présente des variations discontinues (en marches d'escalier) lorsque de nouveaux états résonnants participent au transport (cet effet a été récemment discuté par GUO *et al.* [171] et YAMADA [172] dans le cas du transistor à nanotube de carbone).

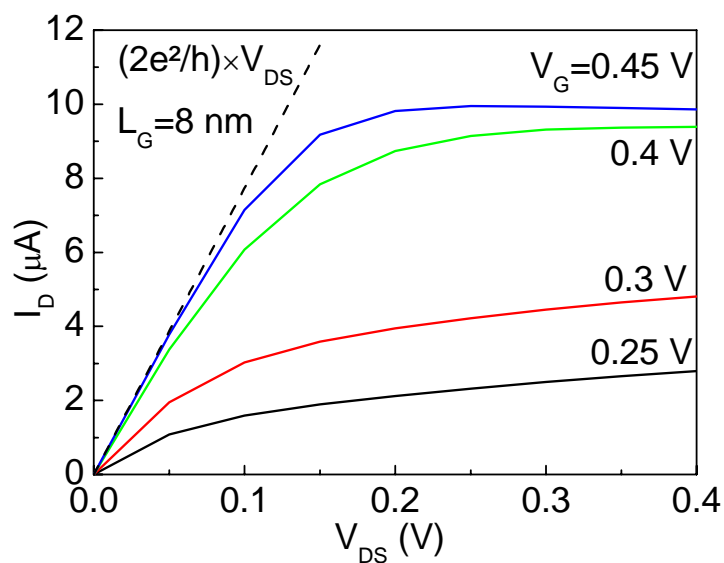


FIG. 4.23: Caractéristiques $I_D - V_{DS}$ sous différentes tensions de grille ($V_G = 0.25 - 0.45 \text{ V}$). La ligne en pointillés représente le courant fourni par le quantum de conductance $G_0 = 2e^2/h$ ($T = 300 \text{ K}$).

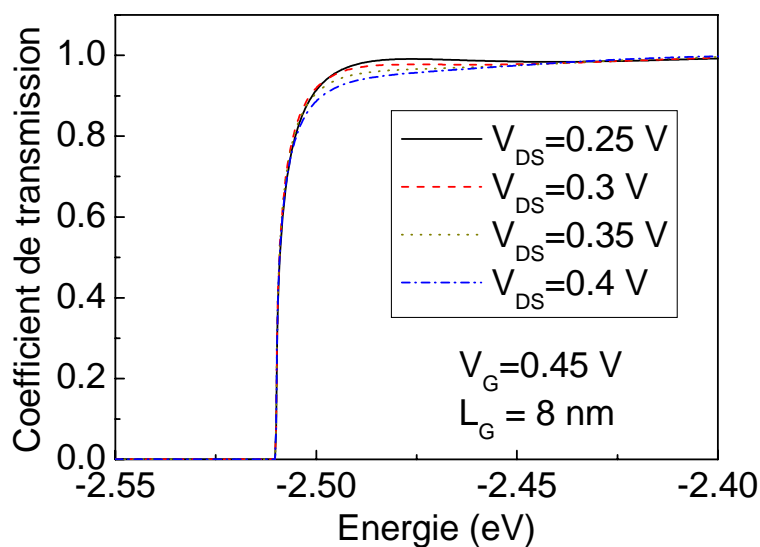


FIG. 4.24: Coefficients de transmission en fonction de l'énergie à $V_G = 0.45 \text{ V}$ pour différents V_{DS} . Le bas de bande de conduction du canal côté source est situé à 2β où $\beta = -1.25 \text{ eV}$.

4.3 Dépendance du courant tunnel source-drain en fonction de la longueur de grille et du travail de sortie de son métal

La section précédente a démontré le rôle néfaste de l'effet tunnel sur les caractéristiques des transistors de taille nanométrique. Dans cette partie, nous étudions en détail l'évolution du courant tunnel source-drain en fonction de la longueur de grille et du travail de sortie de son métal. Le courant de drain est à nouveau divisé en une composante thermo-ionique et une composante tunnel (figure (4.19)).

La figure (4.20) montre le courant de drain du transistor double-grille de longueur $L_G = 10$ nm. Jusqu'à cette échelle, l'impact de la composante tunnel reste minimal, à l'exception des faibles tensions de grille. Lorsque la tension de grille augmente, la disparition progressive de la barrière de potentiel dans le canal annule la composante tunnel et le courant de drain est entièrement gouverné par l'émission thermo-ionique. Calculons maintenant la caractéristique $I_D - V_G$ pour une longueur de grille $L_G = 6$ nm, en conservant les mêmes paramètres que pour la figure (4.20). Nous constatons (figure (4.25)), que la valeur du courant I_{OFF} ($V_G = 0$ V) augmente de plus d'une décade passant de 10^{-9} à 3.10^{-8} A : la barrière de potentiel du canal ne suffit plus à bloquer les électrons injectés depuis la source. Afin de maintenir le I_{OFF} constant, il est nécessaire d'accroître la hauteur de barrière dans le canal en augmentant le travail de sortie ϕ_M du métal de grille. La figure (4.26) représente la caractéristique $I_D - V_G$ du même composant dans lequel le travail de sortie du métal de grille a été augmenté via l'équation de POISSON de 0.3 V pour valoir désormais 4.623 V : le courant I_{OFF} est inférieur à 10^{-9} A, mais le régime de saturation est décalé vers de grandes valeurs de tensions de grille. La pente sous le seuil est également fortement dégradée par la composante tunnel ($S = 134$ m V/décade) . Le courant thermo-ionique présente en revanche une pente sous le seuil bien meilleure, de $S_{th} = 91$ m V/décade. A travers l'étude de 2 longueurs de grille ($L_G = 10$ et 6 nm) il ressort qu'un ajustement du travail de sortie du métal de grille s'impose pour chaque dimension, afin de conserver constant le courant de drain du transistor bloqué. D'autre part, une augmentation du travail de sortie s'accompagne d'un accroissement de la tension de seuil V_{TH} . Dans cette étude, nous définissons la tension seuil comme la valeur de la tension à l'intersection des deux pentes de la courbe $I_D - V_G$ représentée en échelle linéaire (figure (4.27)). L'évolution de I_{OFF} et V_{TH} en fonction de ϕ_M s'interprète en observant la barrière de potentiel d'un dispositif de longueur $L_G = 6$ nm, en fonction du travail de sortie de la grille (figure (4.28)). L'augmentation de la barrière de potentiel avec le travail de sortie conduit à une diminution du courant d'électrons tunnel et donc du courant I_{OFF} . En contre partie, la tension de grille nécessaire pour réduire la barrière de potentiel est plus importante : V_{TH} augmente. La valeur de V_{TH} étant associée à la vitesse de commutation (passage de l'état bloqué à l'état passant), le choix du travail de sortie du métal de grille doit résulter d'un délicat compromis entre performances électriques et vitesse de commutation. Nous avons tracé figure (4.29) les évolutions du courant I_{OFF} et de la tension de seuil V_{TH} en fonction du travail de sortie du métal de grille de longueur $L_G = 6$ nm. Comme expliqué ci-dessus, la tension de seuil croît avec le travail de sortie du transistor alors que le courant I_{OFF} évolue en

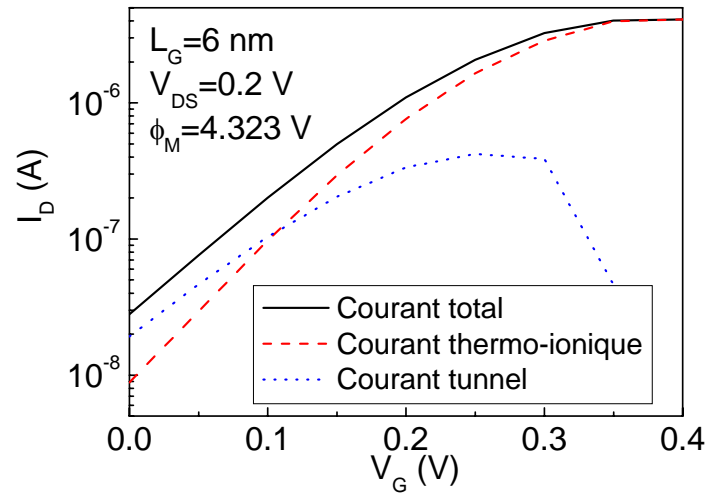


FIG. 4.25: Caractéristique I_D-V_G calculée pour une structure double-grille de longueur $L_G = 6 \text{ nm}$ et un travail de sortie identique à celui du composant de longueur $L_G = 10 \text{ nm}$: $\phi_M = 4.323 \text{ V}$. Le courant I_{OFF} est fortement dégradé.

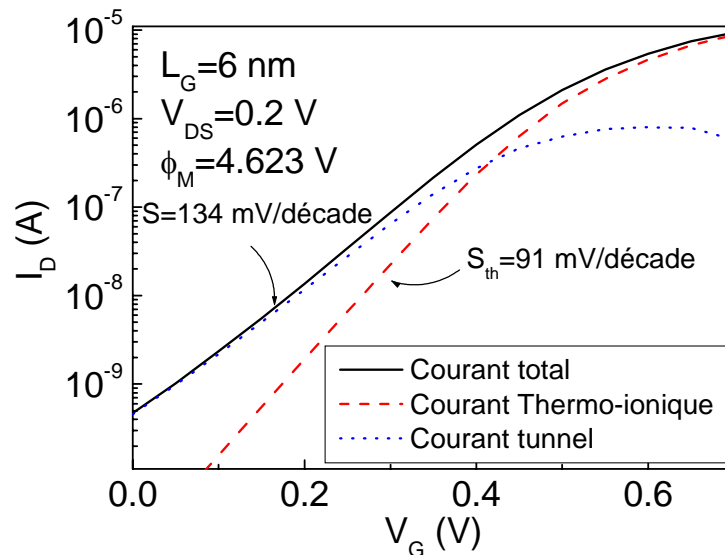


FIG. 4.26: Caractéristique I_D-V_G du dispositif double-grille de longueur $L_G = 6 \text{ nm}$, avec un travail de sortie $\phi_M = 4.623 \text{ V}$. Les autres paramètres sont identiques à ceux de la figure (4.25). Le courant I_{OFF} s'est amélioré ($< 10^{-9} \text{ A}$), au détriment de la tension de seuil qui avoisine 0.5 V . Le contrôle du dispositif s'acquiert en perdant en rapidité (vitesse de commutation entre I_{OFF} et I_{ON}).

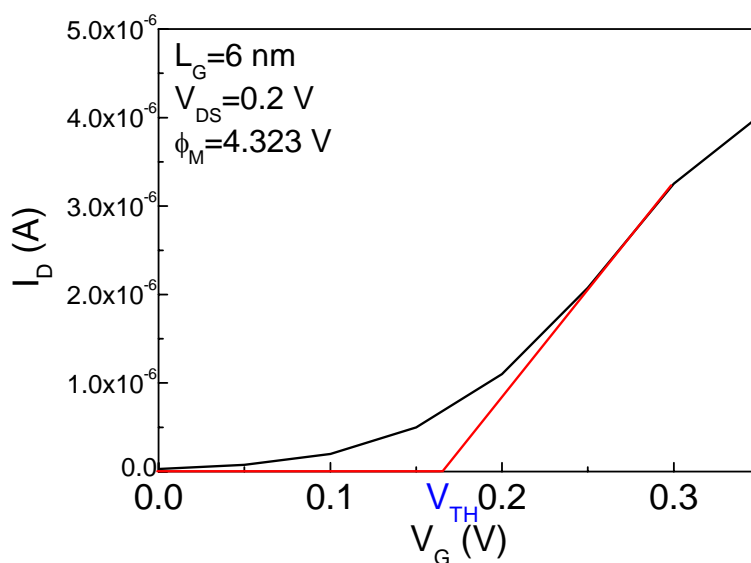


FIG. 4.27: Calcul géométrique de la tension de seuil.

sens inverse. Une variation de 0.5 V améliore ou dégrade dramatiquement le courant sous le seuil et décale également V_{TH} de la même quantité (~ 0.5 V). Pour la longueur de grille considérée, un ϕ_M de 4.47 V semble assurer un faible courant de fuite I_{OFF} (4.10^{-9} A), tout en conservant un V_{TH} raisonnable ($= 0.325$ V). En pratique, l'épaisseur effective d'oxyde peut également être utilisée comme un paramètre supplémentaire afin d'optimiser les performances du dispositif.

L'omniprésence de l'effet tunnel dans les nanoMOSFETs semble jusqu'à présent incontournable. Néanmoins, le courant tunnel source-drain joue-t-il toujours un rôle essentiel dans toutes les configurations? Considérons un dispositif de longueur de grille $L_G = 3$ nm et de travail de sortie identique à celui du transistor de 10 nm ($\phi_M = 4.323$ V). La figure (4.30) représente la caractéristique $I_D - V_G$ associée sous $V_{DS} = 0.2$ V. Pour toutes les valeurs de V_G , le pourcentage du courant tunnel n'excède pas celui du courant thermo-ionique. La diminution de la longueur de barrière de potentiel associée à la réduction de L_G , s'accompagne également d'une diminution de sa hauteur qui renforce la composante thermo-ionique au détriment la partie tunnel. Ainsi, à travail de sortie ϕ_M fixe, la fraction de courant tunnel n'évolue pas toujours en sens opposé de la longueur de grille. Au contraire, la figure (4.31) montre que le pourcentage du courant tunnel admet un maximum à $L_G = 5$ nm si ϕ_M reste égal à 4.323 V pour toutes les dimensions de grille. Une alternative est de modifier le travail de sortie à chaque longueur de grille afin de conserver le même courant $I_{OFF} \cong 10^{-9}$ A. Nous observons alors que la fraction du courant tunnel augmente avec la diminution de la longueur du canal. En deçà de 5 nm, la fraction du courant tunnel sature à sa valeur maximale, indiquant que l'effet tunnel contrôle totalement le courant du transistor. L'inconvénient majeur de ce type de calibrage est qu'il conduit à de très grandes valeurs de travaux de sortie ($\phi_M > 6$ V) et en conséquence entraîne une augmentation considérable de V_{TH} .

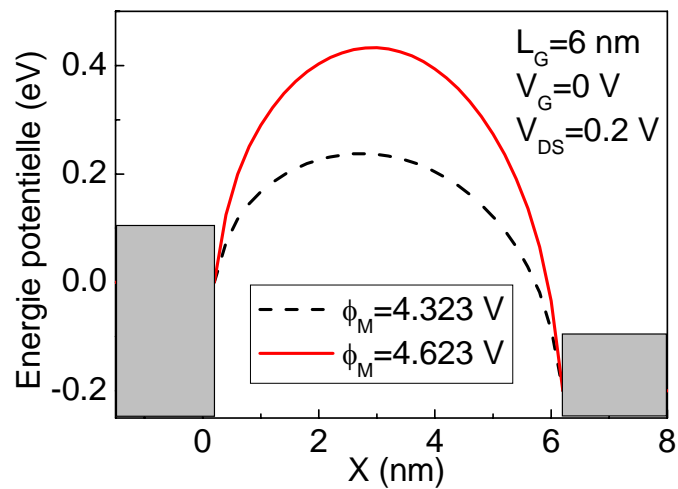


FIG. 4.28: Energie potentielle le long du canal de longueur $L_G = 6$ nm pour deux travaux de sortie de grille ($\phi_M = 4.323$ et 4.623 V). $T = 300$ K.

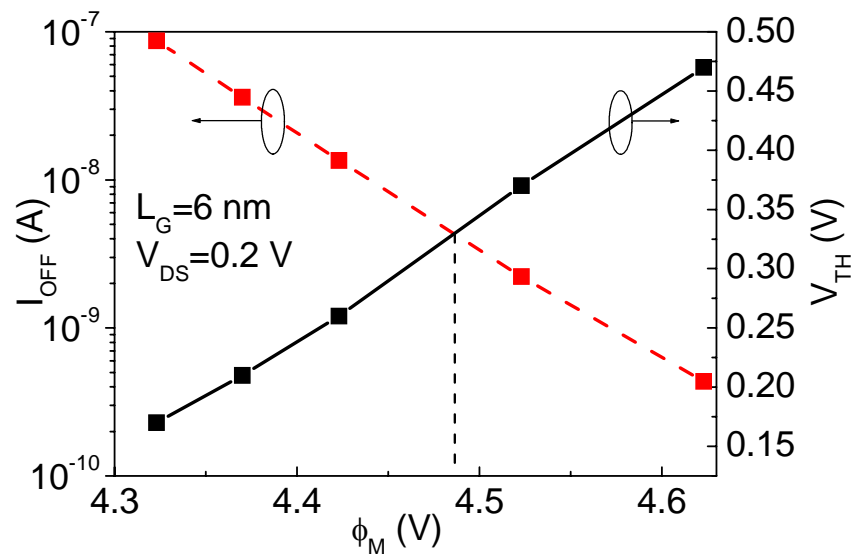


FIG. 4.29: Courant I_{OFF} et tension de seuil V_{TH} en fonction du travail de sortie du métal de grille ϕ_M . Un bon compromis entre performances et vitesse de commutation semble atteint pour $\phi_M = 4.47$ V.

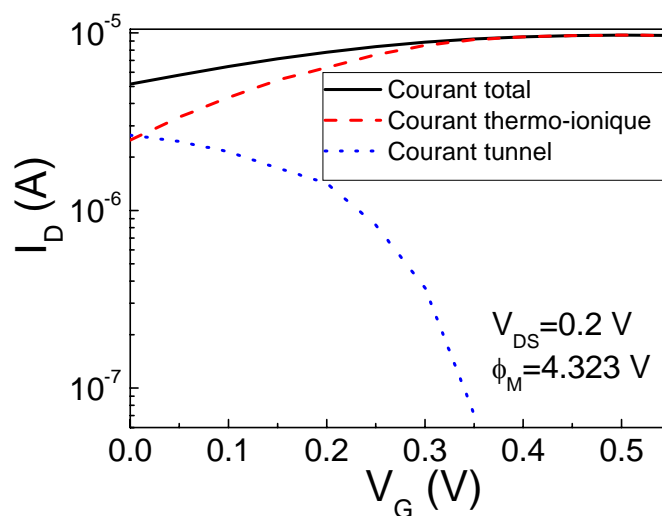


FIG. 4.30: Caractéristique I_D-V_G calculée pour une structure double-grille de longueur de grille $L_G = 3 \text{ nm}$ et $\phi_M = 4.323 \text{ V}$. Étonnamment, la composante tunnel du courant n'est plus dominante $T = 100 \text{ K}$.

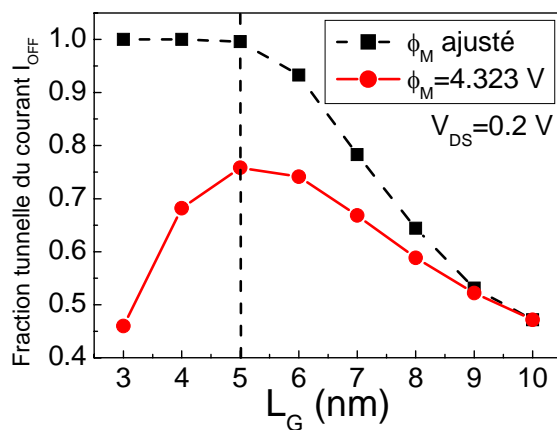


FIG. 4.31: Fraction du courant de drain tunnel @ $V_G = 0 \text{ V}$ en fonction de la longueur de grille, i) pour un travail de sortie de grille fixe $\phi_M = 4.323 \text{ V}$, ii) pour un travail de sortie de grille ajusté afin de maintenir le même courant $I_{OFF} \sim 10^{-9} \text{ A}$.

4.4 Conclusion

Nous avons simulé le transport d'électrons dans un MOSFET double-grille à canal de conduction confiné en utilisant la théorie des liaisons fortes exprimée dans le formalisme des fonctions de GREEN hors-équilibre. Plusieurs résultats sont à souligner. Modéliser le canal par une chaîne linéaire d'atomes ne change pas l'essentiel des propriétés électriques du transistor MOSFET double-grille. Bien que l'effet tunnel modifie les caractéristiques de courant et dégrade la pente sous le seuil, un transistor ultimement confiné répond de façon satisfaisante à sa tension de grille jusqu'à une longueur de grille de ~ 7 nm. En ajustant le travail de sortie des grilles, nous avons montré que la conduction à travers le canal est entièrement gouvernée par l'effet tunnel pour des longueurs inférieures à 5 nm. Par ailleurs, le confinement du canal conduit à une quantification de la conductance en unité de G_0 , le quantum de conductance annoncé par la physique mésoscopique. Enfin, la théorie des liaisons fortes fait apparaître (simplement dans le cas présent) la structure atomique du canal et permet donc de considérer l'influence de défauts cristallins ponctuels (impuretés, lacunes...) de la zone active sur le transport électronique. Ces aspects seront traités dans le prochain chapitre.

Chapitre 5

Modélisation 3D de MOSFETs multi-grilles

La diminution effrénée des dimensions des transistors conduit aujourd'hui à des MOSFETs ultimes SOI de taille nanométrique. Des composants présentant des longueurs de grille aussi petites que 5 et 6 nm [194, 195] ont d'ailleurs été fabriqués. Dans ce contexte, les MOSFETs multi-grilles constituent des candidats intéressants grâce à l'efficacité du contrôle des charges du canal par la tension de grille. Plusieurs groupes ont récemment réalisés des MOSFETs Tri-gate et Gate-all-around (cf § (1.3.3)) à partir de nano-fils de silicium à section rectangulaire [49,196]. Contrairement aux MOSFETs "bulk" et aux double-grilles plan, les multi-grilles ont des architectures 3D dans lesquelles longueur, hauteur et profondeur jouent un rôle principal. Jusqu'à présent, les dimensions transverses des nano-fils n'atteignent pas les 20 nm et la compréhension des composants ultra-minces (section < 10 nm) constitue un défi majeur.

Au chapitre 4 nous avons présenté une solution balistique d'un MOSFET double-grille simplifié dans lequel l'épaisseur du canal était réduite à une chaîne atomique : l'influence du confinement vertical, notamment pour obtenir l'électroneutralité des extensions des réservoirs, n'était pas prise en compte. L'objectif de ce chapitre est de modéliser des architectures innovantes 3D réalistes à l'aide du formalisme des fonctions de GREEN hors-équilibre. Afin de simplifier les étapes numériques, nous adopterons l'approche dite de l'"espace des modes" (pour l'anglais "mode-space"), qui consiste à séparer le traitement des effets du confinement transverse de celui du transport électronique le long de l'axe source-drain (section 5.1). Après avoir appliqué ce modèle au MOSFET Gate-All-Around en détaillant l'évolution des paramètres physiques internes (section 5.2), nous décrirons l'ensemble des architectures 3D (Tri-gate, Π -gate, Ω -gate) et étudierons leur comportement électrique (section 5.3). Enfin nous concluons.

5.1 Approximation de l'espace des modes

La résolution quantique exacte de systèmes 2D ou 3D nécessite un outil de calcul numérique très puissant. De plus en plus de laboratoires et industries s'équipent donc de *cluster*, c'est-à-dire d'une grappe d'ordinateurs connectés entre eux (jusqu'à 150 unités) et travaillant en parallèle sur la même tâche. Cette évolution a incité les chercheurs à proposer des alternatives physiques et/ou numériques susceptibles de réduire la taille des calculs. L'approximation de l'espace des modes est le fruit de cette volonté et permet de transformer, sous certaines conditions, un problème 2D ou 3D en plusieurs problèmes 1D indépendants, et traitables par un simple PC. La section suivante décrit l'approche "mode-space" et définit son domaine de validité.

5.1.1 Confinement quantique dans les architectures MOSFETs innovantes

Les prévisions des industriels indiquent que la géométrie SOI ("Silicon-On-Insulator") de par sa bonne immunité contre les effets canaux courts (cf § (1.2)), pourrait être le dispositif des futures technologies CMOS. Néanmoins, la diminution de la longueur de grille d'une génération de transistors à une autre, s'accompagne également de la réduction des dimensions transverses afin de contenir la progression des effets canaux courts. Le confinement des porteurs de charges dans la section transverse qui en résulte induit une quantification des états perpendiculaires et une distribution de ces derniers en sous-bandes. Ce comportement modifie ostensiblement les caractéristiques électriques des MOSFETs SOI à canaux courts et l'évaluation quantitative de son impact constitue une donnée essentielle.

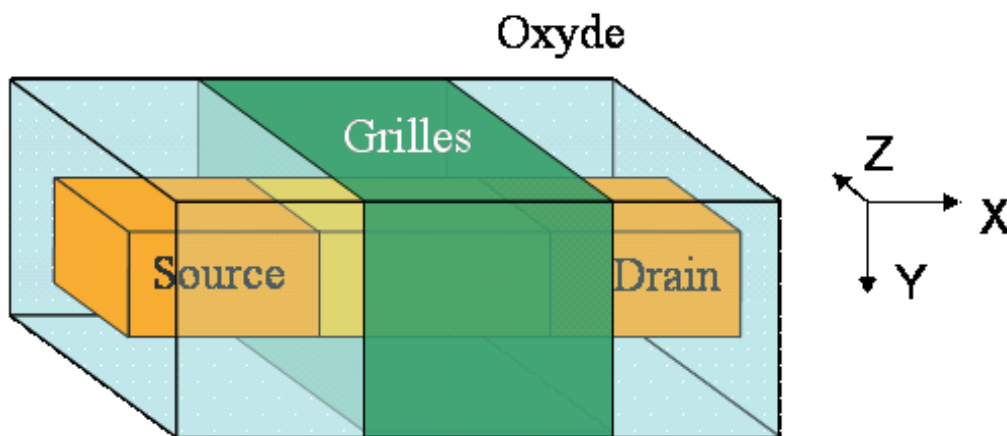


FIG. 5.1: Schéma d'un transistor MOSFET *Gate-All-Around* (GAA).

Considérons le transistor MOSFET *Gate-all-around* représenté sur la figure (5.1). Aux petites dimensions, les effets quantiques affectent sa tension de seuil (confinement suivant les axes \mathbf{y} et \mathbf{z}) et son courant *off* (effet tunnel entre source et drain). Pour traiter un tel système, les approches de modélisation quantique reposent sur une résolution auto-cohérente des équations

de SCHRÖDINGER et de POISSON. La solution de l'équation de SCHRÖDINGER peut être obtenue suivant plusieurs niveaux de complexité en fonction de la nature du dispositif étudié et du degré d'exactitude désiré. Dans le cas de transistors *bulk*, deux groupes ont développé des solutions 2D basées sur le formalisme des fonctions de GREEN hors-équilibre [197, 198, 200]. Leurs approches, qui reposent sur une discrétisation dans l'espace réel de l'hamiltonien 2D exprimé dans l'approximation de la masse effective, sont numériquement très coûteuses. La discrétisation de l'espace réel est néanmoins inévitable pour traiter correctement le transport quantique dans les MOSFETs massifs, car le confinement des porteurs évolue le long du canal et aucune symétrie ne permet de réduire la taille du problème. Il en est tout autrement pour les MOSFETs SOI ultimes, dans lesquels les charges mobiles restent confinées quantiquement tout le long du canal. Ces géométries permettent de contourner la difficulté numérique associée à la résolution dans l'espace réel (sans compromettre l'exactitude du résultat) en exprimant la partie de l'hamiltonien confiné dans l'espace des modes et en limitant les calculs aux seuls modes occupés (≈ 2 ou 3 sous-bandes par vallée du silicium). Comparée avec succès à une résolution complète de l'espace réel dans le cas du transistor SOI double-grille 2D [104, 199] l'approche "mode-space" a été récemment appliquée au transistor à nano-fil de silicium 3D [201, 202]. Afin de justifier intuitivement la solution de l'espace des modes, considérons un plan atomique vertical (\mathbf{y}, \mathbf{z}) du transistor de la figure (5.1) à \mathbf{x} fixé. Les porteurs de ce plan, fortement confinés suivant les deux directions \mathbf{y} et \mathbf{z} , sont distribués dans un réseau discret de sous-bandes. Les énergies des états transverses sont quantifiées et leur éloignement croît avec le confinement. On peut ainsi supposer qu'un transistor GAA ultime, présente un espacement entre sous-bandes suffisant pour ignorer le couplage inter-bande. Chaque sous-bande est ainsi traitée séparément, et le système initialement 3D devient un problème N fois 1D où N représente le nombre de sous-bandes occupées.

5.1.2 Théorie

Considérons le MOSFET GAA représenté sur la figure (5.1). L'interface Si/SiO₂ est perpendiculaire à la direction cristalline $\langle 100 \rangle$. Le système, traité en liaisons fortes, est supposé en première approximation présenter une structure cubique simple (figure (5.2)) et la distance inter-atomique a est identique suivant toutes les directions ($a = 0.2$ nm). Chaque atome possède une orbitale et seules les interactions entre premiers voisins sont envisagées. Cette description est équivalente à basse énergie à l'approximation de la masse effective (cf § (4.1.2)), mais permet de faire apparaître la structure atomique du composant (bien que de façon simpliste) et de modéliser l'influence de défauts ponctuels (cf chapitre 6). La fonction de GREEN est maintenant calculée afin d'obtenir la densité électronique dans la zone active et le courant de drain. En régime balistique la méthode des fonctions de GREEN est équivalente à la résolution directe de l'équation de SCHRÖDINGER avec des conditions aux limites ouvertes. Une description quantique 3D exacte étant numériquement peu envisageable, nous adoptons la représentation de l'espace des modes dans les directions du confinement des grilles. Les différentes étapes de cette approche, dont les conditions de validité seront explicitées par la suite, sont ici décrites. Nous résolvons tout d'abord l'hamiltonien 2D H_{conf} dans chaque plan atomique n perpendiculaire à

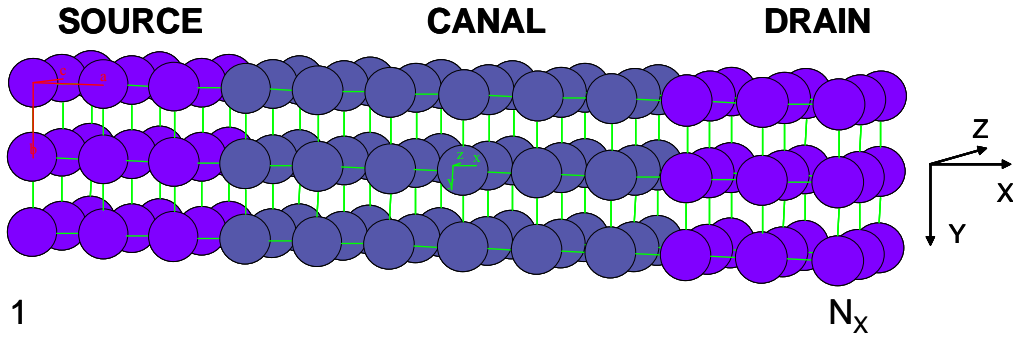


FIG. 5.2: *Le silicium et son oxyde sont supposés, en première approximation, présenter une structure cubique simple. Cette approche, équivalente à l'approximation de la masse effective, nous permettra de mesurer l'impact de défauts ponctuels (cf chapitre 6). Par volonté de clarté la structure des oxydes de grille n'est pas représentée.*

l'axe source-drain :

$$H_{conf} |\psi_{n,i}\rangle = \varepsilon_{n,i} |\psi_{n,i}\rangle, \quad (5.1)$$

avec :

$$\begin{aligned} H_{conf} &= \sum_{l,m} V_{lmn,lmn} |l, m, n\rangle \langle l, m, n| \\ &+ \beta_y \sum_{l,m} |l, m, n\rangle \langle l, m+1, n| \\ &+ \beta_z \sum_{l,m} |l, m, n\rangle \langle l+1, m, n|, \end{aligned} \quad (5.2)$$

où $|l, m, n\rangle$ est le ket représentant l'orbitale de l'atome de coordonnées $z = l \times a$, $y = m \times a$ du plan vertical n . $V_{lmn,lmn}$ représente l'élément de matrice du potentiel électrostatique supposé diagonal dans la base atomique conformément à la discussion de la section (4.1.2) sur la méthode des liaisons fortes. $|\psi_{n,i}\rangle$ et $\varepsilon_{n,i}$ sont respectivement les vecteurs et les valeurs propres des plans transverses. Enfin β_y et β_z sont les couplages entre premiers voisins le long des directions correspondantes. Elles sont reliées aux masses effectives par les expressions suivantes (cf § (2.3.2)) :

$$\beta_y = -\frac{\hbar^2}{2m_y^* a^2}, \text{ et } \beta_z = -\frac{\hbar^2}{2m_z^* a^2}. \quad (5.3)$$

La structure cubique simple est étendue aux oxydes de grille dont la description est incluse dans l'hamiltonien H_{conf} autorisant ainsi la pénétration de la fonction d'onde $|\psi_{n,i}\rangle$. Les mêmes couplages β_y et β_z sont considérés puisqu'il a été vérifié que la différence de gap entre le Si et le SiO₂ déterminait pour l'essentiel la pénétration de l'état transverse dans l'oxyde. Les éléments de matrice du potentiel électrostatique représentant l'oxyde sont donc augmentés de 3 eV par

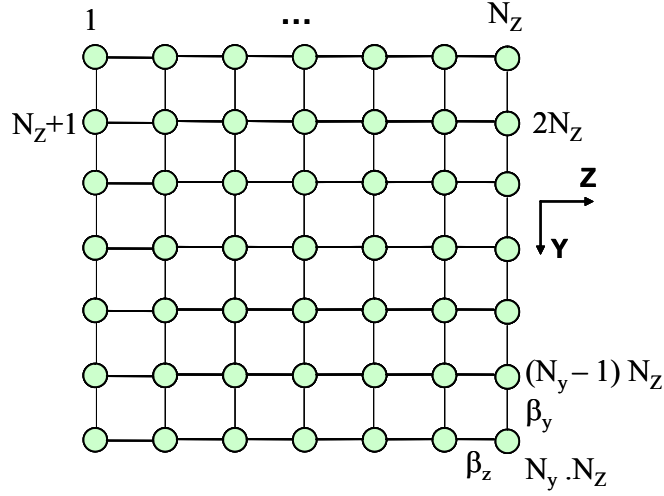


FIG. 5.3: Représentation schématique et numérotation des atomes de la section transverse du MOSFET GAA possédant N_Y atomes suivant la direction \mathbf{y} et N_Z suivant la direction \mathbf{z} . Les énergies de couplage des oxydes ont une influence négligeable face à la barrière de potentiel d'interface et sont donc choisies égales à celles du silicium [203].

rapport à ceux du silicium et l'expression matricielle de l'équation (5.1) s'écrit :

$$H_{conf} = \begin{bmatrix} V_{1,n} & \beta_z & 0 & \cdots & 0 & \beta_y & 0 & \cdots & 0 \\ \beta_z & V_{2,n} & \beta_z & 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \beta_z & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & 0 & \ddots & V_{i,n} & \ddots & \ddots & \ddots & 0 & \beta_y \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ \beta_y & \ddots & \ddots & \ddots & \ddots & V_{N_Z+1,n} & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \beta_z & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \beta_z & V_{(N_Y \cdot N_Z)-1,n} & \beta_z \\ 0 & \cdots & 0 & \beta_y & 0 & \cdots & 0 & \beta_z & V_{N_Y \cdot N_Z,n} \end{bmatrix}, \quad (5.4)$$

en utilisant les notations de la figure (5.3). L'approximation de l'espace des modes consiste ensuite à représenter l'hamiltonien 3D en considérant que les sous-bandes sont indépendantes. En d'autres termes, l'interaction entre $|\psi_{n,i}\rangle$ d'indice i différent est négligée :

$$H_{3D} = \begin{bmatrix} h_1 & 0 & \cdots & \cdots & \cdots \\ 0 & h_2 & 0 & \cdots & \cdots \\ \cdots & 0 & \ddots & 0 & \cdots \\ \cdots & \cdots & 0 & h_i & 0 \\ \cdots & \cdots & \cdots & 0 & \ddots \end{bmatrix}, \quad (5.5)$$

où

$$h_i = \begin{bmatrix} \varepsilon_{1,i} & \beta_x & 0 & \cdots \\ \beta_x & \varepsilon_{2,i} & \beta_x & 0 \\ 0 & \beta_x & \ddots & \beta_x \\ \cdots & 0 & \beta_x & \varepsilon_{N_X,i} \end{bmatrix} \quad (5.6)$$

est l'hamiltonien de la sous-bande i , N_X est le nombre d'atomes le long de l'axe source-drain et β_x est l'énergie de couplage suivant \mathbf{x} reliée à la masse effective par (cf § (4.1.2)) :

$$\beta_x = -\frac{\hbar^2}{2m_x^* a^2}. \quad (5.7)$$

En théorie, les hamiltoniens h_i de toutes les sous-bandes devraient être calculés, mais concrètement seules les premières sous-bandes sont occupées et participent à la conduction électronique. Il en résulte un gain numérique substantiel puisque le problème 3D de dimensions $(N_X \times N_Y \times N_Z)^2$ se transforme en un problème N^1 fois 1D dont la dimension est définie par le nombre d'atomes suivant l'axe \mathbf{x} ($= N_X$)².

Connaissant les hamiltoniens des sous-bandes peuplées (h_i), le transport 1D à travers chaque sous-bande est calculé. La fonction de GREEN retardée de la sous-bande i s'exprime comme :

$$G_i(\varepsilon) = \lim_{\eta \rightarrow 0^+} [(\varepsilon + i\eta)I - h_i - \Sigma^i], \quad (5.8)$$

où Σ^i est la matrice de la self-énergie des contacts qui modélise l'élargissement des niveaux d'énergie du système fini lors du couplage avec les réservoirs d'électrons. Le transport étant toujours 1D, leur expression reste identique à celle du § (4.2.1) au décalage en énergie près (confinement transverse). On obtient donc :

$$\Sigma^i = \Sigma_S^i + \Sigma_D^i. \quad (5.9)$$

Σ_S^i et Σ_D^i s'expriment en fonction de l'énergie de couplage réservoir-zone active ($= \beta_x$) et de la fonction de GREEN de surface des réservoirs. En reprenant les équations (4.34) et (4.36) la self-énergie de la source s'exprime de la façon suivante :

$$\Sigma_S^i = \begin{bmatrix} \beta_x g_S^i \beta_x & 0 & \cdots & \cdots & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 & 0 \\ 0 & \cdots & \cdots & 0 & 0 \end{bmatrix}, \quad (5.10)$$

où g_S^i est la fonction de GREEN du contact de la source qui représente une chaîne semi-infinie

¹ N représente le nombre de sous-bandes.

dont l'énergie potentielle est ajustée à celle de la sous-bande i :

$$g_S^i(\varepsilon) = \frac{\varepsilon + \varepsilon_{1,i}}{2\beta_x^2} \left(1 - \sqrt{1 - \frac{4\beta_x^2}{(\varepsilon + \varepsilon_{1,i})^2}} \right) \quad \text{si } |\varepsilon + \varepsilon_{1,i}| > 2|\beta_x|, \quad (5.11)$$

$$g_S^i(\varepsilon) = \frac{\varepsilon + \varepsilon_{1,i}}{2\beta_x^2} - \frac{i}{|\beta_x|} \sqrt{1 - \left| \frac{\varepsilon + \varepsilon_{1,i}}{2\beta_x} \right|^2} \quad \text{si } |\varepsilon + \varepsilon_{1,i}| < 2|\beta_x|. \quad (5.12)$$

De façon similaire la self-énergie du drain s'écrit :

$$\Sigma_D^i = \begin{bmatrix} 0 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 & 0 \\ 0 & \cdots & \cdots & 0 & \beta_x g_D^i \beta_x \end{bmatrix}, \quad (5.13)$$

où g_D^i est la fonction de GREEN du contact de drain :

$$g_D^i(\varepsilon) = \frac{\varepsilon + \varepsilon_{N_X,i}}{2\beta_x^2} \left(1 - \sqrt{1 - \frac{4\beta_x^2}{(\varepsilon + \varepsilon_{N_X,i})^2}} \right) \quad \text{si } |\varepsilon + \varepsilon_{N_X,i}| > 2|\beta_x|, \quad (5.14)$$

$$g_D^i(\varepsilon) = \frac{\varepsilon + \varepsilon_{N_X,i}}{2\beta_x^2} - \frac{i}{|\beta_x|} \sqrt{1 - \left| \frac{\varepsilon + \varepsilon_{N_X,i}}{2\beta_x} \right|^2} \quad \text{si } |\varepsilon + \varepsilon_{N_X,i}| < 2|\beta_x|. \quad (5.15)$$

De telles self-énergies permettent de réaliser des contacts parfaits exempts de réflexion.

La fonction de GREEN de la sous-bande i obtenue, nous définissons deux nouvelles quantités :

$$\Gamma_S^i = i \left(\Sigma_S^i - (\Sigma_S^i)^\dagger \right) \quad \text{et} \quad \Gamma_D^i = i \left(\Sigma_D^i - (\Sigma_D^i)^\dagger \right), \quad (5.16)$$

qui ont la dimension d'une énergie et représentent la force d'interaction entre les réservoirs et la zone active. Les fonctions de la densité spectrale des contacts s'expriment sous la forme :

$$A_S^i = G_i \Gamma_S^i G_i \quad \text{et} \quad A_D^i = G_i \Gamma_D^i G_i. \quad (5.17)$$

Ce sont des matrices de dimensions (N_X^2) dont les termes diagonaux représentent la densité locale d'états de chaque atome. Bien que le système soit dans un état hors-équilibre, les électrons injectés depuis les réservoirs de la source et du drain conservent leur distribution initiale (transport balistique). La fonction spectrale de la source est donc occupée suivant la distribution de FERMÍ-DIRAC de la source (f_S) alors que la fonction spectrale du drain est occupée suivant la distribution de FERMÍ-DIRAC du drain (f_D).

La matrice de la densité d'électrons 1D associée à la sous-bande i (le long de l'axe source-drain) est définie par :

$$n_i(\varepsilon) = \frac{1}{\pi a} \times [f_S(\varepsilon) A_S^i + f_D(\varepsilon) A_D^i]. \quad (5.18)$$

Pour obtenir la densité électronique totale 3D, il suffit d'intégrer l'équation (5.18) sur l'énergie ε , de sommer les contributions de chaque vallée et sous-bande, de les multiplier par le module au carré de l'état transverse associé $|\psi_{n,i}(y, z)|^2$ et de répéter l'opération pour tous les plans n :

$$n_{l,m,n}^{3D} = \sum_{l,m,n} \sum_i |\langle l, m, n | \psi_{n,i} \rangle|^2 \int_0^{+\infty} n_i(\varepsilon) d\varepsilon, \quad (5.19)$$

où i représente la somme sur les premières sous-bandes de chaque vallée. La densité électronique rétro-agit ensuite sur l'équation de POISSON 3D² afin d'obtenir un système auto-cohérent.

L'auto-cohérence atteinte, le courant de drain est calculé en sommant les courants associés à tous les modes peuplés. Pour ce faire, le système est clivé en deux régions et le courant traversant le plan séparant ces deux régions est déterminé. La matrice de la fonction de GREEN retardée de la sous-bande i du système non-couplé s'écrit :

$$G_i^0(\varepsilon) = \lim_{\eta \rightarrow 0^+} \begin{bmatrix} (\varepsilon + i\eta)I_{11} - (h_i)_{11} - \Sigma_S^i & 0 \\ 0 & (\varepsilon + i\eta)I_{22} - (h_i)_{22} - \Sigma_D^i \end{bmatrix} \quad (5.20)$$

où $(h_i)_{11}$ (respectivement $(h_i)_{22}$) est la matrice de l'hamiltonien dans la région 1 (respectivement dans la région 2) du mode i . Soit C_i la matrice de couplage assurant la connexion pour la sous-bande i entre les régions 1 et 2 :

$$C_i = \begin{bmatrix} 0 & (h_i)_{12} \\ (h_i)_{21} & 0 \end{bmatrix}, \quad (5.21)$$

où $(h_i)_{12}$ et $(h_i)_{21}$ sont les matrices des blocs non-diagonaux de l'hamiltonien du mode i . En se reportant au § (3.3.3), l'expression finale du courant³ traversant la sous-bande i est :

$$I_i = -\frac{4\pi e}{\hbar} \int_{-\infty}^{+\infty} d\varepsilon Tr_1 \left[N_0^i \Lambda_i^\dagger C_i N_0^i C_i \Lambda_i \right] (f_S(\varepsilon) - f_D(\varepsilon)), \quad (5.22)$$

où Tr_1 désigne la trace restreinte à la base des orbitales de la région 1, $N_0^i (= -(1/\pi) \text{Im}(G_0^i))$ est la matrice densité à l'équilibre du système non-couplé, et $\Lambda_i = (I - G_0^i C_i G_0^i C_i)^{-1}$. Enfin, le courant total s'obtient en sommant sur les premiers modes de chaque vallée :

$$I = \sum_i I_i. \quad (5.23)$$

Nous modélisons le cas du Si par six bandes de liaisons fortes dont la partie basse énergie doit reproduire l'approximation de la masse effective. La figure (5.4) représente donc les six ellipsoïdes équivalents du silicium volumique. Le confinement suivant \mathbf{y} et \mathbf{z} , sépare les six vallées en trois

²Les conditions aux limites de l'équation de POISSON 3D seront discutées au § (5.1.4).

³L'équation (5.22) est la généralisation de la formule de LANDAUER (cf § (3.3.1)) qui relie le courant au coefficient de transmission $T(\varepsilon) = 4\pi^2 Tr_1 [N_0 \Lambda^\dagger C N_0 C \Lambda]$.

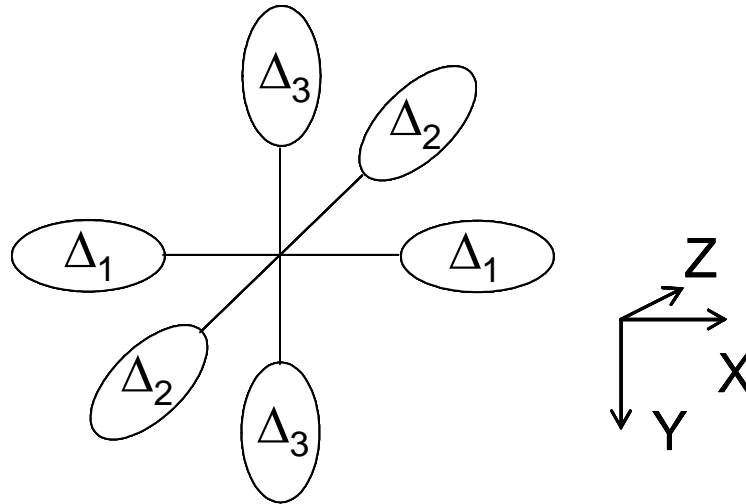


FIG. 5.4: Représentation des six ellipsoïdes équivalents du silicium volumique. Sous l'influence du confinement, trois paires prennent des énergies différentes.

paires dont les masses effectives dans les trois directions sont :

· Vallées Δ_1 :

$$- m_x^* = m_l, \text{ et } m_y^* = m_z^* = m_t,$$

· Vallées Δ_2 :

$$- m_x^* = m_z^* = m_t, \text{ et } m_y^* = m_l,$$

· Vallées Δ_3 :

$$- m_x^* = m_y^* = m_t, \text{ et } m_z^* = m_l.$$

Les masses effectives des vallées Δ_1 dans les directions du confinement sont égales à la masse transverse m_t , qui est environ cinq fois plus faible que la masse longitudinale m_l . Les niveaux d'énergie longitudinale (suivant \mathbf{x}) associés à ces vallées sont plus élevés et ne participent donc pas à la conduction des électrons⁴. En conséquence seules les vallées Δ_2 et Δ_3 seront considérées dans la suite de l'étude.

5.1.3 Justification de l'approche "mode-space"

Jusqu'à présent, nous avons détaillé la méthode de l'espace des modes sans justifier sa validité. Il est cependant essentiel de connaître les conditions sous lesquelles cette approche offre une modélisation correcte du transport électronique. Dans la partie suivante, nous exprimons l'élément de matrice de H_{3D} entre deux états transverses appartenant à différents plans atomiques et soulignons les approximations faites pour simplifier l'hamiltonien.

⁴Rappel : Dans l'approximation de la masse effective, les niveaux d'énergie d'un puits infini 1D s'écrivent : $\varepsilon_i = (\hbar^2 k_i^2)/(2m^*)$. Une faible masse effective augmente donc l'espace entre niveaux d'énergie.

Définissons $|\psi_{n,i}\rangle$ comme le $i^{\text{ème}}$ état transverse du $n^{\text{ème}}$ plan atomique. L'élément de matrice de l'hamiltonien avec $|\psi_{n+1,i'}\rangle$ peut s'écrire :

$$\langle \psi_{n+1,i'} | H_{3D} | \psi_{n,i} \rangle = \beta_x \langle \psi_{n+1,i'} | T_+ \psi_{n,i} \rangle, \quad (5.24)$$

où T_+ représente la translation d'ensemble de l'état $|\psi_{n,i}\rangle$ du plan n au plan $n+1$. Si le potentiel électrostatique le long de l'axe \mathbf{x} n'est pas constant, c'est-à-dire si (V_n est le potentiel du plan n) :

$$T_+ V_n \neq V_{n+1}, \quad (5.25)$$

l'application de la théorie des perturbations au premier ordre à l'état $|T_+ \psi_{n,i'}\rangle$ donne⁵ :

$$|\psi_{n+1,i'}\rangle = |T_+ \psi_{n,i'}\rangle + \sum_{k \neq i'} \frac{\langle T_+ \psi_{n,k} | V_{n+1} - T_+ V_n | T_+ \psi_{n,i'} \rangle}{\varepsilon_{n,i'} - \varepsilon_{n,k}} |T_+ \psi_{n,k}\rangle. \quad (5.28)$$

Le produit scalaire $\langle \psi_{n+1,i'} | T_+ \psi_{n,i} \rangle$ s'écrit donc :

$$\begin{aligned} \langle \psi_{n+1,i'} | T_+ \psi_{n,i} \rangle &= \langle T_+ \psi_{n,i'} | T_+ \psi_{n,i} \rangle + \sum_{k \neq i'} \frac{\langle T_+ \psi_{n,i'} | V_{n+1} - T_+ V_n | T_+ \psi_{n,k} \rangle}{\varepsilon_{n,i'} - \varepsilon_{n,k}} \langle T_+ \psi_{n,k} | T_+ \psi_{n,i} \rangle, \\ &= \delta_{i'i} + \frac{\langle T_+ \psi_{n,i'} | V_{n+1} - T_+ V_n | T_+ \psi_{n,i} \rangle}{\varepsilon_{n,i'} - \varepsilon_{n,i}} (1 - \delta_{i'i}). \end{aligned} \quad (5.29)$$

Considérant deux états de sous-bandes différentes ($i \neq i'$), le terme de couplage de l'hamiltonien (5.24) devient :

$$\begin{aligned} \langle \psi_{n+1,i'} | H_{3D} | \psi_{n,i} \rangle &= \beta_x \langle \psi_{n+1,i'} | T_+ \psi_{n,i} \rangle, \\ &= \beta_x \times \frac{\langle T_+ \psi_{n,i'} | V_{n+1} - T_+ V_n | T_+ \psi_{n,i} \rangle}{\varepsilon_{n,i'} - \varepsilon_{n,i}}, \end{aligned} \quad (5.30)$$

qui est d'autant plus faible que la variation du potentiel électrostatique suivant l'axe source-drain diminue. Si le potentiel est constant dans cette direction, l'équation (5.30) indique que les sous-bandes engendrées par le confinement latéral sont réellement indépendantes et que la matrice de l'hamiltonien du système peut être représentée en sous-blocs diagonaux. Les sous-bandes sont parfaitement découplées les unes des autres et le problème initial à trois dimensions se décompose en plusieurs problèmes 1D indépendants. Cette hypothèse reste valable tant que les variations du profil du potentiel électrostatique d'une section transverse à la suivante sont négligeables. Par exemple, si V_n conserve la même forme, mais qu'il prend différentes valeurs le

⁵Rappel : la correction au premier ordre du $n^{\text{ème}}$ état non perturbé $|\psi_n^{(0)}\rangle$ sous l'effet d'un potentiel V s'écrit :

$$|\psi_n\rangle = |\psi_n^{(0)}\rangle + |\psi_n^{(1)}\rangle, \quad (5.26)$$

avec :

$$|\psi_n^{(1)}\rangle = \sum_{k \neq n} \frac{\langle \psi_k^{(0)} | V | \psi_n^{(0)} \rangle}{\varepsilon_n^{(0)} - \varepsilon_k^{(0)}} |\psi_k^{(0)}\rangle, \quad (5.27)$$

où les $\varepsilon_k^{(0)}$ représentent les énergies propres associées aux états d'ordre 0 : $|\psi_k^{(0)}\rangle$.

long de l'axe \mathbf{x} , les fonctions propres $|\psi_{n,i}\rangle$ restent inchangées bien que les valeurs propres $\varepsilon_{n,i}$ varient.

Pour les transistors SOI, dans lesquels la partie en silicium est uniformément "sandwichée" entre des grilles isolantes, la variation transverse du potentiel est très faible et l'approximation des sous-bandes indépendantes est justifiée. Des vérifications numériques ont été effectuées avec succès dans le cas d'un transistor double-grille 2D [104]. En revanche, les transistors *bulk* ou les SOI dont les dimensions transverses du canal de conduction ne sont pas constantes ne peuvent être modélisés par ce type d'approche.

5.1.4 Résolution de l'équation de Poisson 3D et auto-cohérence

L'équation de POISSON 3D est résolue en utilisant la méthode des différences finies décrite au § (4.2.1) dans le cas à deux dimensions. Le maillage uniforme est ici étendu à la direction \mathbf{z} avec un pas de discrétisation identique à celui de la structure cubique simple adoptée ici pour simuler le silicium ($a = 0.2 \text{ nm}$). Utilisant un développement de TAYLOR au second ordre, les dérivées secondes du potentiel électrostatique $V(x, y, z)$ s'expriment dans le formalisme des différences finies sous la forme :

$$\begin{aligned} \frac{\partial^2 V(x, y, z)}{\partial x^2} &\approx \frac{V(x + \Delta x, y, z) + V(x - \Delta x, y, z) - 2V(x, y, z)}{\Delta x^2}, \\ \frac{\partial^2 V(x, y, z)}{\partial y^2} &\approx \frac{V(x, y + \Delta y, z) + V(x, y - \Delta y, z) - 2V(x, y, z)}{\Delta y^2}, \\ \frac{\partial^2 V(x, y, z)}{\partial z^2} &\approx \frac{V(x, y, z + \Delta z) + V(x, y, z - \Delta z) - 2V(x, y, z)}{\Delta z^2}. \end{aligned} \quad (5.31)$$

Considérant un maillage uniforme de pas égal à a , l'équation de POISSON pour un atome interne au matériau s'écrit :

$$\begin{aligned} &\frac{V_{l+1,m,n} + V_{l-1,m,n} - 2V_{l,m,n}}{a^2} + \frac{V_{l,m+1,n} + V_{l,m-1,n} - 2V_{l,m,n}}{a^2} + \frac{V_{l,m,n+1} + V_{l,m,n-1} - 2V_{l,m,n}}{a^2} \\ &= \frac{\rho_{l,m,n}}{\epsilon}, \end{aligned} \quad (5.32)$$

où $V_{l,m,n}$ est le potentiel électrostatique discrétisé, ρ est densité de charges et l , m et n représentent les $l^{\text{ièmes}}$, $m^{\text{ièmes}}$ et $n^{\text{ièmes}}$ atomes suivant les directions \mathbf{z} , \mathbf{y} et \mathbf{x} respectivement.

Aux interfaces oxyde/silicium, la continuité la composante perpendiculaire de $\epsilon \vec{E}$, impose les relations suivantes :

$$(4V_{l,m,n-1} - 3V_{l,m,n} - V_{l,m,n-2}) \epsilon_{\text{ox}} = (4V_{l,m,n+1} - 3V_{l,m,n} - V_{l,m,n+2}) \epsilon_{\text{Si}}, \quad (5.33)$$

si l'interface est caractérisée par le $n^{\text{ième}}$ plan atomique perpendiculaire à la direction \mathbf{x} ,

$$(4V_{l,m-1,n} - 3V_{l,m,n} - V_{l,m-2,n}) \epsilon_{\text{ox}} = (4V_{l,m+1,n} - 3V_{l,m,n} - V_{l,m+2,n}) \epsilon_{\text{Si}}, \quad (5.34)$$

si l'interface est caractérisée par le $m^{\text{ième}}$ plan atomique perpendiculaire à la direction \mathbf{y} ,

$$(4V_{l-1,m,n} - 3V_{l,m,n} - V_{l-2,m,n}) \epsilon_{\text{Si}} = (4V_{l+1,m,n} - 3V_{l,m,n} - V_{l+2,m,n}) \epsilon_{\text{Si}}, \quad (5.35)$$

si l'interface est caractérisée par le $l^{\text{ième}}$ plan atomique perpendiculaire à la direction \mathbf{z} .

Au niveau des contacts des grilles les conditions aux limites de DIRICHLET restent valables :

$$V_{l,m,n} = V_G - V_{FB}, \quad (5.36)$$

où V_{FB} est la tension de bande plate traduisant la différence du travail de sortie entre le métal de grille et le silicium.

Dans les contacts de source et drain, les conditions aux limites de NEUMANN remplacent celles de DIRICHLET adoptées dans le système 2D du chapitre 4 et imposent désormais un champ électrique nul :

$$\frac{\partial V(x, y, z)}{\partial x, y, z} = 0. \quad (5.37)$$

En régime passant, la majorité des électrons issus de la source est transmise à travers le canal, alors que le comportement opposé est valable pour les électrons injectés depuis le drain. La densité électronique de l'extension de source diminue et celle du drain augmente par rapport aux valeurs d'équilibre. Si le potentiel aux extrémités de la zone active est fixe (conditions de DIRICHLET), l'électroneutralité de ces régions (\equiv concentration des dopants donneurs égale celle de la densité électronique) qui résulte de l'auto-cohérence SCHRÖDINGER-POISSON ne peut être conservée. En conséquence, le potentiel électrostatique des extrémités de la zone active doit "flotter" par rapport aux niveaux de FERMI des grands réservoirs à l'équilibre thermodynamique. Ces conditions sont ici applicables grâce au caractère 3D de la zone active. Dans le MOSFET double-grille ultime décrit au chapitre 4, la zone active se réduisait à un fil de silicium idéalement unidimensionnel. Considérer un potentiel flottant aux extensions de source et de drain, aurait conduit à une chute de tension quasi-nulle dans la zone active, phénomène peu probable sachant que les résistances de contact de systèmes réels ne sont jamais totalement négligeables et qu'elles permettent de maintenir une différence de potentiel entre les extrémités de la zone active. Dans un système balistique 1D dépourvu de résistance de contact, seuls les réservoirs limitent la conduction et encaissent la totalité de la tension appliquée (cf § (3.4.2)).

Toutes les frontières qui ne sont pas au contact d'électrodes de grilles remplissent donc les conditions de champ électrique nul, tant au niveau des atomes de silicium des extrémités du nanofil que pour les atomes de l'oxyde des contours du nanofil. La condition de dérivée nulle se

traduit en différences finies par les égalités ci-dessous :

$$\begin{aligned}
 V_{l,m,n} &= \frac{4V_{l,m,n+1} - V_{l,m,n+2}}{3}, \\
 V_{l,m,n} &= \frac{4V_{l,m+1,n} - V_{l,m+2,n}}{3}, \\
 V_{l,m,n} &= \frac{4V_{l+1,m,n} - V_{l+2,m,n}}{3},
 \end{aligned} \tag{5.38}$$

suivant que la surface considérée est perpendiculaire à la direction x , y ou z respectivement. Les équations (5.32)-(5.38) permettent de calculer de proche en proche les valeurs discrètes du potentiel électrostatique $V_{l,m,n}$.

L'équation de POISSON est ensuite couplée à celle de SCHRÖDINGER afin de converger vers la solution auto-cohérente. L'approche de l'espace des modes, qui traite séparément le confinement quantique et le transport, se décompose en plusieurs étapes (figure (5.5)) :

- Etape 1 : Résoudre l'équation de POISSON afin d'obtenir un potentiel électrostatique d'essai.
- Etape 2 : Résoudre l'équation de SCHRÖDINGER 2D des plans atomiques transverses et calculer les premières sous-bandes électroniques de chaque vallée le long du nano-fil ainsi que leur fonction d'onde associée.
- Etape 3 : Résoudre l'équation de transport 1D dans le formalisme des fonctions de GREEN hors-équilibre pour en retirer les densités électroniques 1D.
- Etape 4 : Calculer les densités de charges 3D en multipliant les densités électroniques 1D obtenues en 3 par le module au carré des modes transverses trouvés en 2.
- Etape 5 : Revenir à l'étape 1 et calculer le nouveau potentiel électrostatique. Si la convergence est atteinte, le courant est calculé dans le formalisme des fonctions de GREEN ; sinon on reprend les étapes 2 et 3.

5.2 Application au MOSFET Gate-All-Around

Nous allons dans la suite appliquer le formalisme précédemment décrit au MOSFET *Gate-all-around* représenté sur la figure (5.6). Le transistor est caractérisé par un nano-fil de silicium de section carrée ($W_{\text{Si}} = T_{\text{Si}} = 4 \text{ nm}$) entouré d'un oxyde de silicium d'épaisseur constante égale à 1nm.

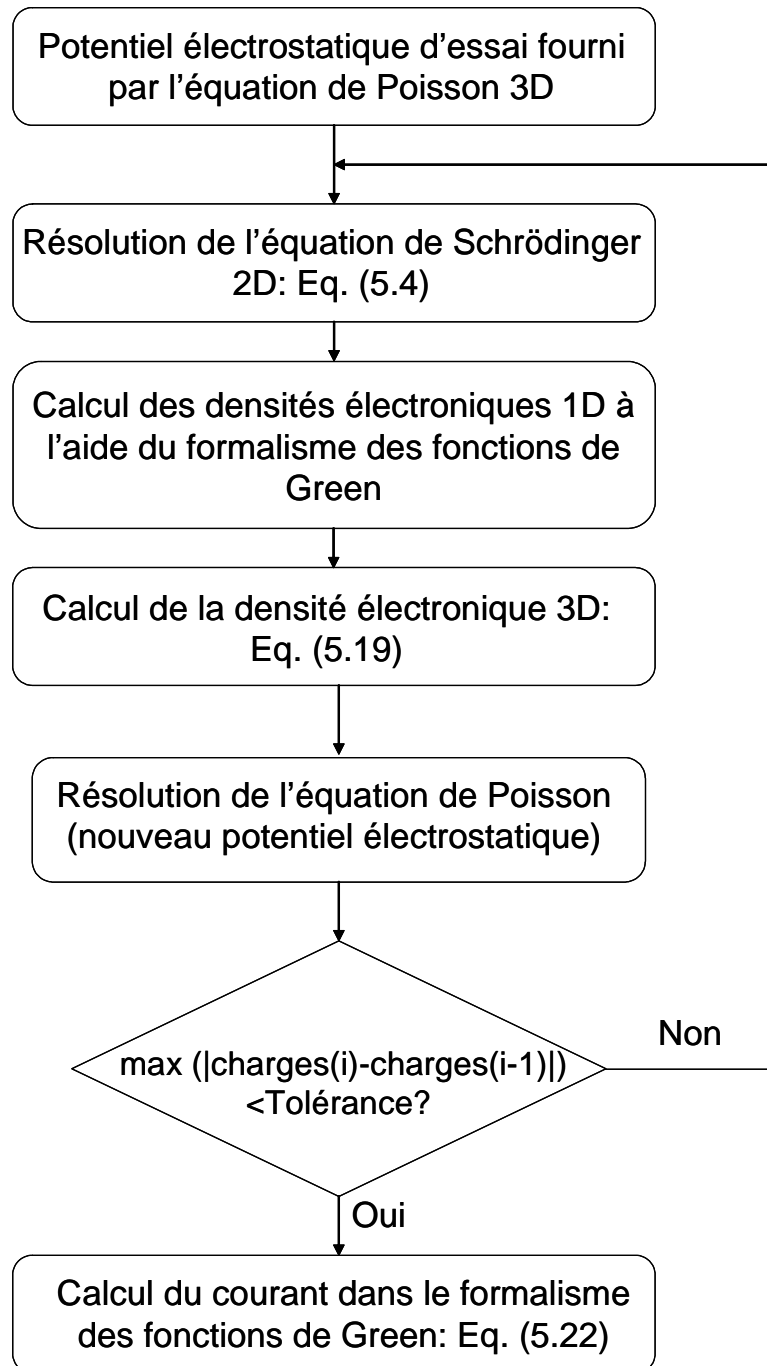


FIG. 5.5: Graphique détaillant les différentes étapes de l'auto-cohérence dans l'approche de l'espace des modes 3D.

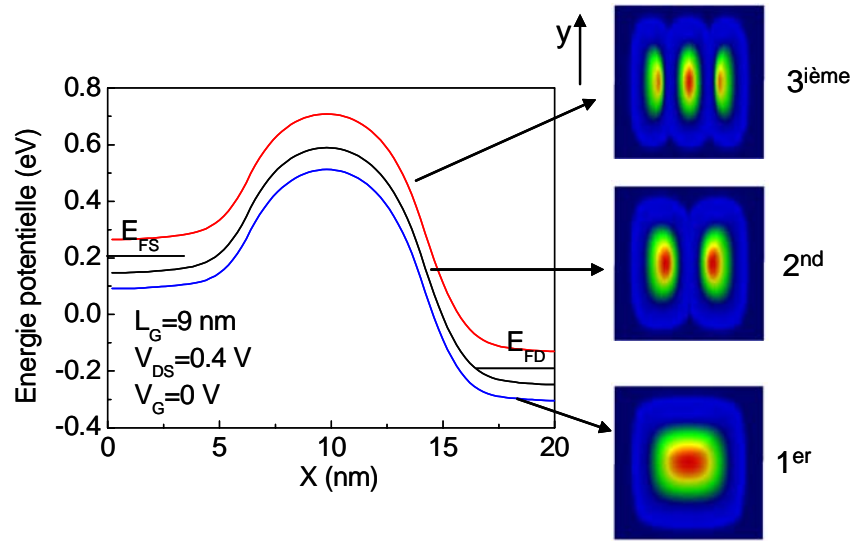


FIG. 5.7: Représentation des trois premières sous-bandes d'énergie associées aux vallées (010) et des modules au carré de leurs états propres transverses. E_{FS} et E_{FD} définissent les niveaux de FERMIE de la source et du drain respectivement.

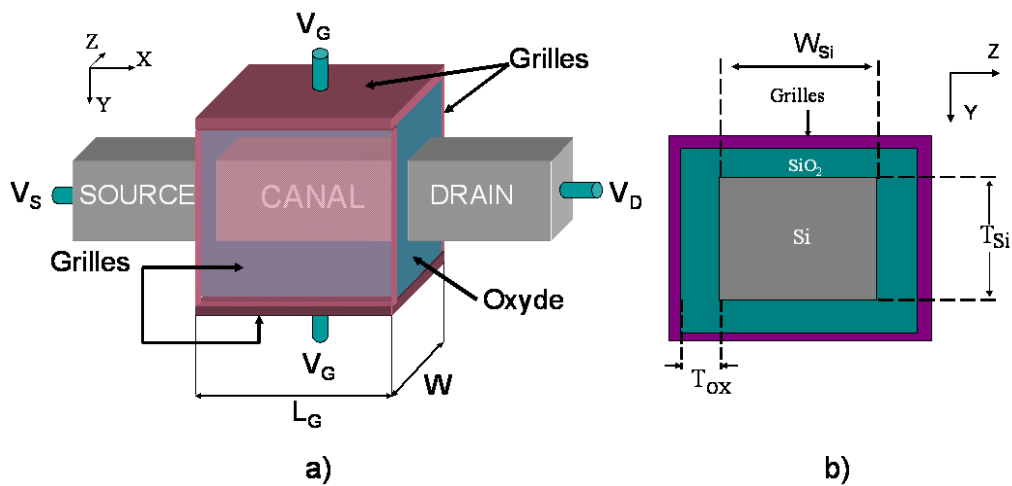


FIG. 5.6: a) Représentation 3D d'un MOSFET Gate-all-around. Par volonté de clarté, l'oxyde de grille n'est pas représenté autour de la source et du drain. b) Définition des principaux paramètres de la section transversale.

La figure (5.7) illustre la séparation de l'équation de SCHRÖDINGER 3D en une équation 2D définissant le confinement suivant les directions y et z et une équation 1D décrivant le transport balistique le long du canal à l'aide du formalisme des fonctions de GREEN. Le confinement transverse induit une discrétisation des énergies qui se répartissent en sous-bandes. L'approximation de l'espace des modes, qui consiste à considérer les sous-bandes indépendantes, permet de ne retenir que celles peuplées d'électrons, c'est-à-dire celles qui se situent sous les niveaux de FERMIE des réservoirs macroscopiques. Afin de légitimer l'application de cette méthode aux transistors confinés, il a été vérifié que les modes restaient rigoureusement identiques d'un plan atomique

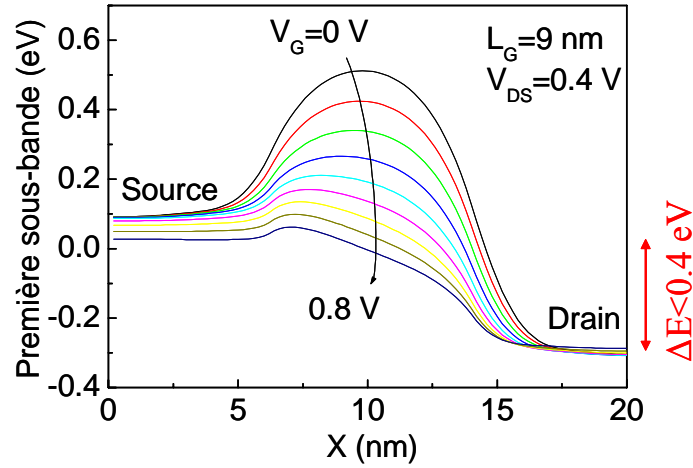


FIG. 5.8: Evolution de la première sous-bande pour V_G variant de 0 à 0.8 V par pas de 0.1 V. La tension appliquée aux extrémités de la zone active est inférieure à V_{DS} (0.4 V), illustrant ainsi la notion de résistance des réservoirs.

à un autre⁶. D'autre part le calcul des sous-bandes et des vecteurs propres de chaque section transverse constitue plus de 70% du temps de calcul cpu total. L'approche "mode-space" peut donc être numériquement optimisée en ne calculant que les fonctions propres de la première section atomique et en les répétant dans les autres plans de la zone active. Les valeurs propres doivent en revanche être déterminées dans tout le système, mais leur obtention est beaucoup plus rapide.

La figure (5.8) représente la première sous-bande des vallées (010) le long de la zone active pour différentes tensions de grille. A faible V_G , la barrière de potentiel du canal interdit le passage des électrons depuis la source vers le drain : le transistor est "bloqué". La résistance du canal est dans ce cas suffisante pour que la totalité de la tension de drain V_{DS} appliquée aux électrodes se retrouve au niveau de la zone active. Lorsque V_G augmente, la barrière de potentiel du canal diminue et les électrons traversent la zone active sans subir de réflexion. Seule la tension drain-source décide du sens du transport. Il en résulte un état hors-équilibre prononcé caractérisé par un déficit de charges dans l'extension de la source et un excédent dans l'extension du drain (figure (5.9)). Plus précisément pour une tension $V_{DS} > 0$, un électron injecté dans l'extension de la source sera transmis à travers le canal, alors que son homologue du drain sera réfléchi. Le nombre d'électrons par niveau d'énergie passe au final de la valeur d'équilibre 2 à 1 dans l'extension de source et à 3 dans celle drain. Le rôle des réservoirs est de s'opposer à cet état hors-équilibre et de maintenir l'électroneutralité en injectant (resp. absorbant) plus d'électrons dans l'extension de la source (resp. du drain). Cependant la quantité d'électrons fournie ou absorbée par un réservoir à travers un seul mode (\equiv sous-bande) est limitée par le quantum de conductance $G_0 = 2e^2/h$ (cf § (3.3.1)). La distribution de FERMI-DIRAC, qui régit la répartition des électrons dans les réservoirs, est à l'origine de cette limitation en imposant un maximum de deux électrons (spin) par niveau d'énergie. Les réservoirs présentent ainsi une

⁶La variation du produit scalaire de vecteurs propres de la sous-bande i appartenant à différentes sections est inférieure au millième ($\langle T^+ \psi_{n,i} | \psi_{n',i} \rangle \approx 0.999$).

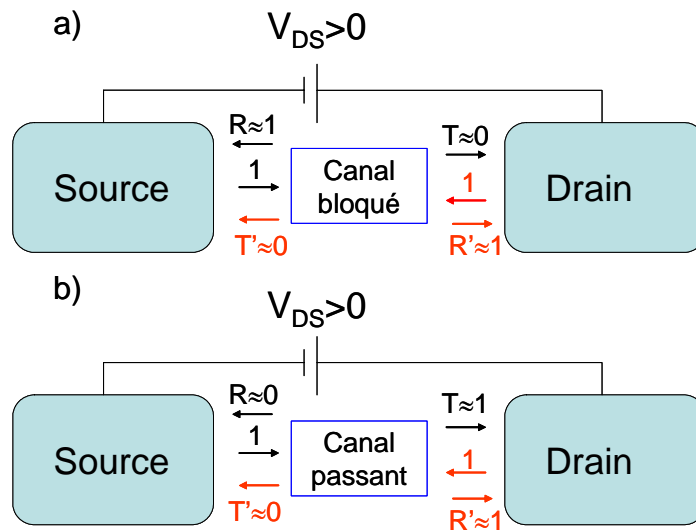


FIG. 5.9: Représentation du transport en termes de probabilités de transmission (T, T') et de réflexion (R, R') d'un état normé ($=1$) dans un : a) transistor passant, b) transistor bloqué.

résistance finie et "encaissent" une partie de la tension V_{DS} appliquée aux électrodes (figure (5.8) pour $V_G = 0.8$ V). L'abaissement des sous-bandes côté source augmente le nombre d'électrons de cette région alors que le décalage positif en énergie des sous-bandes côté drain produit l'effet inverse.

La figure (5.10) montre le coefficient de transmission à travers la barrière source-canal-drain associé aux vallées (010). Malgré un système électrostatique 3D complexe, le coefficient de transmission présente une allure simple en escaliers illustrant la contribution des trois premières sous-bandes d'énergie au processus de transport des porteurs. Concrètement, les énergies pour lesquelles la transmission augmente d'une unité indiquent le début d'une sous-bande et la vitesse de croissance dépend de la forme de la barrière de potentiel entre la source et le drain qui détermine la proportion de la composante tunnel et thermo-ionique. Chaque plateau correspond à un quantum de conductance $G_0 = 2e^2/h$. Nous avons ici un exemple d'évolution des propriétés électriques pour les dimensions mésoscopiques. La conductivité du canal n'est plus continue et change brutalement d'une énergie à une autre. La relation de linéarité entre courant et tension (loi d'OHM) n'est plus vérifiée ou tout du moins ne l'est que sur une plage restreinte de tension où le nombre de sous-bandes participant au transport est constant.

La figure (5.11) représente la densité électronique le long de l'axe source-drain, prise au centre de la section transverse, pour les deux régimes de fonctionnement du transistor. En régime bloqué, la diminution de la quantité de charges du canal suit une loi exponentielle, caractéristique d'une transmission par effet tunnel. Les électrons issus de la source sont clairement dissociés de ceux provenant du drain. A l'état passant, la densité électronique des charges augmente dans le canal : les électrons de la source traversent désormais la zone active. La séparation spatiale des particules de chaque réservoir est moins apparente. Nous observons également des oscillations côté drain, associées aux réflexions au niveau du contact. Les états injectés depuis le drain se

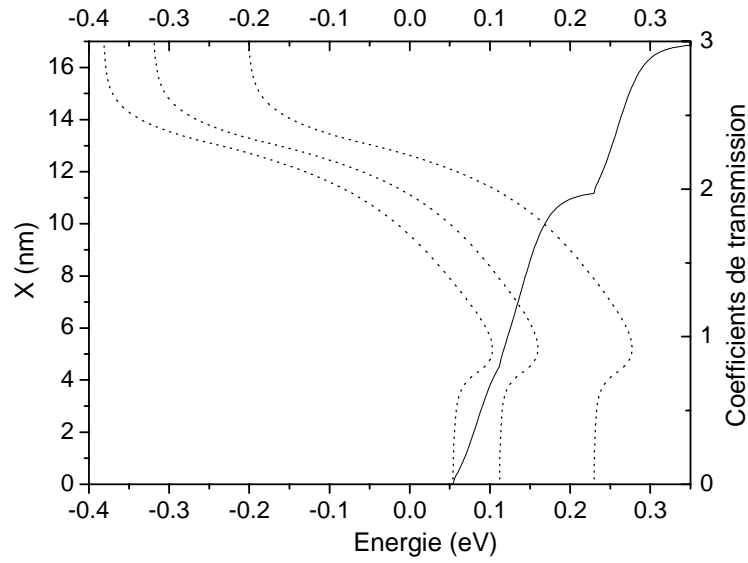


FIG. 5.10: Coefficient de transmission à travers les vallées (010) le long de l'axe source-canal-drain et les sous-bandes associées. $V_{DS} = 0.5$ V, $V_G = 0$ V et $L_G = 9$ nm.

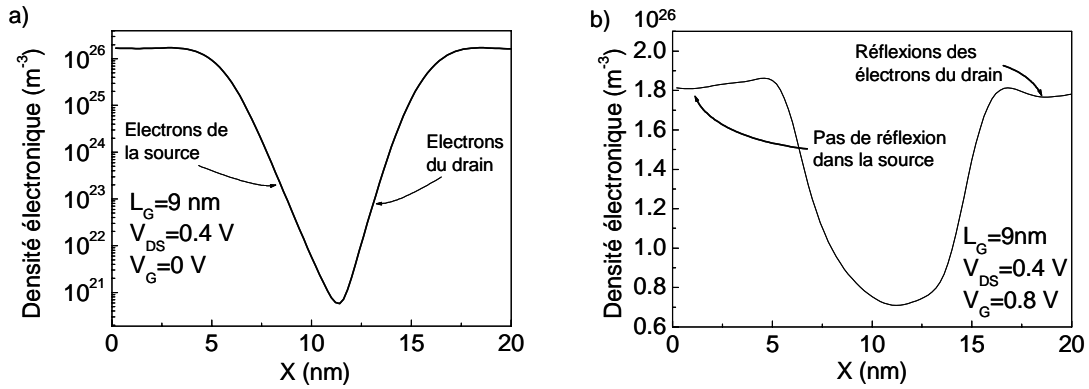


FIG. 5.11: Profil de la densité électronique le long de la zone active du MOSFET GAA simulé : a) en régime bloqué, b) en régime passant. $\theta = 300$ K.

réfléchissent sur la barrière du canal et interfèrent avec les nouveaux états entrants dans la zone active. Côté source, la majorité des électrons traverse le canal et les interférences avec les électrons réfléchis sont plus faibles. Notons également, que la présence d'oscillations repose sur la conservation de la cohérence de phase dans les composants balistiques. En présence de diffusion les interférences sont détruites [202]. Ces oscillations sont plus visibles si l'on étudie la densité d'états locale (équation (5.18) sans les fonctions de distribution) le long de la zone active pour chaque régime (figure (5.12)). Dans l'état bloqué, les extensions de source et de drain présentent des franges d'interférence traduisant les réflexions des états incidents sur la barrière de potentiel du canal et leur avec les nouveaux états entrant dans la zone active. En régime passant, seule l'extension de drain conserve les oscillations. On observe également une densité non nulle sous les sous-bandes ce qui est dû au tunnel source-drain.

Les caractéristiques $I_D - V_G$ calculées pour deux longueurs de grille différentes (figure (5.13))

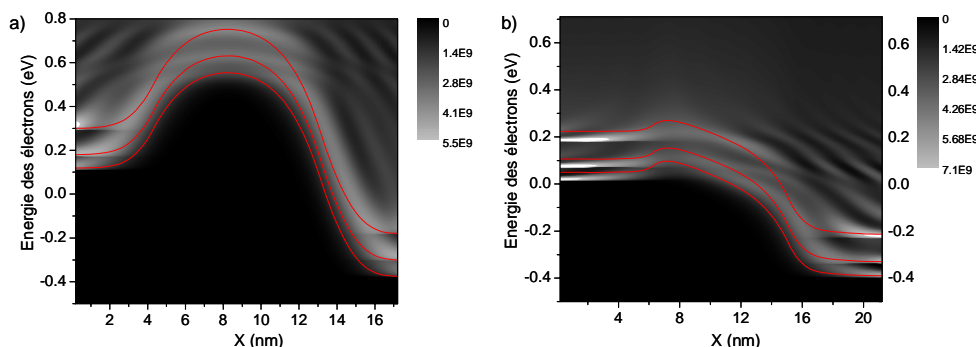


FIG. 5.12: Densités locales d'états et sous-bandes (lignes continues) de la vallées (010) d'un MOSFET GAA a) en régime bloqué ($V_G = 0$ V), b) en régime passant ($V_G = 0.7$ V). $L_G = 9$ nm, $V_{DS} = 0.5$ V.

montrent que l'effet tunnel entre la source et le drain s'amplifie avec la diminution de L_G sous les 10 nm. A ces dimensions, l'épaisseur de la barrière du canal décroît : la composante tunnel à faible V_G dégrade la pente sous le seuil ainsi que le courant I_{off} , alors que le courant thermo-ionique présente des performances sous le seuil bien meilleures. Le transistor devient "tunnel" [204], mais sa tension de grille contrôle toujours le courant du canal.

5.3 Comparaison des différentes architectures émergentes

Les MOSFETs multi-grille, tels que le *Tri-gate*, le *Pi-gate*, le *Omega-gate* et le *Gate-all-around* (figure (5.14a)) sont pressentis comme de sérieux candidats pour continuer la diminution des composants jusqu'à l'échelle nanométrique [32]. Une limitation importante à ce niveau d'intégration est la prédominance des effets quantiques, principalement le confinement du canal (qui induit une séparation des sous-bandes) et la conduction par effet tunnel à travers la barrière de potentiel du canal. Il a été théoriquement prouvé [201] que les transistors à nano-fil de silicium avec deux/trois grilles et un bord flottant (pas de grille arrière) peuvent présenter un meilleur contrôle électrostatique que les MOSFETs double-grille. L'objectif de cette section est précisément de simuler à l'aide du formalisme des fonctions de GREEN des transistors à nano-fils présentant différentes architectures de grille. Nous comparons les performances électriques en fonction des configurations de grille (figure (5.14b)) et explorons la dépendance des performances face aux variations d'épaisseur, de largeur et de longueur du fil de silicium afin d'évaluer les capacités et les limites de chaque architecture. La figure (5.15) illustre l'ensemble des paramètres utilisés pour définir les formes des multi-grilles. Nous considérons toujours un oxyde de silicium d'épaisseur $T_{ox} = 1$ nm. La section transverse de la partie en silicium reste également inchangée ($T_{Si} = W_{Si} = 4$ nm). Seule l'influence de la structure des grilles est dans un premier temps étudiée.

La figure (5.16) compare les caractéristiques $I_D - V_G$ de chaque architecture. Nous constatons que les paramètres sous le seuil ainsi que le courant I_{on} s'améliorent nettement avec l'augmen-

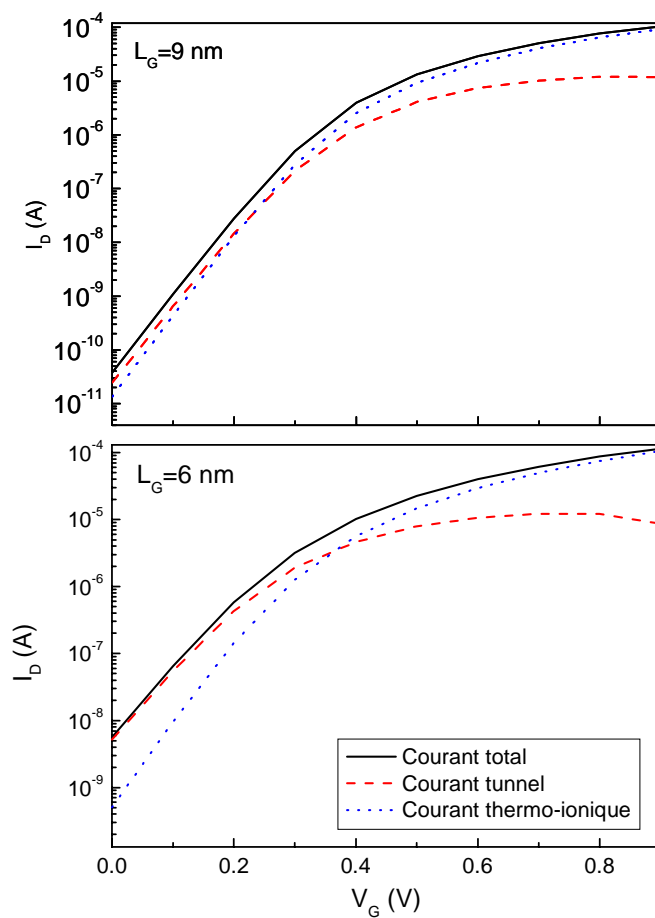


FIG. 5.13: Caractéristiques $I_D - V_G$ calculées pour un MOSFET GAA avec $L_G = 9$ nm et $L_G = 6$ nm. $T_{Si} = W_{Si} = 4$ nm et $V_{DS} = 0.5$ V. La composante tunnel dégrade substantiellement la pente sous le seuil du transistor court.

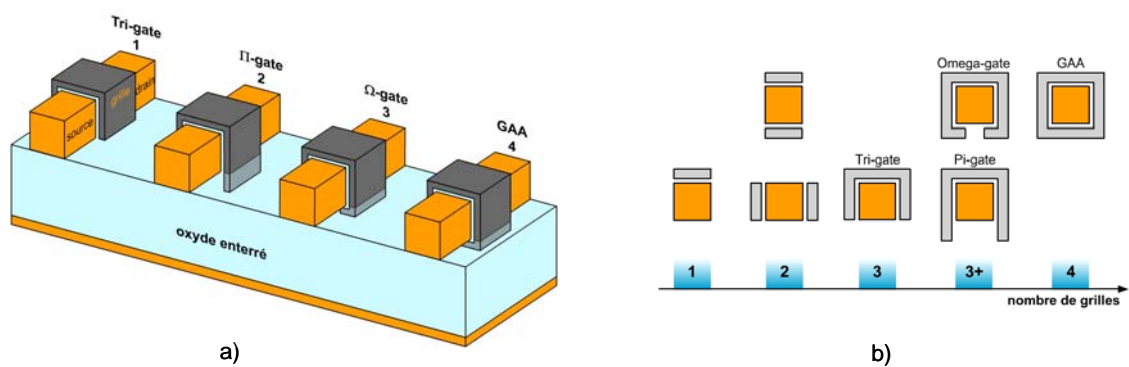


FIG. 5.14: a) Différentes configurations de MOSFETs multi-grilles. b) Sections transverses schématiques classées en fonction du nombre équivalent de grilles.

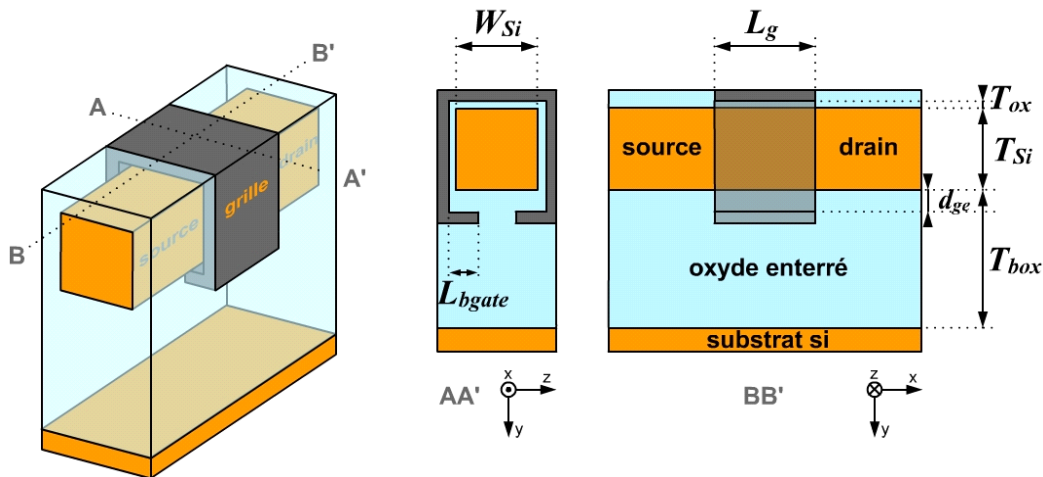


FIG. 5.15: Définition des principaux paramètres associés aux architectures présentées figure (5.14). Pour toutes les structures simulées, la concentration de dopants S/D est 10^{20} cm^{-3} et le canal est intrinsèque. Il n'y a pas de recouvrement de grille avec les régions S/D et les jonctions sont supposées abruptes.

tation du nombre de grilles équivalentes, c'est-à-dire avec le renforcement du contrôle électrostatique des grilles sur la conduction du canal. Cet effet est illustré par la figure (5.17) où la densité électronique dans le plan $y-z$ du milieu du canal est représentée. Sans tenir compte de l'amplitude des densités elle-même, on observe une inversion volumique plus prononcée au centre de la section transverse dans le cas du transistor *Gate-all-around*. L'influence de chaque grille sur les charges apparaît nettement depuis la structure *Tri-gate* jusqu'au MOSFET *Gate-all-around* en passant par le Π -*gate* et le Ω -*gate* et les densités électroniques gagnent en homogénéité avec l'augmentation du nombre équivalent de grilles. Les caractéristiques $I_D - V_G$ pour des transistors plus courts ($L_G = 6 \text{ nm}$) sont représentées figure (5.18). Nous observons que la dégradation en régime sous le seuil est plus importante que dans le cas précédent (*i.e.* $L_G = 9 \text{ nm}$) lorsque le nombre équivalent de grilles est réduit. La figure (5.18) montre également que le GAA contrôle le mieux la barrière du canal. Les coefficients de transmission à travers la barrière source-drain associés aux vallées (010) sont représentés figure (5.19). En régime passant, les courbes en escalier traduisent la contribution des trois premières sous-bandes d'énergie. A énergie donnée l'augmentation du coefficient de transmission avec le nombre équivalent de grilles constitue clairement un indice de qualité du contrôle électrostatique. Le comportement est inversé en régime bloqué, *i.e.* que le GAA fournit la plus petite transparence est le *Tri-gate* la plus élevée. L'influence des extensions de grille pour des structures Π -*gate* et Ω -*gate* est reportée sur la figure (5.20). Cette série de résultats montre que les performances globales de l'architecture *Tri-gate* sont faiblement améliorées par la pénétration verticale de la grille dans l'oxyde enterré. Il en est autrement pour le Ω -*gate* dans lequel un mince recouvrement de la grille arrière se traduit par une forte modification des caractéristiques électriques; ces dernières se rapprochant de celles du MOSFET GAA. L'impact de l'extension de grille est plus visible en observant la pente sous le seuil (S) et le courant I_{on} de chaque architecture séparément (figure (5.21)). Alors qu'une pénétration verticale de 3 nm de l'oxyde d'un Π -*gate* n'induit qu'une amélioration de 7.5% de S et de 15% du

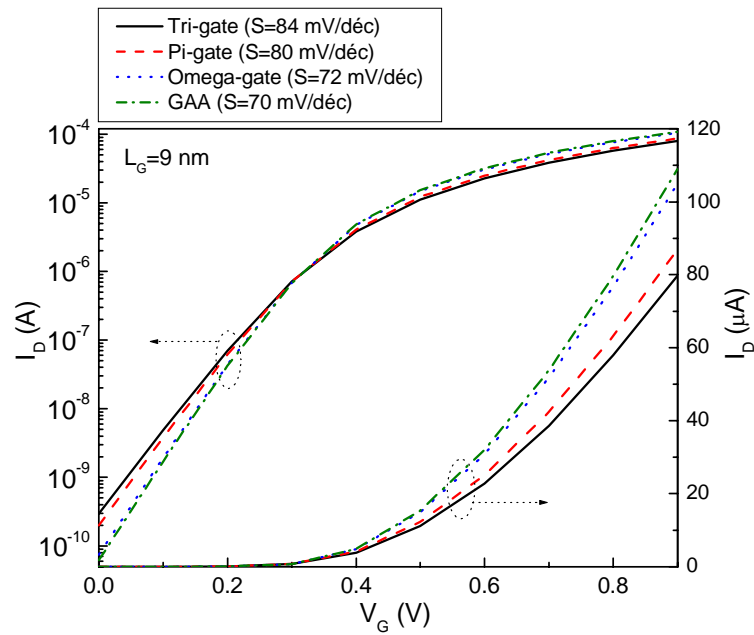


FIG. 5.16: Caractéristiques $I_D - V_G$ des différentes architectures. Augmenter le nombre de grilles améliore de manière générale les performances du transistor. Les transistors Π -gate et Ω -gate présentent respectivement une extension de grille égale à $d_{ge} = 0.8 \text{ nm}$ et $L_{bg} = 0.4 \text{ nm}$. $T_{Si} = W_{Si} = 4 \text{ nm}$.

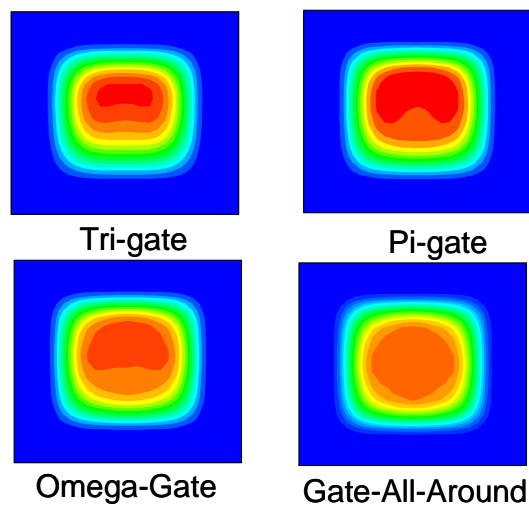


FIG. 5.17: Evolution des densités électroniques dans la section transverse du milieu du canal pour différentes structures. $V_{DS} = 0.5 \text{ V}$, $V_G = 0.7 \text{ V}$. Les autres paramètres sont identiques à ceux de la figure (5.16).

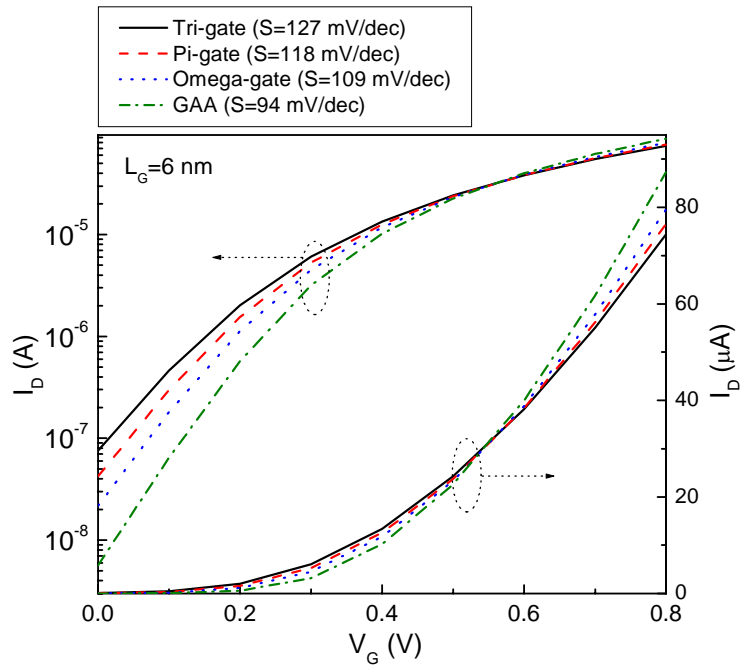


FIG. 5.18: Caractéristiques $I_D - V_G$ des différentes architectures pour une longueur de grille $L_G = 6$ nm. Les autres paramètres sont ceux de la figure (5.16).

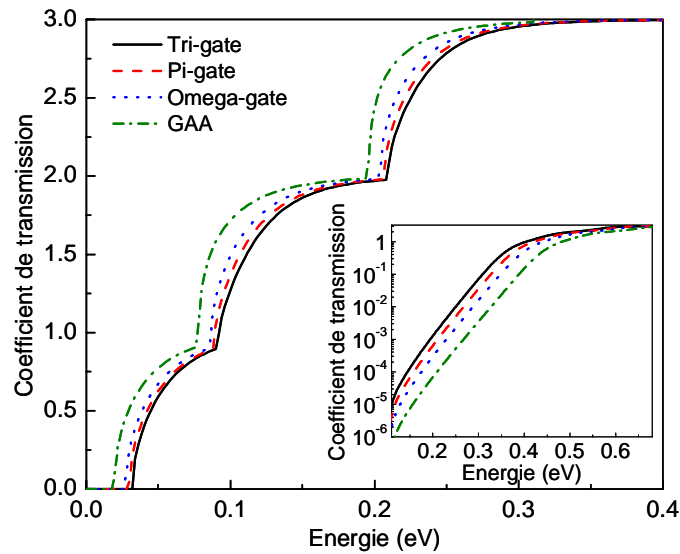


FIG. 5.19: Coefficients de transmission des vallées (010) pour les différentes architectures avec $L_G = 6$ nm en régime passant ($V_G = 0.7$ V). Figure insérée : coefficient de transmission pour les mêmes architectures en régime bloqué ($V_G = 0$ V). Les autres paramètres sont ceux de la figure (5.16)

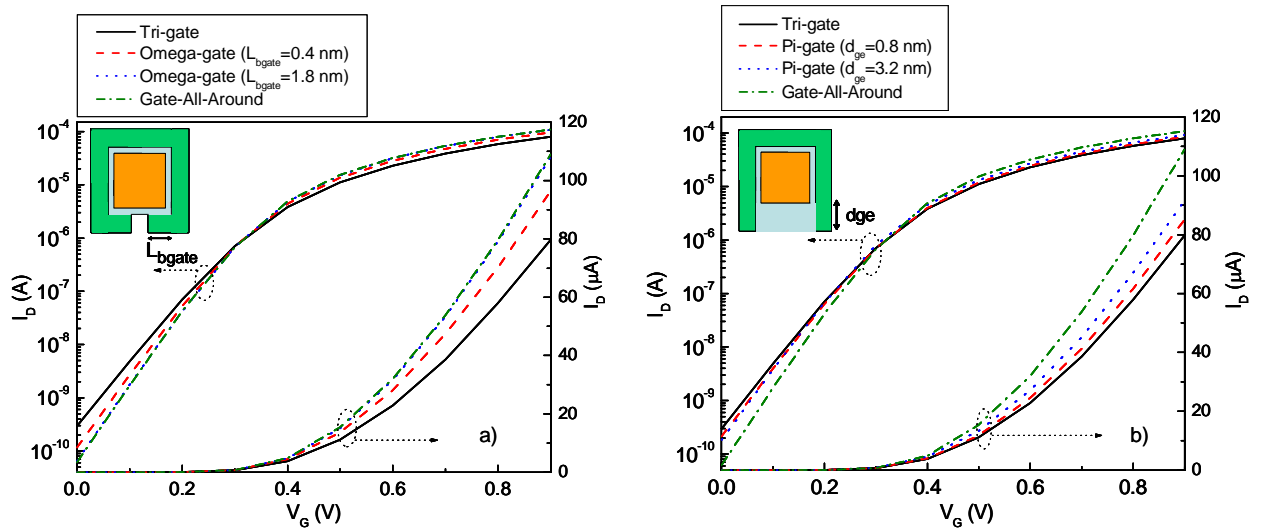


FIG. 5.20: Evolution des caractéristiques de courant $I_D - V_G$ en fonction a) de la pénétration de l'oxyde vertical des Π -gate b) du recouvrement de l'oxyde arrière des Ω -gate. $L_G = 9$ nm, $T_{Si} = W_{Si} = 4$ nm, $V_{DS} = 0.5$ V.

I_{on} , un recouvrement de la grille arrière d'un Ω -gate de 0.4 nm améliore les mêmes paramètres de 10% et 21% respectivement.

Finalement, la figure (5.22) montre l'évolution de la pente sous le seuil en fonction de la section transverse d'un fil de silicium carré pour le GAA et le *Tri-gate*. L'augmentation de la section transverse induit une divergence rapide de S vers des valeurs inacceptables (>110 mV/décade) pour le *Tri-gate* alors que la pente du GAA est moins modifiée. D'autre part, nous observons pour de faibles sections transverses, que l'influence de la configuration de grille s'atténue et que les pentes S des deux structures sont sensiblement les mêmes. Ce comportement est expliqué par le fait qu'en-dessous 2 nm, la combinaison du contrôle électrostatique des grilles et des effets du confinement quantique conduit à des distributions de densité de porteurs similaires dans le film pour toutes les structures de grille. Nous illustrons cet effet figure (5.23) dans le cas le plus défavorable du *Tri-gate*. La structure présente une distribution électronique 2D centrée identique à celle du GAA lorsque la section transverse décroît sous 2 nm. À ce niveau ultime d'intégration, on peut supposer que toutes les architectures multi-grilles deviennent sensiblement équivalentes.

5.4 Conclusion

Dans ce chapitre, nous avons présenté une simulation quantique 3D de transistors à nano-fil de silicium basée sur le formalisme des fonctions de GREEN hors-équilibre. L'approche de l'espace des modes est adoptée pour décomposer l'hamiltonien 3D et ainsi obtenir un gain significatif du temps de calcul. Ceci nous a permis de nous intéresser à des architectures MOSFETs émergentes. Nous avons notamment comparé les performances électriques de différents compo-

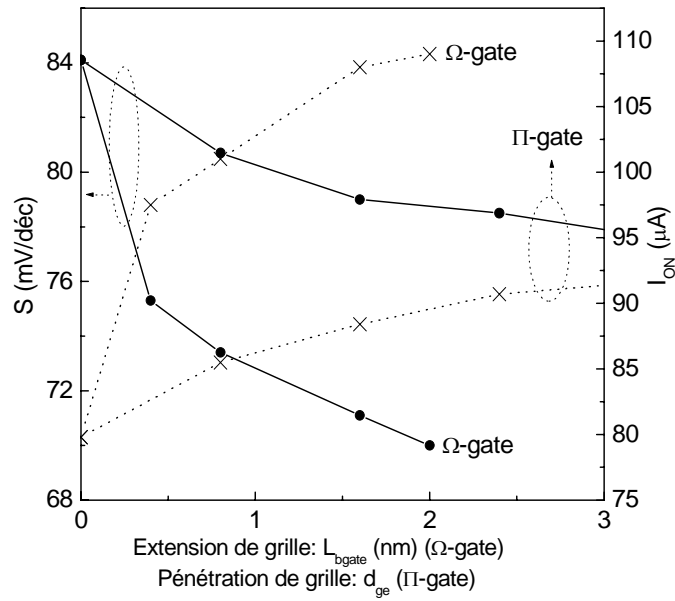


FIG. 5.21: Influence des extensions de grille dans les structures Π -gate et Ω -gate sur la pente sous le seuil S (continu) et le courant I_{on} (pointillés). $L_G = 9$ nm, $T_{Si} = W_{Si} = 4$ nm, $V_{DS} = 0.5$ V.

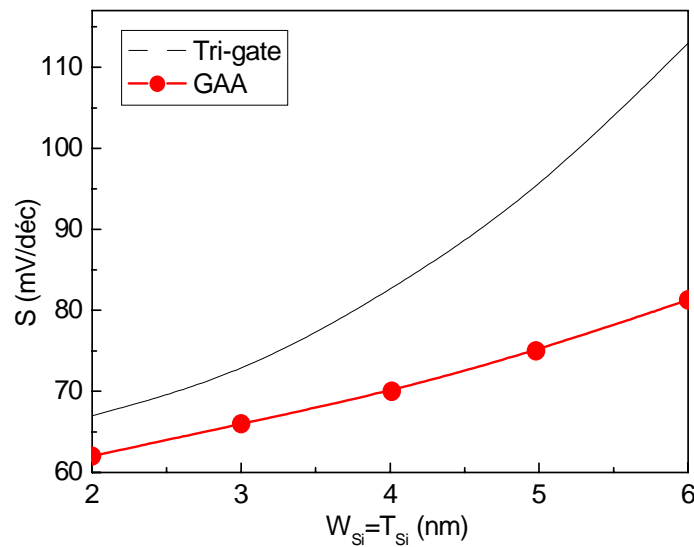


FIG. 5.22: Pente sous le seuil en fonction de la section transversale carrée pour les architectures Tri-gate et GAA. Les autres paramètres sont : $L_G = 9$ nm, $V_{DS} = 0.5$ V.

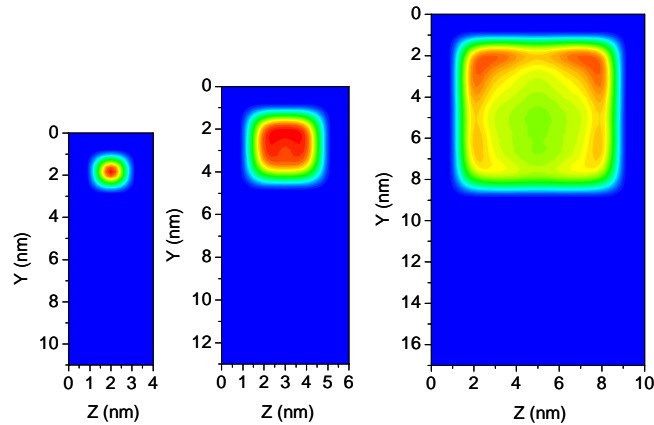


FIG. 5.23: *Distribution de la densité électronique 2D dans le plan y - z d'une structure Tri-gate avec : 1) $T_{Si} = W_{Si} = 2$ nm, 2) $T_{Si} = W_{Si} = 4$ nm, 3) $T_{Si} = W_{Si} = 8$ nm.*

sants multi-grilles en fonction de leur architecture. Le transistor gate-all-around semble présenter les meilleures performances grâce à son excellent contrôle électrostatique des charges du canal et reste le favori pour parvenir à des longueurs de grille sous les 10 nm. La structure Ω -gate qui propose des performances très proches de ce dernier sans nécessiter de grille arrière (plus difficilement réalisable) pourrait bien constituer une autre solution. Enfin, nous avons montré que les différences électriques entre architectures s'estompaient avec la diminution de la section transverse.

Les informations obtenues jusqu'à présent concernent des composants parfaits, c'est-à-dire sans défaut cristallin. Ce cas de figure est en réalité peu probable et la prévision de l'influence de défauts ponctuels dans de tels systèmes soulève des interrogations majeures. Le chapitre suivant tentera d'y répondre en modélisant des défauts ponctuels dans les transistors 3D, toujours à l'aide du formalisme des fonctions de GREEN.

Chapitre 6

Traitement et influence de défauts ponctuels dans les nano-transistors

La diminution de la taille des MOSFETs vers les limites physiques et technologiques pose le choix des structures des prochaines générations CMOS en des termes nouveaux. Parmi les difficultés envisagées, les fluctuations aléatoires du nombre et de la position des impuretés dopantes des composants peuvent détériorer la fiabilité des transistors (variation de la tension de seuil et du courant I_{off}) et les rendre impropres aux opérations CMOS [205, 206]. En effet, lorsque les MOSFETs atteignent l'échelle nanométrique, la variation des paramètres du transistor due au faible nombre de dopants et à leur position dans le canal devient préoccupante. Du fait de la réduction de la tension d'alimentation des transistors ultimes, les nouveaux circuits intégrés sont de surcroît plus sensibles aux fluctuations des caractéristiques des MOSFETs et les positions aléatoires des dopants se répercutent désormais sur les performances et la fonctionnalité des circuits analogiques et numériques.

Dans ce chapitre, nous utilisons une description quantique basée sur le formalisme des fonctions de GREEN hors-équilibre pour étudier l'influence de défauts ponctuels sur les paramètres électriques de structures MOSFETs émergentes. La première section présente la solution analytique d'un défaut ponctuel représenté par un potentiel intra-atomique, introduit dans une chaîne linéaire à l'équilibre. La section (6.2) applique cette approche au MOSFET double-grille dont le canal, ultimement confiné, est réduit à une chaîne atomique (cf chapitre 4). Cependant, l'influence du potentiel coulombien associé à l'écrantage par les charges du canal de l'impureté ionisée est négligé. La section (6.3) étend donc le modèle à une description quantique plus réaliste, capable de simuler un défaut dans des dispositifs émergents 3D décrits au chapitre 5 à l'aide de l'approximation "mode-space". La section (6.4) étudie l'évolution des paramètres physiques qui accompagne la présence d'un défaut et discute l'influence de la position et du type de ce dernier sur les caractéristiques électriques. La dernière section établit les conclusions et perspectives.

6.1 Défauts ponctuels dans une chaîne à l'équilibre

A l'échelle nanométrique, les propriétés électroniques des semiconducteurs sont particulièrement sensibles à la présence d'un défaut ponctuel. Il est reconnu que les défauts ponctuels brisent la périodicité du réseau et introduisent des états localisés dans le gap du matériau. Un état localisé peut être de deux types :

- peu profond (*shallow*) dont l'orbitale est délocalisée sur une vingtaine de distances inter-atomiques,
- profond (*deep*) qui correspond à une orbitale très localisée, incluant principalement les premiers et seconds voisins du défaut.

Le formalisme des fonctions de GREEN permet un traitement de ces défauts assez efficace. Nous considérons dans cette section un défaut neutre pour lequel le potentiel perturbateur est situé dans son voisinage et négligeons l'effet d'une éventuelle queue coulombienne. La longueur d'écrantage dans les semiconducteurs usuels étant de l'ordre de la distance inter-atomique, la description en liaisons fortes d'une impureté substitutionnelle neutre peut se réduire, en première approximation, à un potentiel perturbateur U_d sur le site du défaut. Soit un cristal parfait d'hamiltonien H et d'opérateur de GREEN G . L'introduction d'une impureté peut se représenter par une perturbation U (matrice associée au potentiel intra-atomique U_d) qui conduit à un nouvel hamiltonien $H' = H + U$ et un nouvel opérateur de GREEN G' . L'équation de SCHRÖDINGER du système perturbé est :

$$(\varepsilon - H) |\Psi'\rangle = U |\Psi'\rangle. \quad (6.1)$$

La solution $|\Psi'\rangle$ de ce système différentiel est la somme d'une solution générale $|\Psi\rangle$ de l'équation sans second membre, c'est-à-dire :

$$(\varepsilon - H) |\Psi\rangle = 0, \quad (6.2)$$

et d'une solution particulière qui peut s'écrire $GU |\Psi'\rangle$. Nous obtenons :

$$|\Psi'\rangle = |\Psi\rangle + GU |\Psi'\rangle. \quad (6.3)$$

L'équation (6.3) est valable pour les énergies ε appartenant à la bande d'énergie autorisée dans le système parfait. Dans le cas contraire, c'est-à-dire si ε appartient à un intervalle d'énergie interdit, nous avons :

$$|\Psi'\rangle = GU |\Psi'\rangle. \quad (6.4)$$

Les solutions de cette équation définissent les états localisés et diffèrent de la solution triviale nulle si et seulement si [193] :

$$\det(I - GU) = 0, \quad (6.5)$$

où I est la matrice identité. L'équation (6.5) traduit tout l'intérêt des fonctions de GREEN : l'ensemble des caractéristiques du système (position des niveaux d'énergie, extension de la fonction d'onde correspondante ainsi que la modification de la densité d'états) est lié à la seule connais-

sance de la fonction de GREEN du système non perturbé et à la forme du potentiel perturbateur. Par exemple, la densité d'états locale à l'équilibre d'un atome i s'exprime sous la forme¹ :

$$N_{ii}(\varepsilon) = -\frac{1}{\pi} \text{Im}(G_{ii}(\varepsilon)). \quad (6.6)$$

Nous considérons tout d'abord une chaîne atomique infinie à l'équilibre présentant les mêmes constantes de couplage β et distance inter-atomique a qu'au chapitre 4². Chaque atome possède une orbitale et seules les interactions entre premiers voisins sont envisagées. Nous calculons ensuite la fonction de GREEN de l'atome origine 0 de la chaîne infinie (figure (6.1)a). D'après les équations (4.34) et (4.35), la fonction de GREEN de l'extrémité d'une chaîne semi-infinie à l'équilibre s'écrit :

$$G_{00}^0(\varepsilon) = \frac{\varepsilon}{2\beta^2} \left(1 - \sqrt{1 - \frac{4\beta^2}{\varepsilon^2}} \right) \quad \text{si } |\varepsilon| > 2|\beta|, \quad (6.7)$$

$$G_{00}^0(\varepsilon) = \frac{\varepsilon}{2\beta^2} - \frac{i}{|\beta|} \sqrt{1 - \left| \frac{\varepsilon}{2\beta} \right|^2} \quad \text{si } |\varepsilon| < 2|\beta|, \quad (6.8)$$

La fonction de GREEN de la chaîne infinie se construit alors à partir de deux chaînes semi-infinies dont nous couplons les orbitales pendantes par un potentiel perturbateur β . Nous appliquons l'équation de DYSON en définissant G^0 la fonction de GREEN de deux chaînes semi-infinies isolées et G la fonction de GREEN d'une chaîne infinie :

$$G_{00} = G_{00}^0 + G_{00}^0 \beta G_{10}, \quad (6.9)$$

et

$$G_{10} = G_{11}^0 \beta G_{00}. \quad (6.10)$$

Mais $G_{00}^0 = G_{11}^0$, et en injectant (6.10) dans (6.9), il résulte :

$$G_{00} = \frac{G_{00}^0}{1 - G_{00}^0 \beta^2 G_{00}^0}. \quad (6.11)$$

Nous substituons ensuite G_{00}^0 par les expressions (6.7) et (6.8), la fonction de GREEN de l'atome 0 de la chaîne infinie est :

$$G_{00}(\varepsilon) = -\frac{i}{\sqrt{4\beta^2 - \varepsilon^2}}, \quad \text{si } |\varepsilon| \leq 2|\beta|, \quad (6.12)$$

$$G_{00}(\varepsilon) = \pm \frac{1}{\sqrt{\varepsilon^2 - 4\beta^2}}, \quad \text{si } |\varepsilon| \geq 2|\beta|, \quad (6.13)$$

¹Voir appendice sur les fonctions de GREEN.

² $\beta = -1.25$ eV et $a = 4$ Å.

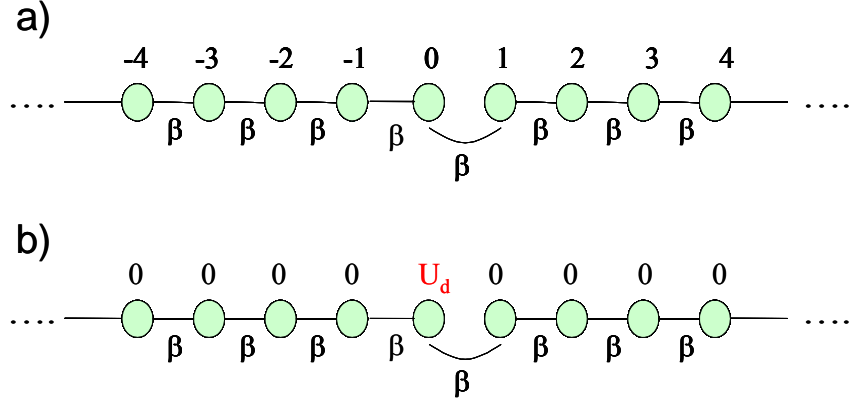


FIG. 6.1: a) Chaîne linéaire atomique infinie dont les atomes sont numérotés de $-\infty$ à $+\infty$. b) La perturbation induite par un défaut ponctuel se traduit par un potentiel intra-atomique U_d localisé sur le site du défaut. β représente l'énergie de couplage.

le signe de l'équation (6.13) étant déterminé par celui de ε . Utilisant l'équation (6.6), la densité d'états locale associée est :

$$N_{00}(\varepsilon) = \frac{1}{\pi} \times \frac{1}{\sqrt{4\beta^2 - \varepsilon^2}}. \quad (6.14)$$

Nous obtenons la densité d'états usuelle 1D divergente en $1/\sqrt{X}$ en bord de bande.

La perturbation induite par un défaut ponctuel neutre (=non ionisé) peut se décrire en ajoutant un terme d'énergie potentielle U_d sur le site du défaut. Les autres interactions, y compris celles du défaut, restent inchangées. La matrice du potentiel perturbateur est donc diagonale, carrée et de rang égal au nombre d'orbitales considérées pour décrire l'impureté (1 dans notre cas). Notant G_{d00} la fonction de GREEN du défaut introduit sur le site 0 (figure (6.1)b), l'équation de DYSON s'écrit :

$$G_{d00} = (1 - G_{00}U_d)^{-1} G_{00}. \quad (6.15)$$

Si $|\varepsilon| \leq 2|\beta|$, la fonction de GREEN du système perturbé devient, ((6.12)) dans (6.15) :

$$G_{d00}(\varepsilon) = -\frac{i\sqrt{4\beta^2 - \varepsilon^2} - U_d}{4\beta^2 - \varepsilon^2 + U_d^2}, \quad (6.16)$$

à partir de laquelle est déduite la densité d'états locale du défaut :

$$N_{00}(\varepsilon) = \frac{1}{\pi} \times \frac{\sqrt{4\beta^2 - \varepsilon^2}}{4\beta^2 - \varepsilon^2 + U_d^2}. \quad (6.17)$$

En présence d'un défaut ponctuel, la densité d'états locale dans la bande d'énergie diminue avec l'augmentation de la perturbation $|U_d|$. L'introduction de défauts atténue l'allure caractéristique de la densité d'états 1D, dans laquelle la divergence des bords de bandes disparaît ($|\varepsilon| = 2|\beta|$). Nous rappelons que la relation de dispersion d'une chaîne linéaire caractérisée par une énergie

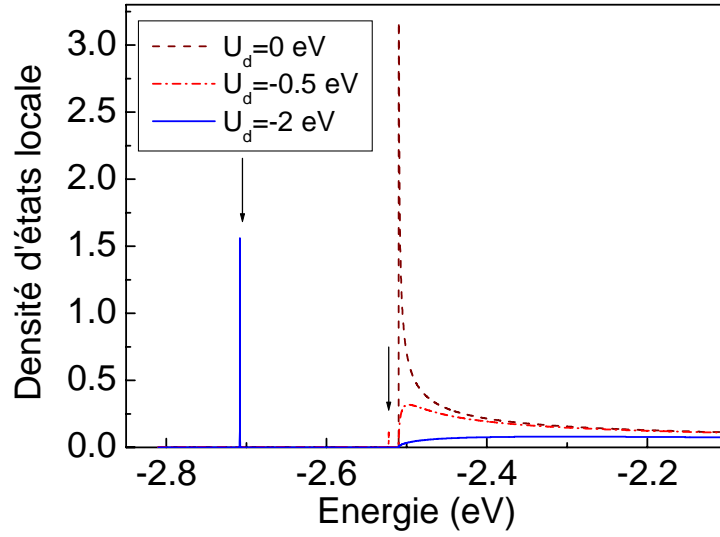


FIG. 6.2: Densité d'états localisée $N_{00}(\varepsilon)$ de l'impureté en fonction du potentiel attractif U_d . On observe la disparition progressive de la densité d'états 1D et la création du niveau localisé sous la bande d'énergie (indiqué par les flèches).

de couplage β et une distance inter-atomique a est donnée par l'expression :

$$\varepsilon = 2\beta \cos ka, \quad (6.18)$$

où k est le vecteur d'onde.

Pour $|\varepsilon| > 2|\beta|$, le calcul de l'énergie de l'état localisé s'obtient en résolvant l'équation (6.5), et conduit à deux valeurs possibles en-dehors de la bande d'énergie :

$$\det(I - GU) = 0 \rightarrow 1 - G_{00}U_d = 0 \rightarrow G_{00} = \frac{1}{U_d}, \quad (6.19)$$

d'où

$$\pm \frac{1}{\sqrt{\varepsilon_d^2 - 4\beta^2}} = \frac{1}{U_d}, \quad (6.20)$$

et finalement,

$$\varepsilon_d = \pm \sqrt{U_d^2 + 4\beta^2}. \quad (6.21)$$

Si $U_d < 0$, les propriétés de la fonction de GREEN [190] excluent la solution positive. La valeur négative de U_d correspond à un potentiel attractif qui crée un état localisé, donneur d'électrons, en-dessous la bande d'énergie. La profondeur de l'état localisée augmente avec $|U_d|$. La figure (6.2) montre l'évolution de la densité d'états locale en fonction de la perturbation U_d . Elle illustre d'une part la création de l'état localisé dont la profondeur augmente avec $|U_d|$ et d'autre part la diminution associée de $N_{00}(\varepsilon)$ afin d'assurer la normalisation. De plus, l'augmentation du pic de l'état localisé avec $|U_d|$ reflète une localisation croissante de l'électron sur le niveau de l'impureté. Nous pouvons également évaluer l'extension de l'état localisé en appliquant l'équation maîtresse de la méthode des variations (cf § (4.1.2)). Dans une base d'orbitales atomiques $\{|n\rangle\}_{n \in \mathbb{Z}}$, l'état

lié $|\Psi_d\rangle$ peut s'écrire :

$$|\Psi_d\rangle = \sum_n a_n |n\rangle, \quad (6.22)$$

où les $|a_n|^2$ définissent le poids de l'état $|\Psi_d\rangle$ sur l'orbitale atomique $|n\rangle$.

· L'équation $\langle n|H - \varepsilon|\Psi_d\rangle = 0$ (méthode des variations) se traduit sur le défaut (en 0) par :

$$(\varepsilon - U_d) a_0 = \beta (a_{-1} + a_1). \quad (6.23)$$

· Loin du défaut (sur $i^{\text{ième}}$ atome), nous obtenons :

$$\varepsilon a_i = \beta (a_{i-1} + a_{i+1}). \quad (6.24)$$

Un tel système admet des solutions en $(a_l)_{l \in \mathbb{Z}}$ de la forme :

$$a_l = a_0 K^{|l|}. \quad (6.25)$$

Il en résulte deux équations :

$$\begin{aligned} (\varepsilon - U_d) &= 2\beta K \quad (\text{sur le défaut}) \\ \varepsilon &= \beta \left(K + \frac{1}{K} \right) \quad (\text{loin du défaut}) \end{aligned} \quad (6.26)$$

La résolution du système (6.26) fournit les solutions suivantes :

$$K = -\frac{U_d}{2\beta} \pm \sqrt{\left(\frac{U_d}{2\beta}\right)^2 + 1}, \quad (6.27)$$

où le choix du signe est déterminé par l'opposé de celui de U_d afin de conserver le bon niveau localisé ε_d en injectant K dans l'équation (6.26). La figure (6.3) montre les probabilités de présence ($|\langle n|\Psi_d\rangle|^2$) calculées à partir de l'équation (6.27) pour différentes valeurs de potentiels U_d . Notons qu'un comportement symétrique se produit pour une valeur positive de U_d . Dans ce cas, l'état localisé a une énergie donnée par la solution positive de l'équation (6.21) et correspond à un niveau accepteur d'électrons.

L'influence d'un défaut ponctuel est également visible sur le coefficient de transmission. La figure (6.4) montre l'effet d'un défaut donneur ($U_d < 0$) sur la transparence calculée à l'aide du formalisme des fonctions de GREEN présenté au paragraphe (4.2). On constate une diminution de la transmission pour des défauts profonds, due à la réduction de $N_{00}(\varepsilon)$. Des potentiels très attractifs ($U_d = -20$ eV) conduisent au cas limite de la liaison brisée de la lacune, pour laquelle la transmission s'annule. La limite $U_d = -\infty$ décale le niveau localisé vers les énergies infiniment négatives, et l'impureté n'interagit plus avec les états du reste du système.

Cette approche simple permet néanmoins de simuler une gamme variée de défauts, depuis l'impureté peu profonde jusqu'à la lacune. Nous avons étudié l'influence du défaut ponctuel

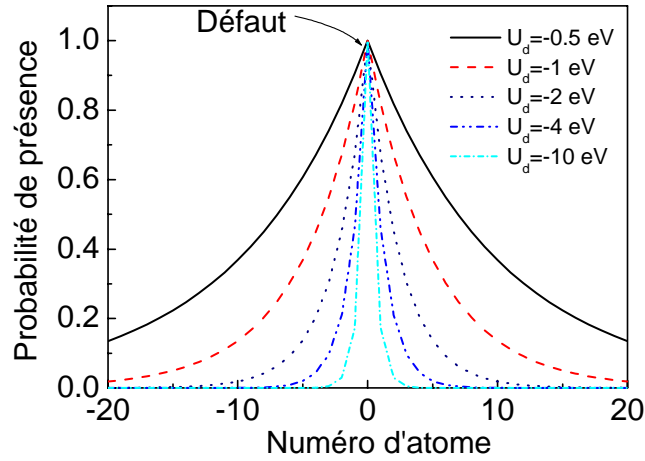


FIG. 6.3: Extension des états localisés en fonction du potentiel perturbateur du défaut U_d .

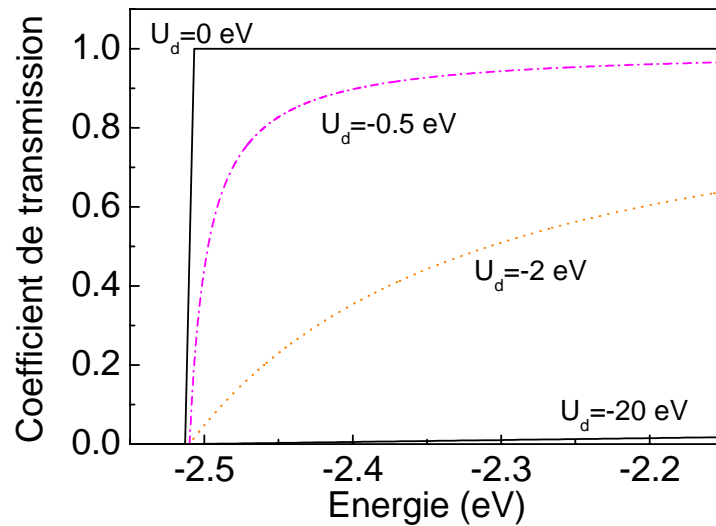


FIG. 6.4: Coefficients de transmission à travers une chaîne infinie à l'équilibre pour différentes valeurs du potentiel attractif U_d . L'état localisé induit une diminution de la transparence qui converge, pour de fortes valeurs de $|U_d|$, vers la limite du cas lacunaire.

sur la densité d'états locale et le coefficient de transmission. Dans la prochaine section, nous utilisons ce modèle pour traiter l'influence des défauts ponctuels sur le courant d'un MOSFET double-grille ultimement confiné.

6.2 Influence du défaut dans un MOSFET double-grille

Nous considérons le modèle du MOSFET double-grille décrit au chapitre 4 dans lequel une impureté est introduite dans le canal. L'impureté est toujours modélisée par un potentiel perturbateur intra-atomique U_d placé sur le site du défaut. Dans le cadre du système polarisé, le potentiel total est la superposition du potentiel auto-cohérent du système sans défaut et du potentiel local U_d caractérisant le défaut. Cette approche semble raisonnable puisque le premier est macroscopique alors que le second est purement local, les deux étant usuellement séparables. Dans toute cette section la longueur de grille L_G sera constante, égale à 10 nm.

Afin de mieux comprendre les effets engendrés par un défaut ponctuel, nous comparons les caractéristiques électriques avec et sans défaut. Il faut tout d'abord distinguer le potentiel répulsif ($U_d > 0$) du potentiel attractif ($U_d < 0$) qui créent respectivement un niveau accepteur et donneur d'électrons. La figure (6.5) montre l'influence de la position du défaut ponctuel le long du canal sur les caractéristiques du courant de drain (I_D) en fonction de la tension de grille (V_G). Un défaut accepteur est dans un premier temps introduit dans le canal à deux nanomètres de la source, puis à deux nanomètres du drain. Deux amplitudes de potentiel de site sont étudiées : $U_d = 1$ eV et $U_d = 3$ eV. La comparaison avec le régime du réseau parfait montre que le courant sous le seuil en présence d'une impureté, quelle que soit sa position dans le canal et son potentiel intra-atomique, est toujours plus faible. Les pentes sous le seuil restent inchangées, mais les courants associés aux défauts sont décalés vers le bas.

Comme discuté dans la section précédente, l'impureté induit une diminution de la densité d'états, compensée par la création d'un état localisé au-dessus du continuum d'énergie. Il en résulte une augmentation de la composante tunnel du courant total qui réduit le coefficient de transmission. La figure (6.5) montre également qu'un défaut donneur introduit près de la source dégrade plus fortement le courant que le même défaut positionné près du drain. L'impureté près de la source produit une importante réflexion des électrons qui peuvent facilement ré-intégrer le contact en réduisant le courant total. Lorsqu'une impureté identique est introduite dans la seconde partie du composant, les électrons se réfléchissent toujours, mais la barrière de potentiel du canal ainsi que la polarisation du drain s'opposent à leur retour vers la source : le courant est moins dégradé. Cette analyse est confirmée par la figure (6.6), dans laquelle les coefficients de transmission sont représentés dans l'état *off* ($V_G = 0$ V) pour les cinq cas précédemment décrits. La transparence du système sans défaut est toujours plus importante que celles associées à un défaut accepteur. Pour un potentiel U_d donné, les transparences sont assez similaires en bas de bande d'énergie du canal côté source (-2.5 eV $\lesssim \varepsilon \lesssim -2.3$ eV). En revanche, le coefficient de transmission correspondant à un défaut introduit près du drain augmente plus rapidement

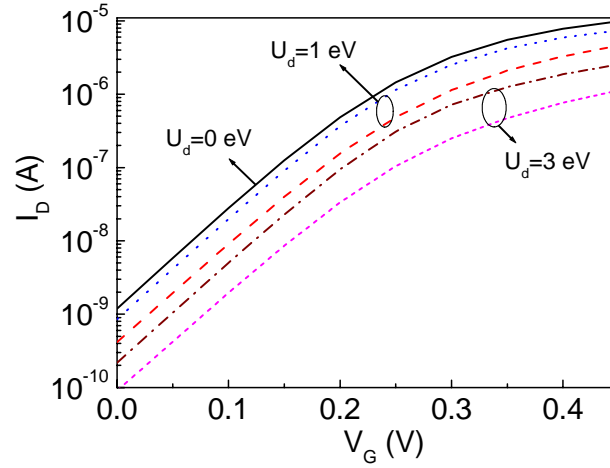


FIG. 6.5: Caractéristiques I_D-V_G d'un système sans défaut (ligne continue), et dans le modèle de l'impureté accepteuse ($U_d > 0$). L'influence de la position de l'impureté le long du canal est étudiée : le défaut est d'abord introduit à deux nanomètres de la source ($U_d = 1$ eV en ligne discontinue, $U_d = 3$ eV en "petits tirets"), puis à deux nanomètres du drain ($U_d = 1$ eV en pointillés, $U_d = 3$ eV en "pointillé-tirets"). $V_{DS} = 0.3$ V.

aux hautes énergies ($\varepsilon \gtrsim -2.2$ eV). Cette évolution traduit l'influence de V_{DS} sur le transport électronique près du drain, région dans laquelle la probabilité qu'un électron soit totalement réfléchi par un défaut est faible dès que l'énergie cinétique de la particule augmente. Nous étudions maintenant l'effet d'un défaut donneur d'électrons ($U_d < 0$) sur les caractéristiques I_D-V_G . La figure (6.7) compare les courants de drain d'un système contenant un défaut donneur avec celui d'un réseau parfait. Deux positions d'impuretés sont toujours considérées (à deux nanomètres de la source et à deux nanomètres du drain) pour deux amplitudes de potentiel intra-atomique. La présence d'un défaut près du drain, redonne les résultats du défaut accepteur, c'est-à-dire que le courant est globalement diminué par rapport à la limite du cristal sans défaut. En revanche, la présence d'une impureté peu profonde ($U_d = -1$ eV) près de la source produit un courant sous le seuil plus important que celui donné par le composant parfait.

La figure (6.8), qui représente les coefficients de transmission dans l'état *off* pour plusieurs valeurs du potentiel du défaut, clarifie les résultats du courant. En bas de la bande d'énergie de la source ($\varepsilon \simeq -2.5$ eV), la présence d'une impureté peu profonde ($|U_d| < 2$ eV) s'accompagne d'une plus grande transparence que celle du composant parfait. Pour de plus grandes énergies, l'impureté introduite près de la source donne des coefficients de transmission supérieurs à ceux fournis par la même impureté située près du drain. Près de la source, le défaut donneur crée un état localisé sous la bande d'énergie qui induit un effet de tunnel assisté à travers la barrière de potentiel du canal. Les pics, visibles dans la transparence, pour des impuretés peu profondes, illustrent clairement l'effet tunnel via l'état localisé. Si le défaut est plus profond ($U_d = -5$ eV), les électrons ne peuvent plus "tunneler" à travers l'état localisé qui se comporte alors comme une perturbation diminuant la probabilité de transmission des électrons. Dans la situation opposée où un défaut donneur est introduit près du drain, la tension V_{DS} décale vers le bas le niveau de l'état localisé qui n'interagit pas avec les électrons de forte énergie ($\varepsilon > -2.3$ eV) : le coefficient de

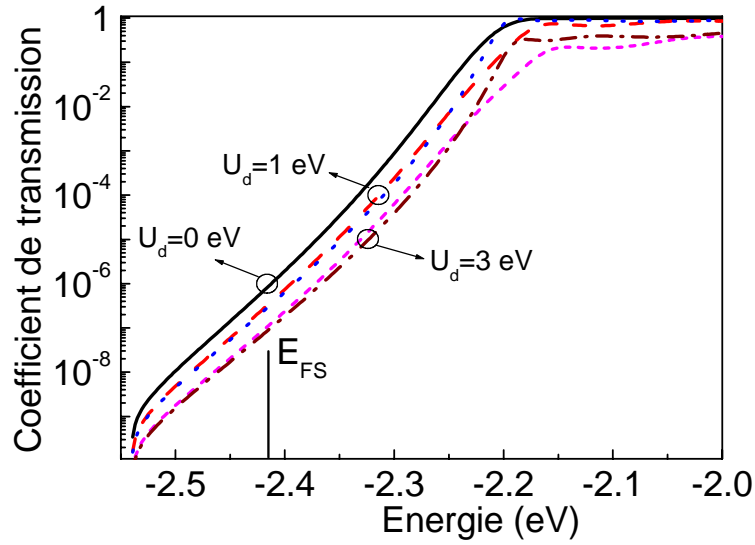


FIG. 6.6: Coefficients de transmission à travers le canal pour un système sans défaut (ligne continue) ; pour un système contenant un défaut accepteur d'abord introduit à deux nanomètres de la source ($U_d = 1$ eV en ligne discontinue, $U_d = 3$ eV en "petits tirets"), puis à deux nanomètres du drain ($U_d = 1$ eV en pointillés, $U_d = 3$ eV en "pointillé-tirets"). $V_{DS} = 0.3$ V, $V_G = 0$ V et E_{FS} est le niveau de FERMI du contact de la source.

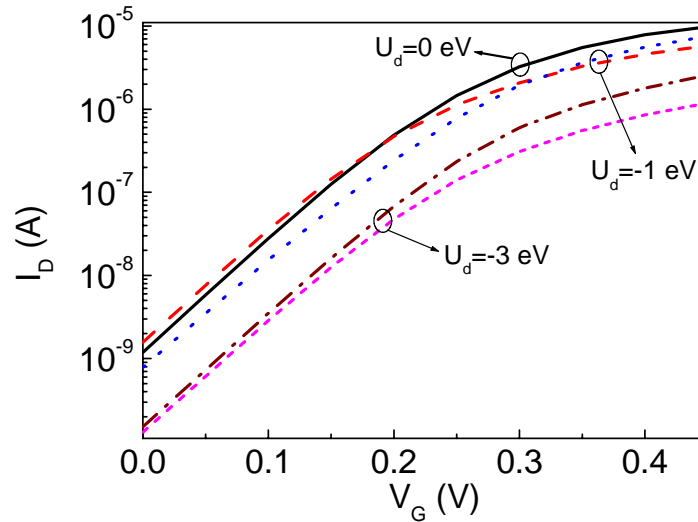


FIG. 6.7: Caractéristiques I_D - V_G dans un système sans défaut (ligne continue), et dans le modèle du défaut donneur ($U_d < 0$). L'influence de la position de l'impureté le long du canal est étudiée : le défaut est d'abord introduit à deux nanomètres de la source ($U_d = -1$ eV en ligne discontinue, $U_d = -3$ eV en "petits tirets"), puis à deux nanomètres du drain ($U_d = -1$ eV en pointillés, $U_d = -3$ eV en "pointillé-tirets"). $V_{DS} = 0.3$ V.

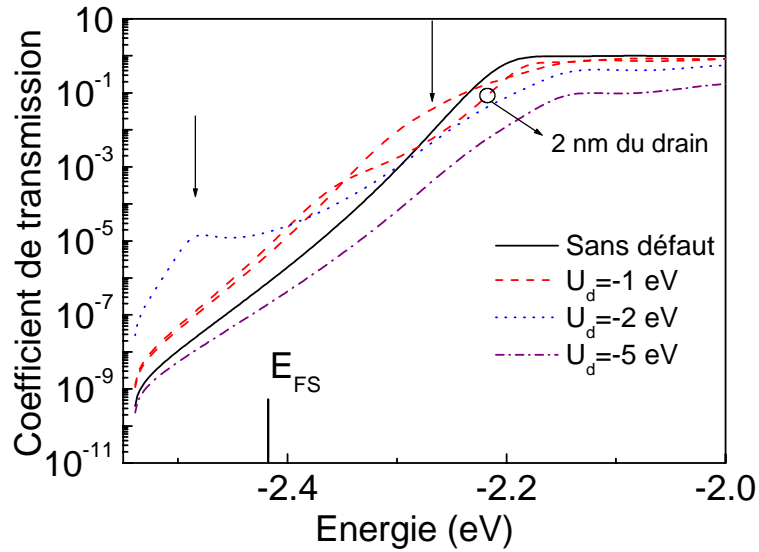


FIG. 6.8: Coefficients de transmission à travers le canal dans le cas d'un cristal parfait, d'un défaut donneur introduit à 2 nm de la source et à 2 nm du drain (comme indiqué). Les flèches indiquent les pics de transmission associés à l'effet tunnel assisté à travers les états localisés. $V_{DS} = 0.3$ V, $V_G = 0$ V et E_{FS} est le niveau de FERMÍ du contact de source.

transmission ne présente pas de pic. Enfin la figure (6.9) montre les caractéristiques $I_D - V_{DS}$ dans l'état passant ($V_G = 0.4$ V) du composant sans défaut, et pour plusieurs composants contenant une impureté introduite à 2 nm de la source. En présence de défauts ponctuels, le courant simulé est toujours inférieur à celui obtenu avec un réseau parfait. Dans l'état passant, la diminution de la barrière de potentiel du canal atténue fortement la composante tunnel du courant total et l'état localisé créé par le défaut induit uniquement de la diffusion. Nous avons également représenté sur la figure (6.9) le courant fourni par le quantum de conductance G_0 (cf § (3.3.1)), duquel s'éloignent les caractéristiques des composants lorsque U_d devient fortement attractif : cette évolution illustre la diminution du coefficient de transmission pour un état localisé profond. L'état localisé d'un niveau très profond ($U_d < -5$ eV) ne participe plus au transport électronique et le comportement du défaut ponctuel se rapproche progressivement de celui d'une lacune à travers laquelle aucun électron ne peut être transmis. Dans le cas d'un canal avec un seul mode de conduction, le composant agit alors comme un transistor bloqué en permanence.

Commentaires

Nous avons développé une approche basée sur le formalisme des fonctions de Green hors-équilibre, qui simule un MOSFET double-grille à canal unique de conduction en présence de défauts ponctuels. Nous avons de façon générale obtenu une diminution du courant de drain traduisant la réflexion des électrons par le défaut. Cependant la présence d'un défaut donneur introduit près de la source peut produire un transport par effet tunnel assisté et conduire à des courants sous le seuil plus élevés que ceux obtenus avec un cristal parfait : un réseau idéal n'est pas la limite supérieure du courant dans les fils quantiques.

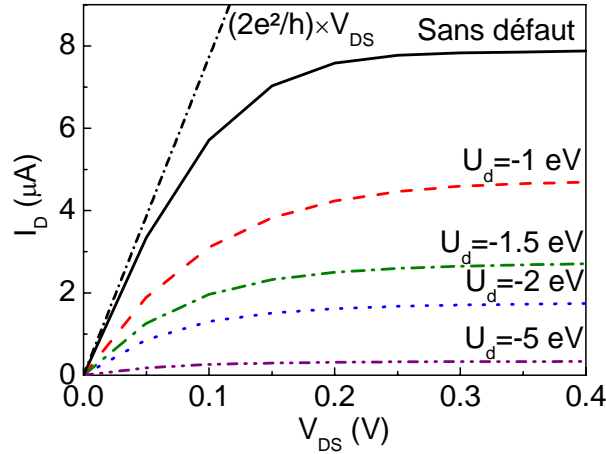


FIG. 6.9: Caractéristiques $I_D - V_{DS}$ d'un composant parfait ou dans lequel une impureté est introduite à 2 nm de la source. Considérant plusieurs valeurs du potentiel attractif U_d , le courant dans l'état passant en présence d'impureté varie de 5 à 55% de la limite du cristal sans défaut. $V_G = 0.4$ V

Le présent traitement souffre néanmoins de ne pas inclure le potentiel de COULOMB associé à une impureté ionisée et écrantée par les électrons du reste du système. Les sections suivantes se proposent donc d'étendre ce modèle à des dispositifs réalistes 3D en considérant la queue coulombienne du défaut chargé comme une partie du potentiel macroscopique total.

6.3 Défauts ponctuels ionisés dans les MOSFETs multi-grilles

Considérons le transistor MOSFET *gate-all-around* (GAA) de la figure (6.10) dans lequel l'interface Si/SiO₂ est perpendiculaire à la direction cristalline $\langle 100 \rangle$. Le système, traité en liaisons fortes, est supposé en première approximation présenter une structure cubique simple (figure (6.11)) dans laquelle la distance inter-atomique a est identique suivant toutes les directions ($a = 0.2$ nm). Les réservoirs de source et drain ont une distribution de dopants simulée par un dopage continu égal à 10^{26} m⁻³. En revanche, la concentration des impuretés dopantes du canal, considérée homogène dans les composants macroscopiques, subit désormais des fluctuations aléatoires spatiales. En effet, un canal dopé de type P avec une concentration 10^{25} m⁻³, et dont les dimensions sont $L_G \times T_{Si} \times W_{Si} = 10 \times 5 \times 3$ nm³, contient en moyenne 1.5 impureté. Cette étude s'intéresse donc à l'influence sur les paramètres physiques et les propriétés électriques de la position et du type d'impureté du canal.

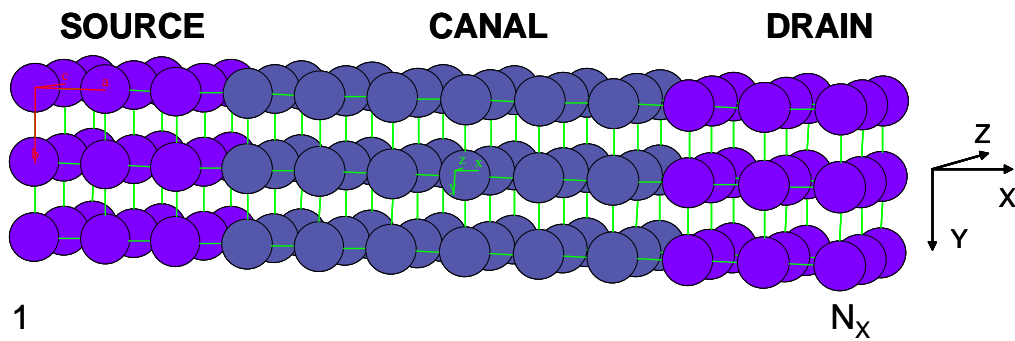


FIG. 6.11: Le silicium et son oxyde sont supposés, en première approximation, présenter une structure cubique simple dans laquelle seules les interactions entre premiers voisins sont considérées. Par volonté de clarté la structure des oxydes de grille n'est pas représentée.

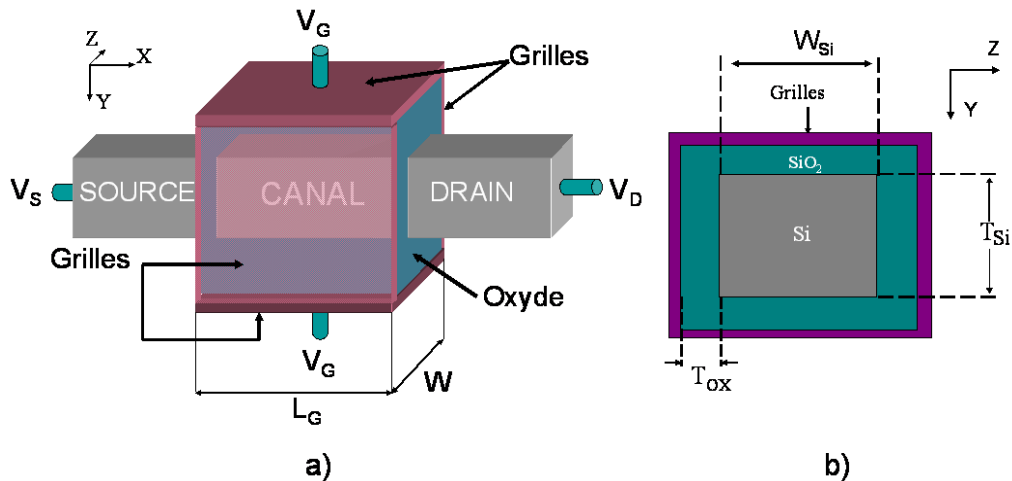


FIG. 6.10: Schéma d'un transistor MOSFET Gate-All-Around (GAA) : a) vue en perspective, b) coupe transverse.

Une impureté ionisée est définie par un potentiel d'interaction coulombien à longue distance combiné à un potentiel intra-atomique que nous avons précédemment traité dans le cas d'une impureté neutre (non ionisée). La queue coulombienne, qui résulte des interactions électrostatiques du défaut avec les électrons du système, est une variation lente et macroscopique du potentiel électrostatique total. En revanche, le potentiel intra-atomique génère des variations très localisées qui modifient au voisinage du défaut les états transverse 2D associés au confinement des plans atomiques y - z .

L'approche "mode-space" utilisée dans le chapitre 5, et qui consiste à séparer le traitement du confinement de celui du transport, n'est donc plus valide. Les états propres transverse varient le long de l'axe du transport et les différentes sous-bandes d'énergie ne sont plus indépendantes. Une solution est cependant de dissocier les deux termes du potentiel du défaut en incluant le potentiel de COULOMB macroscopique dans l'auto-cohérence de l'approche "mode-space" et de traiter le potentiel local intra-atomique une fois l'auto-cohérence atteinte, à l'aide d'une

description 3D dans l'espace réel. Le potentiel intra-atomique n'induisant qu'une perturbation très localisée, son influence sur le courant total devrait être de faible envergure dans un système réaliste 3D et cautionne donc cette décomposition.

Nous adoptons de nouveau l'approche dans laquelle l'équation de POISSON 3D est résolue de façon auto-cohérente avec l'équation de SCHRÖDINGER 3D. L'approximation "mode-space" divise l'équation de SCHRÖDINGER 3D en une équation 2D incluant le confinement de la section transverse (plan $\mathbf{y-z}$), et une équation 1D décrivant le transport balistique le long de l'axe source-drain à l'aide du formalisme des fonctions de GREEN hors-équilibre. La zone active (silicium et oxydes de grilles) est représentée par une version simplifiée de la théorie des liaisons fortes avec une seule orbitale par atome et en considérant seulement les interactions entre premiers voisins³.

Nous reprenons un calcul des charges en tout point similaire à celui du chapitre 5 dans lequel le potentiel macroscopique coulombien de l'impureté est incorporé. L'hamiltonien 2D (H_{conf}) de chaque plan atomique n perpendiculaire à l'axe source-drain est tout d'abord résolu :

$$H_{conf} |\psi_{n,i}\rangle = \varepsilon_{n,i} |\psi_{n,i}\rangle, \quad (6.28)$$

avec :

$$\begin{aligned} H_{conf} &= \sum_{l,m} V_{lmn,lmn}^{def} |l, m, n\rangle \langle l, m, n| \\ &+ \beta_y \sum_{l,m} |l, m, n\rangle \langle l, m+1, n| \\ &+ \beta_z \sum_{l,m} |l, m, n\rangle \langle l+1, m, n|, \end{aligned} \quad (6.29)$$

où $|l, m, n\rangle$ est le ket représentant l'orbitale de l'atome de coordonnées $z = l \times a$, $y = m \times a$ du $n^{ième}$ plan vertical. $V_{lmn,lmn}^{def}$ est l'élément de matrice du potentiel électrostatique supposé diagonal dans la base atomique. Il est la somme du potentiel $V_{lmn,lmn}$ fourni par l'équation de POISSON, et du potentiel coulombien $V_{lmn,lmn}^{Coul}$ défini par :

$$V_{lmn,lmn}^{Coul} = \pm \frac{1 - e^{-\sqrt{(l-l')^2 + (m-m')^2 + (n-n')^2}/L_C}}{4\pi\epsilon_0\epsilon_{Si} \times \left(\sqrt{(l-l')^2 + (m-m')^2 + (n-n')^2}\right)}, \quad (6.30)$$

où l' , m' et n' caractérisent les coordonnées du défaut ponctuel et L_C est une longueur caractéristique, égale à 10^{-10} m, permettant de lisser l'allure du potentiel coulombien au voisinage du défaut afin d'éviter qu'il ne diverge. Le signe de $V_{lmn,lmn}^{Coul}$ détermine le type de l'impureté ($<0 \rightarrow$ défaut donneur, $>0 \rightarrow$ défaut accepteur). $|\psi_{n,i}\rangle$ et $\varepsilon_{n,i}$ sont respectivement les vecteurs et les valeurs propres des plans transverses. Enfin β_y et β_z sont les énergies de couplages entre premiers voisins le long des directions correspondantes. Elles sont reliées aux masses effectives

³Cette représentation est équivalente à l'approximation de la masse effective (cf § (4.1.2)).

par les expressions suivantes (cf § (4.1.2)) :

$$\beta_y = -\frac{\hbar^2}{2m_y^*a^2}, \text{ et } \beta_z = -\frac{\hbar^2}{2m_z^*a^2}. \quad (6.31)$$

La structure cubique simple est étendue aux oxydes de grille dont la description est intégrée dans l'hamiltonien H_{conf} autorisant ainsi la pénétration des fonctions d'onde $|\psi_{n,i}\rangle$. Les mêmes couplages β_y et β_z sont conservés dans les oxydes puisqu'il a été vérifié que la différence de gap entre le Si et le SiO₂ (≈ 3 eV) déterminait pour l'essentiel la pénétration de l'état transverse dans l'oxyde.

Si la perturbation engendrée par le potentiel coulombien reste macroscopique et s'étend sur plusieurs plans atomiques, les électrons se répartissent en sous-bandes indépendantes et la matrice de l'hamiltonien 3D est diagonale par blocs :

$$H_{3D} = \begin{bmatrix} h_1 & 0 & \cdots & \cdots & \cdots \\ 0 & h_2 & 0 & \cdots & \cdots \\ \cdots & 0 & \ddots & 0 & \cdots \\ \cdots & \cdots & 0 & h_i & 0 \\ \cdots & \cdots & \cdots & 0 & \ddots \end{bmatrix}, \quad (6.32)$$

où

$$h_i = \begin{bmatrix} \varepsilon_{1,i} & \beta_x & 0 & \cdots \\ \beta_x & \varepsilon_{2,i} & \beta_x & 0 \\ 0 & \beta_x & \ddots & \beta_x \\ \cdots & 0 & \beta_x & \varepsilon_{N_X,i} \end{bmatrix} \quad (6.33)$$

est l'hamiltonien de la sous-bande i , N_X est le nombre d'atomes le long de l'axe source-drain et β_x est l'énergie de couplage suivant x reliée à la masse effective par (cf § (4.1.2)) :

$$\beta_x = -\frac{\hbar^2}{2m_x^*a^2}. \quad (6.34)$$

Nous calculons alors la fonction de GREEN retardée de la $i^{\text{ième}}$ sous-bande :

$$G_i(\varepsilon) = \lim_{\eta \rightarrow 0^+} [(\varepsilon + i\eta)I - h_i - \Sigma^i], \quad (6.35)$$

où ε est l'énergie et Σ^i est la self-énergie des contacts qui modélise l'élargissement des niveaux d'énergie du système fini lors des couplages avec les réservoirs d'électrons :

$$\Sigma^i = \Sigma_S^i + \Sigma_D^i. \quad (6.36)$$

Σ_S^i et Σ_D^i traduisent le couplage des extrémités de la sous-bande i de la zone active avec les réservoirs macroscopiques de la source et du drain respectivement. Leur expression est détaillée au paragraphe (5.1.2). Nous définissons ensuite deux nouvelles quantités qui représentent la force

d'interaction entre les réservoirs et la zone active :

$$\Gamma_S^i = i \left(\Sigma_S^i - (\Sigma_S^i)^\dagger \right) \text{ et } \Gamma_D^i = i \left(\Sigma_D^i - (\Sigma_D^i)^\dagger \right), \quad (6.37)$$

et qui ont la dimension d'une énergie.

Les fonctions de la densité spectrale des contacts s'expriment alors sous la forme :

$$A_S^i = G_i \Gamma_S^i G_i \text{ et } A_D^i = G_i \Gamma_D^i G_i. \quad (6.38)$$

Il en résulte l'expression de la matrice densité électronique le long de la sous-bande i :

$$n_i(\varepsilon) = \frac{1}{\pi a} \times [f_S(\varepsilon) A_S^i + f_D(\varepsilon) A_D^i]. \quad (6.39)$$

La densité électronique totale 3D s'obtient finalement en intégrant l'équation (6.39) sur l'énergie, en sommant les contributions de chaque vallée et sous-bandes, et en les multipliant par le module au carré de l'état transverse $|\langle l, m, n | \psi_{n,i} \rangle|^2$:

$$n_{l,m,n}^{3D} = \sum_{l,m,n} \sum_i |\langle l, m, n | \psi_{n,i} \rangle|^2 \int_0^{+\infty} n_i(\varepsilon) d\varepsilon, \quad (6.40)$$

où i représente les premières sous-bandes de chaque vallée, c'est-à-dire celles peuplées d'électrons. La densité rétro-agit ensuite sur l'équation de POISSON 3D. Le nouveau potentiel électrostatique $V_{lmn,lmn}$ obtenu, nous lui ajoutons le potentiel coulombien centré autour du site du défaut ponctuel $V_{lmn,lmn}^{Coul}$ défini par (6.30). A chaque boucle de SCHRÖDINGER-POISSON, l'impact du potentiel coulombien est ainsi pris en compte. Lorsque l'auto-cohérence est atteinte, le courant du canal est calculé en clivant virtuellement le système en deux régions dans le plan du défaut ponctuel (figure (6.12) et section (3.3.3)). A ce stade nous ne calculons pas le courant de chaque sous-bande, mais utilisons une description dans l'espace réel 3D afin d'inclure par la suite le potentiel intra-atomique du défaut. Cette approche nous permet de modéliser le potentiel localisé du défaut sans appréhender un couplage entre sous-bandes puisque la validité du traitement 3D dans l'espace réel ne repose pas sur l'indépendance des sous-bandes. Le passage de l'espace des modes à l'espace réel s'effectue en calculant la fonction de GREEN de chaque surface S_1 et S_2 . Concrètement si $G_i^1(\varepsilon)$ est la fonction de GREEN de la sous-bande i dans la région 1, la fonction de GREEN de la surface S_1 s'écrit :

$$\langle l, m, n | G_{S_1}(\varepsilon) | l', m', n' \rangle = \sum_i \langle l, m, n | \psi_{n,i} \rangle \langle \psi_{n,i} | G_i^1(\varepsilon) | \psi_{n',i} \rangle \langle \psi_{n',i} | l', m', n' \rangle, \quad (6.41)$$

De même, si $G_i^2(\varepsilon)$ est la fonction de GREEN de la sous-bande i dans la région 2, la fonction de GREEN de la surface S_2 est :

$$\langle l, m, n | G_{S_2}(\varepsilon) | l', m', n' \rangle = \sum_i \langle l, m, n | \psi_{n,i} \rangle \langle \psi_{n,i} | G_i^2(\varepsilon) | \psi_{n',i} \rangle \langle \psi_{n',i} | l', m', n' \rangle. \quad (6.42)$$

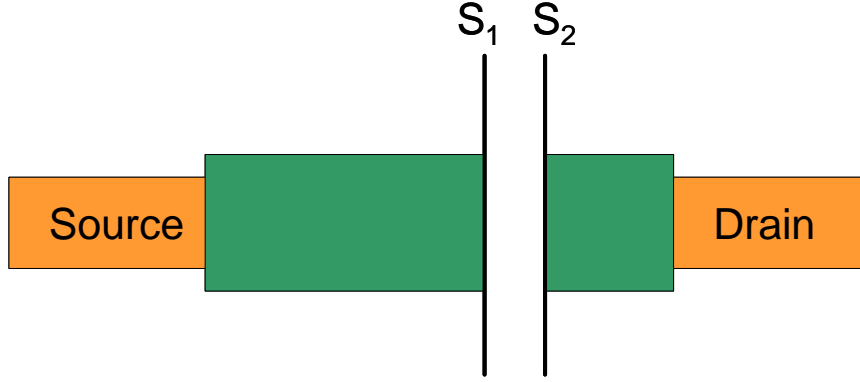


FIG. 6.12: *Calcul du courant : le système 3D est virtuellement clivé en deux régions dans le plan du défaut afin de déterminer les charges circulant à travers les sections S_1 et S_2 . Le transport étant balistique, le nombre d'électrons transmis est constant le long de la zone active, et le plan du clivage n'a pas d'incidence sur le résultat.*

La somme sur i des équations (6.41) et (6.42) s'étend sur tous les modes transverses de façon à prévoir le couplage par le potentiel intra-atomique de sous-bandes très éloignées. Enfin, les kets $|l, m, n\rangle$ et $|l', m', n'\rangle$ représentent les orbitales des surfaces S_1 puis S_2 respectivement.

Le potentiel intra-atomique de l'impureté U_d est ensuite pris en compte à l'aide de l'équation de DYSON. Suivant que le défaut se situe sur la surface S_1 ou S_2 , la fonction de GREEN associée sera modifiée en considérant le potentiel U_d comme une perturbation. On a donc :

$$G'_{S_1}(\varepsilon) = (I_S - G_{S_1}(\varepsilon)U)^{-1} G_{S_1}(\varepsilon), \quad (6.43)$$

si le défaut est sur la surface S_1 , et :

$$G'_{S_2}(\varepsilon) = (I_S - G_{S_2}(\varepsilon)U)^{-1} G_{S_2}(\varepsilon), \quad (6.44)$$

si le défaut est sur la surface S_2 . U désigne la matrice de perturbation associée au défaut, qui dans le cas présent a tous ses éléments nuls excepté celui de l'orbitale du site du défaut, égal à U_d .

La matrice de la fonction de GREEN du système découpé exprimée dans l'espace réel s'écrit :

$$G_S^0(\varepsilon) = \begin{bmatrix} G'_{S_1}(\varepsilon) & 0 \\ 0 & G_{S_2}(\varepsilon) \end{bmatrix}, \quad (6.45)$$

si le défaut est sur la surface S_1 . La matrice de couplage entre les deux régions est définie par les matrices des blocs non-diagonaux de l'hamiltonien des surfaces S_1 et S_2 :

$$C = \begin{bmatrix} 0 & H_{12} \\ H_{21} & 0 \end{bmatrix}. \quad (6.46)$$

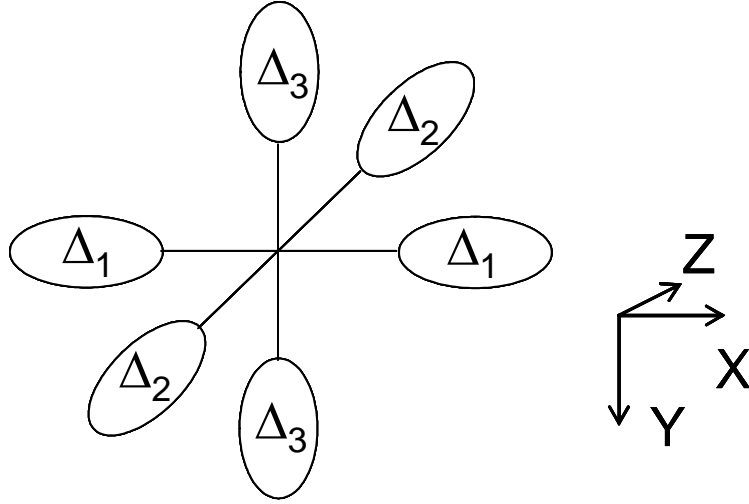


FIG. 6.13: Représentation des six ellipsoïdes équivalents du silicium volumique. Sous l'influence du confinement, trois paires de vallées prennent des énergies différentes.

En se reportant au paragraphe (5.1.2), l'expression finale du courant est :

$$I = -\frac{4\pi e}{\hbar} \int_{-\infty}^{+\infty} d\varepsilon \operatorname{Tr}_1 \left[N_{11}^0 \Lambda_{11}^\dagger C_{12} N_{22}^0 C_{21} \Lambda_{11} \right] (f_S(\varepsilon) - f_D(\varepsilon)), \quad (6.47)$$

où Tr_1 désigne la trace restreinte à la base des orbitales de la surface S_1 , $N^0 = [(-1/\pi) \operatorname{Im}(G_S^0)]$ est la matrice de la densité à l'équilibre du système non-couplé et $\Lambda = [I - G_S^0 C G_S^0 C]^{-1}$.

Le calcul doit être bien entendu répété pour toutes les vallées du silicium. Nous considérons toujours les six ellipsoïdes équivalents du silicium volumique définis par les masses effectives suivantes :

· Vallées Δ_1 :

$$- m_x^* = m_l, \text{ et } m_y^* = m_z^* = m_t,$$

· Vallées Δ_2 :

$$- m_x^* = m_z^* = m_t, \text{ et } m_y^* = m_l,$$

· Vallées Δ_3 :

- $m_x^* = m_y^* = m_t$, et $m_z^* = m_l$. Pour des raisons de confinement et de valeurs de masses effectives discutées dans le paragraphe (5.1.2), seules les vallées Δ_2 et Δ_3 participent efficacement au transport et seront traitées dans cette étude.

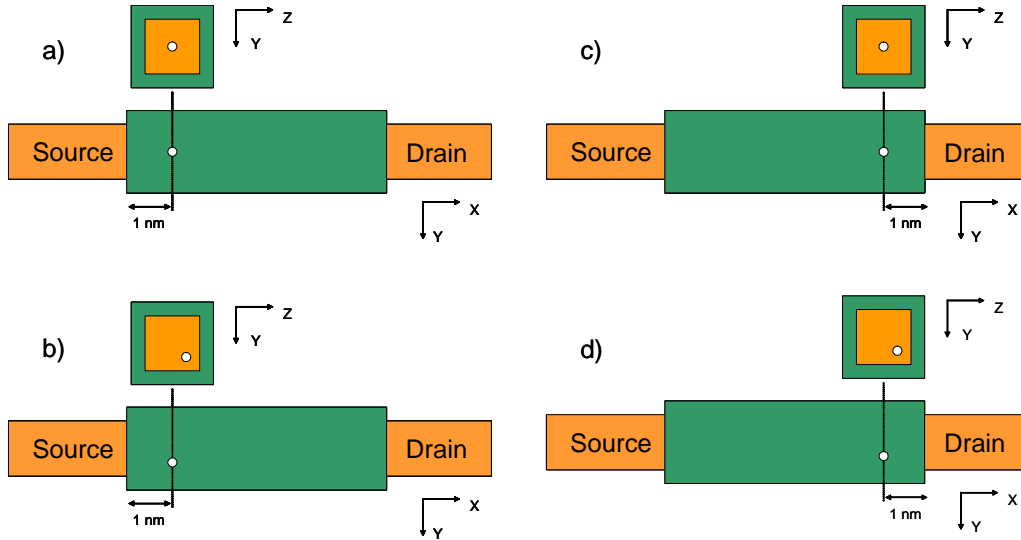


FIG. 6.14: Afin d'analyser l'influence d'une impureté sur le transport, nous définissons quatre positions références de l'impureté dopante dans le canal (cercles blancs) : a) défaut introduit à 1 nm de la source, centré dans la section transverse ; b) défaut introduit à 1 nm de la source, excentré à 0.4 nm des bords du fil de silicium ; c) défaut introduit à 1 nm du drain, centré dans la section transverse ; d) défaut introduit à 1 nm du drain, excentré à 0.4 nm des bords du fil de silicium.

6.4 Résultats

Nous considérons dans cette section le transistor MOSFET *gate-all-around* de la figure (6.10) avec un dopage homogène de type N de 10^{20} cm^{-3} dans la source et le drain. La longueur de grille L_G vaut 8 nm et les dimensions transverses sont : $W_{\text{Si}} = T_{\text{Si}} = 3 \text{ nm}$. L'épaisseur de l'oxyde de grille est également un paramètre fixé à $T_{\text{ox}} = 1 \text{ nm}$. Le canal présente une concentration de défauts accepteurs $N_A = 10^{19} \text{ cm}^{-3}$, ce qui correspond en moyenne à 0.72 impureté dans le volume du canal $V (= L_G \times W_{\text{Si}} \times T_{\text{Si}} = 72 \text{ nm}^3)$. Le dopage de cette région revêt en conséquence une concentration microscopique hautement aléatoire dont nous tenons compte en n'introduisant qu'une seul défaut accepteur dans cette région. La suite de ce travail consiste à mesurer l'influence de la position et du type (ionisé ou pas) de l'impureté sur les caractéristiques électriques du MOSFET GAA. Nous modélisons dans un premier temps une impureté localisée en début de canal au centre de la section transverse (figure (6.14)a).

La figure (6.15) représente les deux premières sous-bandes de la vallée Δ_2 (ou vallée (010)) superposées à celles d'un composant au canal intrinsèque, c'est-à-dire sans impureté. La présence de l'impureté affecte fortement les sous-bandes au voisinage du défaut et les décale vers des valeurs supérieures. Néanmoins, les sous-bandes ne s'interceptent pas et restent indépendantes, ce qui valide l'utilisation de l'approche "mode-space". La figure (6.16) compare le module au carré du premier état propre transverse avec et sans impureté ionisée. L'état est visuellement modifié, mais le produit scalaire des deux fonctions d'onde atténue cet première intuition et montre qu'ils sont en fait très similaires. Nous obtenons une moyenne un produit scalaire égale à

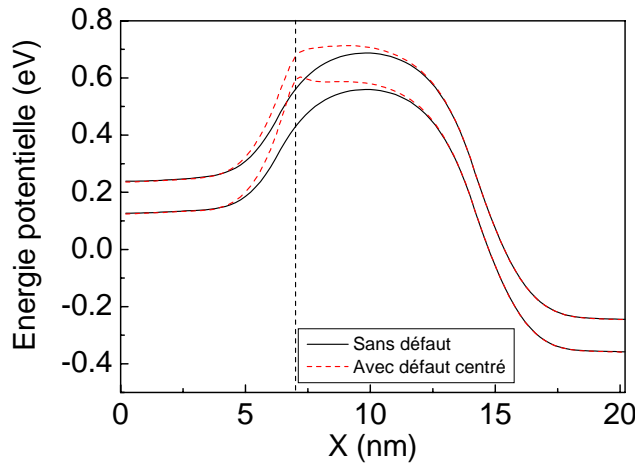


FIG. 6.15: Sous-bandes le long de l'axe source-drain du transistor sans défaut (lignes continues), et du transistor dopé avec un défaut accepteur introduit à 1 nm de la source centré dans la section transverse (lignes discontinues). La ligne verticale indique la position du défaut. $V_{DS} = 0.5$ V et $V_G = 0$ V.

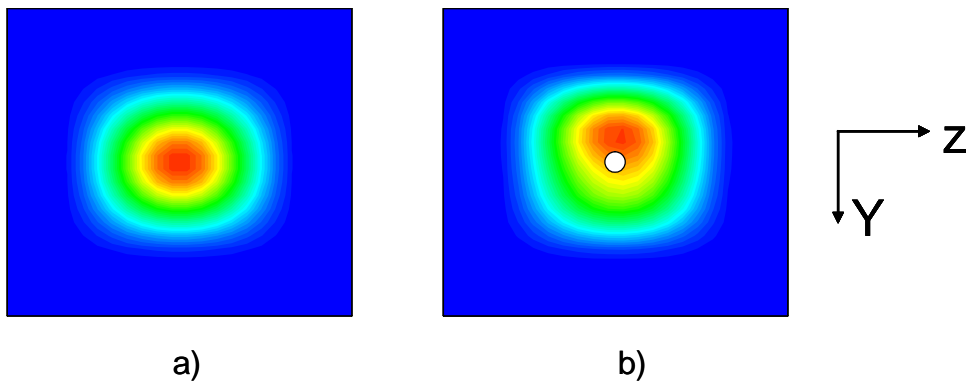


FIG. 6.16: Modules au carré des états transverse de la première sous-bande a) dans une section sans défaut; b) dans une section contenant l'impureté dopante décrite figure (6.15). Le cercle blanc localise l'impureté.

0.95 traduisant la conservation de la fonction d'onde en présence du défaut et le léger impact du potentiel coulombien dans les directions transverse fortement confinées. L'approche de l'espace des modes, dont la validité nécessite une faible variation de l'état transverse le long de la zone active, est donc confortée. De plus, la barrière de potentiel source-drain augmente et l'on peut s'attendre à une diminution des charges dans le canal. La figure (6.17) illustre cette évolution en représentant la densité électronique le long du canal au centre de la section transverse. Les charges du composant dopé sont localement inférieures à celles d'un composant intrinsèque. Plus précisément, seule la densité électronique associée à l'injection de la source est réduite, alors que celle de la composante du drain reste inchangée. Les électrons de drain n'interagissent pas avec une impureté introduite dans la première moitié du canal. Au vue des premiers effets de la présence d'une impureté ionisée sur les quantités locales, une influence significative de la position et du type de l'impureté est envisageable.

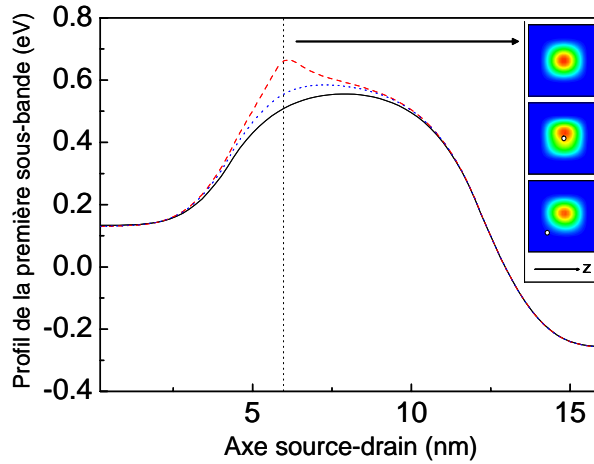


FIG. 6.18: Comparaison de la première sous-bande en fonction de la position du dopant : canal intrinsèque (ligne continue), impureté introduite à 2 nm de la source, centrée (tirets), puis excentrée (pointillés). $V_{DS} = 0.4$ V. La ligne verticale indique la position du défaut alors que la figure insérée représente les modules au carré du premier état transverse dans les trois cas. Les cercles blancs localisent l'impureté.

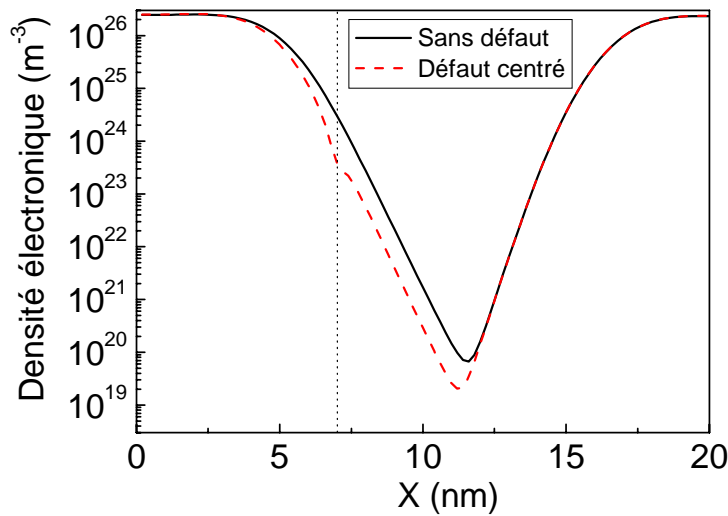


FIG. 6.17: Influence d'un défaut ponctuel sur la densité électronique le long de l'axe source-drain au centre de la section transverse. La ligne discontinue indique la position de l'impureté. $V_{DS} = 0.5$ V et $V_G = 0$ V.

La figure (6.18) représente la première sous-bande de la vallée Δ_1 pour différentes positions dans la section transverse de l'impureté introduite près de la source. Nous constatons que l'influence du défaut se fait davantage ressentir lorsque ce dernier est au centre de la section transverse. À l'opposé, une impureté située dans les coins génère une sous-bande très proche de celle du canal sans défaut. La figure insérée montre que le module au carré du premier état transverse associé à l'impureté excentrée est très peu modifié par rapport à celui de réseau parfait. Son produit scalaire avec le même vecteur propre d'une section sans défaut vaut 0.99.

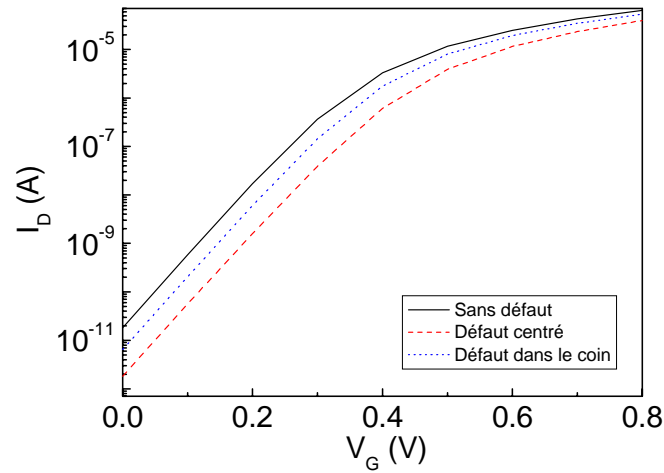


FIG. 6.19: Caractéristiques $I_D - V_G$ associées à trois configurations de dopage : sans impureté (ligne continue), impureté introduite à 2 nm de la source, centrée (tirets), puis excentrée (pointillés). $U_d = 2$ eV, et $V_{DS} = 0.4$ V.

Les conséquences sur les caractéristiques de courant $I_D - V_G$ correspondantes sont représentées sur la figure (6.19). Le courant du canal sans impureté est le plus élevé dans tous les régimes (sous le seuil et saturé) alors que celui obtenu avec un défaut localisé au centre de la section transverse est minimum. Entre ces deux extrema, la présence d'un défaut dans les contours du fil de silicium est moins influente en termes de réduction de courant : le courant sous le seuil atteint la moitié de celui d'un réseau parfait. De plus, nous pouvons noter que la différence entre les courbes n'est pas un décalage global de la caractéristique, mais plutôt une variation de la pente sous le seuil, c'est-à-dire de la transconductance. En somme, modifier la position transverse de l'impureté induit un changement des propriétés de transport qui se répercute directement sur le courant de drain. La figure (6.20) compare la caractéristique d'une impureté neutre (sans potentiel coulombien) à celle d'une impureté ionisée. On observe que la variation du courant en présence d'une impureté neutre est négligeable et redonne quasi-fidèlement le courant de drain d'un composant sans défaut. Une impureté neutre ne suffit donc pas à bloquer un mode de conduction (ou une sous-bande). Le mode de transport d'un composant 3D n'est pas affecté par ce type de défaut isolé, que les électrons "contournent" par les atomes voisins non perturbés.

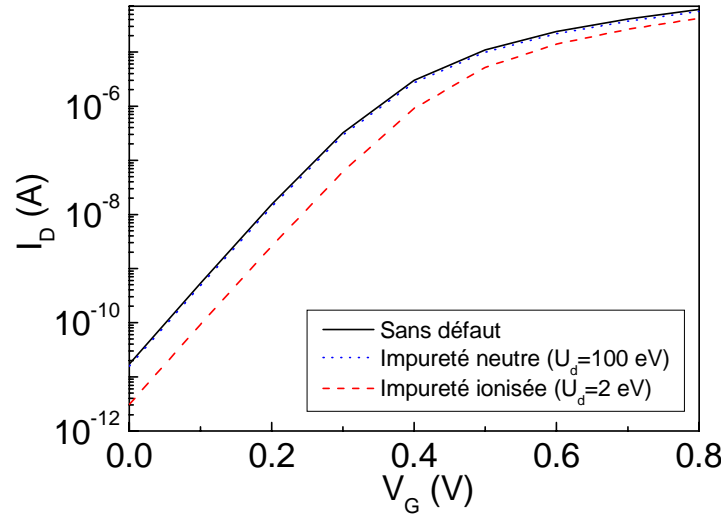


FIG. 6.20: Caractéristiques $I_D - V_G$ associées à trois configurations de dopage : sans impureté (ligne continue), impureté introduite à 2 nm de la source, neutre (pointillés), impureté ionisée (pointillés). Impureté neutre : $U_d = 100$ eV, impureté ionisée : $U_d = 2$ eV. $V_{DS} = 0.4$ V.

6.5 Conclusion

A l'aide d'une simulation en liaisons fortes exprimée dans le formalisme des fonctions de GREEN, nous avons développé un modèle pour évaluer les effets d'un dopage discret du canal de MOSFETs nanométriques. L'effet du type et de la position le long de la zone active d'une impureté neutre a d'abord été étudié dans le transistor double-grille à canal unique de conduction. Nous avons détaillé les différents processus de transport électronique et illustré le cas particulier du transport par effet tunnel assisté. Ce modèle a ensuite été étendu au cas du défaut accepteur introduit dans le canal d'un transistor MOSFET *gate-all-around*. L'influence de quatre⁴ positions significatives de l'impureté a été analysé dans un canal de 8 nm et une section transversale de 9 nm² en termes de quantités physiques (densité électronique, états quantiques transverses, sous-bandes électroniques) et de courant de drain. La variation de position de l'impureté ionisée modifie substantiellement les propriétés de transport qui se traduit par des modifications de la transconductance. En conséquence, le courant de drain peut subir une diminution de plus de 70% en fonction de la position de l'impureté ionisée. Une impureté introduite près de la source au centre de la section transversale dégrade très fortement les caractéristiques, alors que le même défaut excentré dans les coins du fil de silicium a peu d'impact. De plus, les résultats des simulations ont montré qu'une impureté neutre (sans potentiel coulombien) n'avait pas ou très peu d'influence sur le courant total : un défaut neutre isolé ne détériore pas les propriétés d'un transistor de taille nanométrique. Ce travail s'est limité à l'étude d'impuretés dans le canal. Le présent traitement peut sans modification s'appliquer également à l'étude de défauts donneurs dans les réservoirs de source et de drain. Enfin, une autre voie de recherche serait de simuler

⁴Calculs pour les deux positions près du drain en cours d'interprétation.

l'influence de tels défauts ponctuels (donneurs et accepteurs) dans les architectures émergentes simulées au chapitre 5. Ces travaux sont en cours de développement.

Conclusion

Le présent manuscrit fait état d'une étude théorique approfondie des principaux effets quantiques présents dans les dispositifs MOSFETs des générations futures. Les objectifs de cette étude étaient les suivants :

- d'une part de développer les méthodes physiques susceptibles de décrire le transport électronique de transistors nanométriques,
- d'autre part de clarifier les phénomènes physiques sous-jacents,
- d'étudier l'impact de la miniaturisation sur les caractéristiques électriques des dispositifs MOSFETs émergents et plus particulièrement sur les architectures multi-grilles.

Le premier chapitre de ce manuscrit fait la synthèse des principaux dispositifs pressentis pour poursuivre l'élan d'intégration et détaille les aptitudes et les limites de chacun. Un accent particulier a été mis sur le cas prometteur du transistor à nano-tube de carbone. Nous espérons qu'une telle synthèse, bien que connue pour certains, servira comme point de départ à d'autres travaux dans le milieu de la physique des composants.

La description des méthodes de physique théorique a fait l'objet de la seconde partie. Nous avons présenté les différentes approches de calcul de bandes d'énergie ainsi que leur application au transport quantique. En les comparant aux approches classiques (équation de transport de BOLTZMANN), nous avons clairement illustré l'intérêt du formalisme des fonctions de GREEN hors-équilibre pour traiter des systèmes ouverts tels que les MOSFETs. Ce formalisme permet à travers le concept de self-énergie de se focaliser sur le système étudié (zone active) et de modéliser les grands réservoirs d'électrons sans augmenter la taille du problème. Nous avons également décrit l'influence de la résistance des réservoirs sur le transport de systèmes nanoscopiques et par là même, l'importance d'utiliser un modèle auto-cohérent.

Le chapitre 4 applique le formalisme des fonctions de GREEN hors-équilibre au MOSFET double-grille ultime dans lequel l'axe source-canal-drain est représenté par une simple chaîne atomique décrite en liaisons fortes. Ce modèle montre qu'un confinement vertical ultime permet de conserver les propriétés électriques des MOSFETs conventionnels. Nous avons ensuite étudié en détail l'influence de l'effet tunnel source-drain sur les caractéristiques électriques en fonction de la longueur de grille et de son travail de sortie. Les simulations ont indiqué que les performances restent tolérables (en termes de pente sous le seuil et de courant *off*) jusqu'à des longueurs de grille de 7 nm. D'autre part, et contrairement à l'idée reçue, l'amplitude du courant tunnel n'augmente pas systématiquement avec la réduction de la longueur de grille. Il faut à chaque dimension ajuster le travail de sortie de la grille pour conserver le même courant *off*.

L'étape suivante a été d'étendre le même formalisme (fonctions de GREEN) à des dispositifs plus réalistes 3D. Dans le chapitre 5, nous nous sommes donc intéressés aux architectures

MOSFETs émergentes, *i.e.* le *Tri-gate*, le *Pi-gate*, l'*Omega-gate*, et le *Gate-all-around*. Pour ce faire, nous avons utilisé l'approximation de l'espace des modes qui consiste à séparer l'équation de SCHRÖDINGER 3D en une équation 2D traitant le confinement transverse et une équation 1D décrivant le transport quantique suivant l'axe source-drain. Le système initialement 3D se décompose ainsi en plusieurs problèmes 1D traitables par un simple PC. Cette approche nous a permis de combiner l'influence du confinement transverse à celle de l'effet tunnel source-drain. Nous avons aussi illustré numériquement et expliqué l'impact des résistances des réservoirs sur le potentiel électrostatique de la zone active dans les composants balistiques. Les quatre architectures précédemment citées ont ensuite été comparées. Pour des sections transverses assez larges (≥ 4 nm), une amélioration des performances électriques est visible depuis le MOSFET *Tri-gate* jusqu'au *Gate-all-around* en passant par le *Pi-gate* et le *Omega-gate*. Plus précisément, les simulations ont montré que les caractéristiques du *Pi-gate* restent très voisines de celles du *Tri-gate* (même pour des pénétrations verticales significatives des grilles) alors que le MOSFET *Omega-gate* présente un comportement très proche de celui du *Gate-all-around*. Pour de plus faibles sections en revanche, la différence entre configurations de grille s'atténue et l'effet du confinement quantique écrante, en partie au moins, l'avantage électrostatique du *Gate-all-around* par rapport au *Tri-gate*. Toutes les architectures présentent alors un comportement sensiblement identique pour des sections inférieures à 2 nm.

A ce niveau d'intégration, le dopage des dispositifs n'est plus homogène. Le nombre de défauts ponctuels et leur localisation dans le composant jouent un rôle essentiel. Nous avons donc développé dans le chapitre 6 un traitement de défauts ponctuels dans les transistors multi-grilles. Un premier modèle a été élaboré dans le cas du transistor double-grille ultimement confiné du chapitre 4. Nous avons à ce stade étudié, l'influence du type et de la localisation d'un défaut ponctuel neutre le long du canal sur le courant total. Il résulte des simulations qu'un défaut ponctuel neutre induit globalement une diminution du courant. Une exception a cependant été rapportée dans le cas particulier d'un défaut donneur introduit dans la première moitié du canal avec un potentiel intra-atomique bien défini. Un second modèle fut ensuite développé pour décrire les impuretés ionisées dans les transistors 3D réalistes du chapitre 5. Le traitement auto-cohérent sépare la partie coulombienne du défaut de son potentiel intra-atomique. Une étude approfondie a également été menée en fonction du type de défaut et de sa position dans le composant. L'introduction d'un défaut centré dans la section transverse dégrade beaucoup plus fortement le courant qu'un même défaut localisé sur les contours du film. De plus, l'influence d'un défaut neutre (sans potentiel coulombien) sur le système est négligeable.

Ce travail a finalement permis d'établir une vue globale du transport quantique dans les transistors ultimes et de pointer les architectures les plus performantes. Néanmoins, la méthode utilisée pour décrire les matériaux (silicium et oxyde de silicium) est une version simplifiée des liaisons fortes, équivalente à l'approximation de la masse effective. Une modélisation atomique plus rigoureuse, notamment basée sur un modèle de liaisons fortes type SP³, permettrait de mesurer l'influence de la non-parabolicité des bandes d'énergie sous de fortes tensions. De plus, nous avons considéré des réservoirs parfaits, c'est-à-dire à l'équilibre thermodynamique et qui

n'introduisaient pas de réflexion de contacts. Une étude pertinente serait de vérifier l'évolution des caractéristiques des transistors en fonction de la probabilité d'injection des électrons dans la zone active, *i.e.* en fonction de la forme des réservoirs. Enfin, au vue des dernières publications, il semblerait qu'un transistor nanométrique ne soit pas totalement balistique. Aussi, l'intégration dans les modèles précédents d'un traitement de diffusion des électrons par les phonons constituerait une thématique à explorer.

Bibliographie

- [1] G.E. Moore, "Trends in silicon device technology," *IEDM Tech. Digest*, p.12 (1968).
- [2] G.E. Moore, "Progress in digital integrated electronics," *IEDM Tech. Digest*, p.11 (1975).
- [3] "<http://public.itrs.net>," (2003).
- [4] L.V. Keldysh, "Quantum transport equations for high electric fields," *Sov. Phys. -JETP* **20**, 1018 (1965).
- [5] J. Rammer, and H. Smith, "Quantum field-theoretical methods in transport theory of metals," *Rev. Mod. Phys.* **62**, 323 (1986).
- [6] D.K. Ferry, L.A. Akers, and E.W. Greeneich, "Ultra large scale integrated microelectronics," (Prentice-Hall, Englewood Cliffs, NJ, 1988).
- [7] W.K. Henson, N. Yang, S. Kubicek, E.M. Vogel, J.J. Wortman, K.D. Meyer, and A. Naem, "Analysis of off-leakage currents and impact on off-state power consumption for CMOS technology in the 100 nm regime," *IEEE Trans. Electron Dev.* **47**, 1393 (2000).
- [8] S. Barraud, P. Dollfus, S. Galdin, and P. Hesto, "Short-range and long-range Coulomb interactions for 3D Monte Carlo device simulation with discrete impurity distribution," *Solid-State Electron.* **46**, 1061 (2002).
- [9] D. Goguenheim, C. Trapes, and A. Bravaix, "Comparison of degradation modes in 1.2-2.1 nm thick SiO₂ oxides submitted to uniform and hot carrier injections in NMOSFETS," *J. Non-Cryst. Solids* **322**, 183 (2003).
- [10] S. Miyazaki, M. Narasaki, M. Ogasawara, and M. Hirose, "Chemical and electronic structure of ultrathin zirconium oxide films on silicon as determined by photoelectron spectroscopy," *Solid-State Electron.* **46**, 1679 (2002).
- [11] Y. Taur, and T. Ning, "Fundamentals of VLSI Devices" (Cambridge University Press Cambridge, UK, 1998).
- [12] J.B. Jacob, and D. Antoniadis, "Channel profile engineering for MOSFETs with 100 nm channels lengths," *IEEE Trans. Electron Dev.* **42**, 870 (1995).
- [13] K. Rim, J.L. Hoyt, and J.F. Gibbons, "Fabrication and analysis of deep submicron strained-Si n-MOSFET's," *IEEE Trans. Electron Dev.* **47**, 1406 (2000).
- [14] F. Payet, N. Cavassilas, and J.L. Autran, "Simulation of hole phonon-velocity in strained Si/SiGe metal-oxide-semiconductor transistor," *J. Appl. Phys.* **95**, 717 (2004).
- [15] N. Cavassilas, M. Bescond, F. Payet, and J.L. Autran, "Simulation of a free SiGe strained-Si double-gate nano-transistor," submitted to *J. Appl. Phys.* (2004).

- [16] T.A. Langdo *et al.*, "Preparation of novel SiGe-free strained Si on insulator substrates," *Int. SOI Conf. Proc.*, p.211 (2002).
- [17] A.J. Auberton-Hervé, "SOI : materials to systems," *IEDM Tech. Digest*, p.6, (1996).
- [18] L.P. Allen, W. Skinner, and A. Cate, "Status and technology of SOI substrate material," *IEEE International SOI Conf. Proc.*, p.5 (2001).
- [19] D. Munteanu, "Modélisation et caractérisation des transistors SOI : du pseudo-MOSFET au MOSFET submicronique ultra-mince," *Thèse INPG* (1999).
- [20] G. Reichert, "Etude en haute température des transistors MOS submicroniques fabriqués sur silicium sur isolant," *Thèse INPG* (1998).
- [21] S. Wind, D. Frank, and H. Wong, "Scaling silicon MOS device to their limits," *Microelectronics Engg.* **32**, 271 (1996).
- [22] R.H. Yan, A. Ourmazd, and K.F. Lee, "Scaling the Si MOSFET : From bulk to SOI to bulk," *IEEE Trans. Electron Dev.* **39**, 1704 (1992).
- [23] K. Suzuki, T. Tanaka, Y. Tosaka, H. Horie, and Y. Arimoto, "Scaling theory for double-gate SOI MOSFETs," *IEEE Trans. Electron Dev.* **40**, 2326 (1993).
- [24] T. Tanaka, K. Suzuki, H. Horie, and T. Sugii, "Ultrafast operation of V_{TH} -adjusted p^+-n^+ double-gate SOI MOSFET," *IEEE Electron Dev. Lett.* **15**, 386 (1994).
- [25] T. Sekigawa, and Y. Hayashi, "Calculated threshold voltage characteristics of an X MOS transistor having an additional bottom gate," *Solid-State electron.* **27**, 827 (1984).
- [26] D. Hisamoto, T. Kaga, Y. Kawamoto, and E. Takeda, "A fully depleted lean-channel transistor (DELTA)- a novel vertical ultra thin SOI MOSFET," *IEDM Tech. Digest.*, p.833 (1989).
- [27] H.S.P. Wong, D.J. Frank, Y. Taur, and J.M. C. Stork, "Design and performance considerations for sub-0.1 μm double-gate SOI MOSFETs," *IEDM Tech. Digest*, p.747 (1994).
- [28] X. Huang, W.C. Lee, C. Kuo, D. Hisamoto, L. Chiang, J. Kedzierski *et al.*, "Sub 50-nm FinFET : PMOS," *IEDM Tech. Digest*, p.67 (1999).
- [29] G. Pei, J. Kedzierski, P. Oldiges, M. Jeong, and E.C.C. Kan, "FinFET design considerations based on 3-D simulation and analytical modeling," *IEEE Trans. Electron Dev.* **49**, 1411 (2002).
- [30] J.T. Park, and J.P. Colinge, "Multiple-gate SOI MOSFETs : device design guidelines," *IEEE Trans. Electron Dev.* **49**, 2222 (2002).
- [31] J.T. Park, J.P. Colinge, "Pi-Gate SOI MOSFETs," *IEEE Electron Dev. Lett.* **22**, 405 (2001).
- [32] J.P. Colinge, "Multiple-Gate SOI MOSFETs," *Solid-State Electron.*, **48**, 897 (2004).
- [33] S. Miyano, M. Hirose, and F. Masuoka, "Numerical analysis of a cylindrical thin-pillar transistor (CYNTHIA)," *IEEE Trans. Electron Dev.* **39**, 1876 (1992).
- [34] A. Nitayama, H. Takato, N. Okabe, K. Sunouchi, K. Hieda, F. Origuchi and F. Masuoka, "Multi-pillar surrounding gate transistor (M-SGT) for compact and high-speed circuits," *IEEE Trans. Electron Dev.* **38**, 579 (1991).

-
- [35] J.P. Colinge, J.W. Park, and W. Xiong, "Threshold voltage and subthreshold slope of multiple-gate SOI MOSFETs," *IEEE Electron Dev. Lett.* **24**, 515 (2003).
- [36] S. Cristoloveanu, B.J. Blalock, F. Allibert, B.M. Dufrene, and M.M. Mojarradi, "The four-gate transistor," *Proc. ESSDERC*, p.323 (2002).
- [37] K. Akarvardar, B. Dufrene, S. Cristoloveanu, B.J. Blalock, T. Higashino, M.M. Mojarradi, and E. Kolawa, "Multi-bias dependence of threshold voltage subthreshold swing, and mobility in G⁴-FETs," *Proc. ESSDERC*, p.127 (2003).
- [38] H.C. Lin, M.F. Wang, F.J. Hou, H.N. Lin, C.Y. Lu, J.T. Liu, and T.Y. Huang, "High-Performance P-channel Schottky-barrier SOI FinFET featuring self-aligned PtSi source/drain and electrical junctions," *IEEE Electron Dev. Lett.* **24**, 102 (2003).
- [39] H. Noda, F. Murai, and S. Kimura, "Short channel characteristics of Si MOSFET with extremely shallow source and drain regions formed by inversion layers," *IEEE Trans. Electron Dev.* **41**, 1831 (1994).
- [40] J. Kedzierski *et al.*, "Metal-gate FinFET and fully-depleted SOI devices using total gate silicidation," *IEDM Tech. Digest*, p.247 (2002).
- [41] Z. Ren, R. Venugopal, S. Datta, and M. Lundstrom, "Examination of design and manufacturing issues in a 10 nm double-gate MOSFET using nonequilibrium Green's function simulation," *IEDM Tech. Digest*, p.107 (2001).
- [42] A. Asenov, "Random dopant induced threshold voltage lowering and fluctuations in sub-0.1 μm MOSFETs : a 3D 'atomistic' simulation study," *IEEE Trans. Electron Dev.* **45**, 2505 (1998).
- [43] A. Asenov, and A. Saini, "Suppression of random dopant-induced threshold voltage fluctuations in sub-0.1 μm MOSFETs with epitaxial and δ -doped channels," *IEEE Trans. Electron Dev.* **46**, 1718 (1999).
- [44] H. Majima, H. Ishikuro, and T. Hiramoto, "Threshold voltage increase by quantum mechanical narrow channel effect in ultra-narrow MOSFETs," *IEDM Tech. Digest*, p.379 (1999).
- [45] J.P. Colinge, X. Baie, V. Bayot, and E. Grivei, "A silicon-on-insulator quantum wire," *Solid-State Electron.* **39**, 49 (1996).
- [46] B. Yu *et al.*, "FinFET scaling to 10 nm gate length," *IEDM Tech. Digest*, p.251 (2002).
- [47] F.L. Yang *et al.*, "25 nm CMOS Omega FETs," *IEDM Tech. Digest*, p.255 (2002).
- [48] B.H. Lee *et al.*, "Performance enhancement on sub-70 nm strained silicon SOI MOSFETs on ultra-thin thermally mixed strained silicon/SiGe on insulator (TM-SGOI) substrate with raised S/D," *IEDM Tech. Digest*, p.946 (2002).
- [49] Samuel Harrison *et al.*, "Highly performant double gate MOSFET realized with SON process," *IEDM Tech. Digest*, p.449 (2003).
- [50] H.W. Kroto, J.R. Heath, S.C. O'Brien, R.F. Curl, and R.E. Smalley, "C₆₀ : Buckminsterfullerene," *Nature (London)* **318**, 162 (1985).
- [51] T.W. Odom, J.L. Huang, P. Kim, C.M. Lieber, "Atomic structure and electronic properties of single-walled carbon nanotubes," *Nature (London)* **391**, 62 (1998).

-
- [52] S. Iijima, "Helical microtubules of graphitic carbon," *Nature (London)* **354**, 56 (1991).
- [53] S. Iijima and T. Ichihashi, "Single-shell carbon nanotubes of 1-nm diameter," *Nature (London)* **363**, 603 (1993).
- [54] A. Thess, R. Lee, P. Nikolaev *et al.*, "Cristalline ropes of metallic carbon nanotubes," *Science* **273**, 483 (1996).
- [55] P. Nikolaev, M. Bronikowski, R.K. Bradley, F. Rohmund, D.T. Colbert, K.A. Smith, and R.E. Smalley, "Gas-phase catalytic growth of single-walled carbon nanotubes from carbon monoxide," *Chem. Phys. Lett.* **313**, 91 (1999).
- [56] O. Smiljanic, B.L. Stansfield, J.P. Dodelet, A. Serventi and S. Désilets, "Gas-phase synthesis of SWNT by an atmospheric pressure plasma jet," *Chem. Phys. Lett.* **356**, 189 (2002).
- [57] N. Hamada, S.I Sawada, and A. Oshiyama, "New one-dimensional conductors : graphitic microtubules," *Phys. Rev. Lett.* **68**, 1579 (1992).
- [58] J.W.G. Wildoer, L.C. Venema, A.G. Rinzler, R.E. Smalley, and C. Dekker, "Electronic structure of atomically resolved carbon nanotubes," *Nature (London)* **391**, 59 (1998).
- [59] P.G. Collins, and P. Avouris, "Nanotubes for electronics," *Scientific American*, 62 (december 2000).
- [60] S. Frank, P. Poncharal, Z.L. Wang, and W.A. de Heer, "Carbon nanotube quantum resistors," *Science* **280**, 1744 (1998).
- [61] Y. Imry, and R. Landauer, "Conductance viewed as transmission," *Rev. Mod. Phys.* **71**, S306 (1999).
- [62] S. Sanvito, Y.K. Kwon, D. Tománek, and C.J. Lambert, "Fractional quantum conductance in carbon nanotubes," *Phys. Rev. Lett.* **84**, 1974 (2000).
- [63] M.P. Anantram, "Current-carrying capacity of carbon nanotubes," *Phys. Rev. B* **62**, R4837 (2000).
- [64] M.P. Anantram, L. Delzeit, A. Cassell, J. Han, and M. Meyyappan, "Nanotubes in nanoelectronics : transport, growth and modeling," *Physica E* **11**, 118 (2001).
- [65] J.Y. Yi, and J. Bernholc, "Atomic structure and doping of microtubules," *Phys. Rev. B* **47**, 1708 (1993).
- [66] Y. Miyamoto, A. Rubio, X. Blase, M.L. Cohen, and S. Louie, "Ionic cohesion and electron doping of thin carbon tubules with alkali atoms," *Phys. Rev. Lett.* **74**, 2993 (1995).
- [67] J. Appenzeller, J. Knoch, V. Derycke, R. Martel, S. Wind, and Ph. Avouris, "Field-modulated carrier transport in carbon nanotube transistors," *Phys. rev. Lett.* **89**, 126801 (2002).
- [68] J. Appenzeller, J. Knoch, R. Martel, V. Derycke, S. Wind, and Ph. Avouris, "Short-channel like effects in Schottky barrier carbon nanotube field-effect-transistors," *IEDM Tech. Digest* p.285 (2002).
- [69] J. Guo, S. Datta, and M. Lundstrom, "A numerical study of scaling issues for Schottky barrier carbon nanotube transistors," *IEEE Trans. Electron Dev.* **51**, 172 (2004).
- [70] S. Tans, A. Verschueren, and C. Dekker, "Room temperature transistor based on a single carbon nanotube," *Nature (London)* **393**, 49 (1998).

-
- [71] R. Martel, "Single- and multi-wall carbon nanotube field-effect transistors," *Appl. Phys. Lett.* **73**, 2447 (1998).
- [72] X. Liu, C. Lee, and C. Zhou, "Carbon nanotube field-effect-inverters," *Appl. Phys. Lett.* **79**, 3329 (2001).
- [73] A. Bachtold, P. Hadley, T. Nakashini, and C. Dekker, "Logic circuits with carbon nanotube transistors," *Science* **294**, 1317 (2001).
- [74] S.J. Wind, J. Appenzeller, R. Martel, V. Derycke, and Ph. Avouris, "Vertical scaling of carbon nanotube field-effect-transistors using the top gate electrodes," *Appl. Phys. Lett.* **80**, 3817 (2002).
- [75] "IBM research". IBM. <http://www.research.ibm.com/>.
- [76] www.cs.virginia.edu/brochure/museum.html
- [77] A. Javey, J. Guo, Q. Wang, M. Lundstrom, and H. Dai, "Ballistic carbon nanotube field-effect-transistor," *Nature (London)* **424**, 654 (2003).
- [78] A. Javey, J. Guo, M. Paulsson, Q. Wang, D. Mann, M. Lundstrom, and H. Dai, "High-field quasiballistic transport in short carbon nanotubes," *Phys. Rev. Lett.* **92**, 106804 (2004).
- [79] M. Radosavljević, S. Heinze, J. Tersoff, and Ph. Avouris, "Drain voltage scaling in carbon nanotube transistors," *Appl. Phys. Lett.* **83**, 2435 (2003).
- [80] M. Menon, A.N. Andriotis, D. Srivastava, I. Ponomareva, and L.A. Chernozatonskii, "Carbon nanotube "T junctions" : formation pathways and conductivity," *Phys. Rev. Lett.* **91**, 145501 (2003).
- [81] M. Terrones, F. Banhart, N. Grobert, J.C. Charlier, H. Terrones, and P.M. Ajayan, "Molecular Junctions by Joining Single-Walled Carbon Nanotubes," *Phys. Rev. Lett.* **89**, 075505 (2002).
- [82] A. Bachtold, P. Hadley, T. Nakanishi, and C. Dekker, "Logic circuits based on carbon nanotubes," *Physica E* **16**, 42 (2003).
- [83] Y.C. Tseng, P. Xuan, A. Javey, R. Malloy, Q. Wang, J. Bokor, and H. Dai, "Monolithic integration of carbon nanotube devices with silicon MOS technology," *Nano-Lett.* **4**, 123 (2004).
- [84] M. Lundstrom, "Moore's law forever?," *Science* **299**, 210 (2003).
- [85] J.R. Tucker, and T.C. Shen, "Can single-electron integrated circuits and quantum computers be fabricated in silicon?," *Int. J. Circ. Theor. Appl.* **28**, 553 (2000).
- [86] J.D. Meindl, Q. Chen, J.A. Davis, "Limits on silicon nanoelectronics for terascale integration," *Science* **293**, 2044 (2001).
- [87] D.R. Hartree, *Proc. Cambridge Phil. Soc.* **24**, 89 (1928).
- [88] F. Seitz, "The Modern Theory of Solids" (Mc Graw-Hill, New York, 1940).
- [89] C. Cohen-Tannoudji, B. Diu, et F. Lalöe, "Mécanique quantique Tome I" (Editions Hermann).
- [90] P. Hohenberg, and W. Kohn, "Inhomogeneous electron gas," *Phys. Rev.* **136**, 864 (1964).
- [91] W. Kohn, and L.J. Sham, *Phys. Rev.* **140**, 1133 (1965).

-
- [92] O. Zakharov, A. Rubio, X. Blase, M.L. Cohen, and S.G. Louie, "Quasiparticle band structures of six II-VI compounds : ZnS, ZnSe, ZnTe, CdS, CdSe, and CdTe," *Phys. Rev. B* **50**, 10780 (1994).
- [93] X. Zhu, and S.G. Louie, "Quasiparticle band structure of thirteen semiconductors and insulators," *Phys. Rev. B* **43**, 14142 (1991).
- [94] M.L. Cohen, and T.K. Bergstresser, "Band structures and pseudopotential form factors for fourteen semiconductors of the diamond and zinc-blende structures," *Phys. Rev. B* **141**, 789 (1966).
- [95] J.C. Slater, and G.F. Koster, "Simplified LCAO method for the periodic potential problem," *Phys. Rev.* **94**, 1498 (1954).
- [96] C. Krzeminski, C. Delerue, and G. Allan, "Tight-binding description of the electronic response of a molecular device to an applied voltage," *J. Phys. Chem.* **105**, 6321 (2001).
- [97] A.J. Lu, and B.C. Pan, "Nature of single vacancy in achiral carbon nanotubes," *Phys. Rev. Lett.* **92**, 105504 (2004).
- [98] W. Ren, and J. Wang, "Electronic transport through single-wall nicked carbon nanotubes," *Phys. Rev. B* **69**, 33306 (2004).
- [99] Y.N. Niquet, C. Delerue, G. Allan, and M. Lannoo, "Method for tight-binding parametrization : Application to silicon nanostructures," *Phys. Rev. B* **62**, 5109 (2000).
- [100] M. Rhoifing, and S.G. Louie, "Excitonic effects and the optical absorption spectrum of hydrogenated Si clusters," *Phys. Rev. Lett.* **80**, 3320 (1998).
- [101] C. Delerue, "Electronic structure and electron-paramagnetic-resonance properties of intrinsic defects in GaAs," *Phys. Rev. B* **44**, 10525 (1991).
- [102] N. Cavassilas, F. Aniel, K. Boujdaria, and G. Fishman, "Energy-band structure of GaAs and Si : a *sps** k.p method," *Phys. Rev. B* **64**, 115207 (2001).
- [103] D. Munteanu and J.L. Autran, "Two-Dimensional modeling of quantum ballistic transport in ultimate double-gate SOI devices," *Solid-State Electron.* **47**, 1219 (2003).
- [104] 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).
- [105] J. Guo and M.S. Lundstrom, "A computational study of thin-body, double-gate, Schottky barrier MOSFETs," *IEEE Trans. Electron Dev.*, **49** (11), 1897 (2002).
- [106] M. Städele, "Influence of source-to-drain tunneling on the subthreshold behavior of sub-10 nm double-gate MOSFETs," *Proc. ESSDERC*, p.135 (2002).
- [107] M. Städele, A. Di Carlo, P. Lugli, F. Sacconi, and B. Tuttle, "Atomistic tight-binding calculations for the investigation of transport in extremely scaled SOI transistors," *IEDM Tech. Digest*, p.229 (2003).
- [108] G. Fiori and G. Iannaccone, "Modeling of ballistic nanoscale metal-oxide-semiconductor field-effect-transistors," *Appl. Phys. Lett.* **81** (19), 3672 (2002).
- [109] A. Pirovano, A.L. Lacaita, "Two-dimensional quantum effects in nanoscale MOSFETs," *IEEE Trans. Electron. Dev.* **49**, 25 (2002).

-
- [110] Y. Naveh and K.K. Likharev, "Modeling of 10-nm-scale ballistic MOSFETs," *IEEE Electron Dev. Lett.* **21**, 242 (2000).
- [111] J.H. Rhew, Z. Ren, and M.S. Lundstrom, "A numerical study of ballistic transport in a nanoscale MOSFET," *Solid-State Electron.* **46**, 1899 (2002).
- [112] K. Uchida, H. Watanabe, A. Kinoshita, J. Koga, T. Numata, and S.-I. Takagi, "Experimental study on carrier transport mechanism in ultrathin-body SOI n- and p-MOSFETs with SOI thickness less than 5 nm," *IEDM Tech. Digest*, p.47 (2000).
- [113] K. Uchida, K. Matsuzawa, J. Koga, S.I. Takagi, and A. Toriumi, "Enhancement of the hot-electron generation rate in Schottky source metal-oxide-semiconductor field-effect-transistors," *Appl. Phys. Lett.* **76**, 3992 (2000).
- [114] T. Mukaiyama, K.I. Saito, H. Ishikuro, M. Takamiya, T. Saraya, and T. Hiramoto, "Fabrication of gate-all-around MOSFET by silicon anisotropic etching technique," *Solid-State Electron.* **42**, 1623 (1998).
- [115] R.P. Andres, J.D. Bielefeld, J.I. Henderson, D.B. Janes, V.R. Kolagunta, C.P. Kubiak, W.I. Wahoney, and R.G. Gsifchin, "Self-assembly of a two-dimensional superlattice of molecular linked metal clusters," *Science* **273**, 1690 (1996).
- [116] L.A. Bumm, J.J. Arnold, M.T. Cygan, T.D. Dunbar, T.P. Burgin, L. JonesII, D.L. Allara, J.M. Tour, and P.M. Weiss, "Are single molecular wires conducting?," *Science* **271**, 1705 (1996).
- [117] S.J. Tans, M.H. Devoret, H. Dai, A. Thess, R.E. Smalley, L.J. Geerlings, and C. Dekker, "Individual single-wall carbon nanotubes as quantum wires," *Nature (London)* **386**, 474 (1997).
- [118] J. Chen, M.A. Reed, A.M. Raevlett, and J.M. Tour, "Large On-Off ratios and negative differential resistance in a molecular electronic device," *Science* **286**, 1550 (1999).
- [119] B. Alpers, I. Rubinstein, G. Hodes, D. Porath, and O. Millo, "Energy level tunneling spectroscopy and single electron charging in individual CdSe quantum dots," *Appl. Phys. Lett.* **75**, 1751 (1999).
- [120] D.L. Klein, R. Roth, A.K.L. Andrew, A.P. Alivisatos, and P.L. McEuen, "A single-electron transistor made from a cadmium selenide nanocrystal," *Nature (London)* **389**, 699 (1997).
- [121] C. Delerue, and M. Lannoo, "Nanostructures, theory and modelling," Springer, collection Nanoscience and Technology (2004).
- [122] R. Landauer, "Spatial variation of currents and fields due to localized scatterers in metallic conduction," *IBM J. Res. Dev.* **1**, 223 (1957).
- [123] V.K. Khanna, "Physics of carrier-transport mechanism and ultra-small scale phenomena for theoretical modelling of nanometer MOS transistors from diffusive to ballistic regimes of operation," *Physics Reports* **398**, 67 (2004).
- [124] M.V. Fischetti, and S.E. Laux, "Monte carlo analysis of electron transport in small semiconductor devices including band-structure and space-charge effects," *Phys. Rev. B* **38**, 9721 (1988).

-
- [125] S.E. Laux, and M.V. Fischetti, "Monte-Carlo simulation of sub-micrometer Si n-MOSFET's at 77 and 300 K," *IEEE Electron Dev. Lett.* **9**, 467 (1988).
- [126] E. Sangiorgi, and M.R. Pinto, "A semi-empirical model of surface scattering for MC simulation of silicon n-MOSFET," *IEEE Trans. Electron Dev.* **39**, 356 (1992).
- [127] M.V. Fischetti, and S.E. Laux, "Monte carlo study of electron transport in silicon inversion layers," *Phys. Rev. B* **48**, 2244 (1993).
- [128] M.V. Fischetti, and S.E. Laux, "Understanding hot-electron transport in silicon devices : Is there a shortcut?," *J. Appl. Phys.* **78**, 1058 (1995).
- [129] W.J. Gross, D. Vasileska, and D.K. Ferry, "Three-dimensional simulation of ultrasmall metal-oxide-semiconductor field-effect-transistors : the role of the discrete impurities on the device terminal characteristics," *J. Appl. Phys.* **91**, 3737 (2002).
- [130] P. Dollfus, A. Bournel, S. Galdin, S. Barraud, and P. Hesto, "Effect of discrete impurities on electron transport in ultrashort MOSFET using 3-D MC Simulation," *IEEE Trans. Electron Dev.* **51**, 749 (2004).
- [131] M. Le Bellac et F. Montessagne, "Thermodynamique Statistique : équilibre et hors équilibre : cours, exercices et problème résolu," Dunod (2001).
- [132] A. Gnudi, D. Ventura, G. Baccarani, and F. Odeh, "Macroscopic and microscopic approach for the simulation of short devices," *Semiconductors, Part II*. Springer, new York (1994).
- [133] G. Baccarani, and M.R. Wordeman, "An investigation of steady-state velocity overshoot in silicon," *Solid-State Electron.* **28**, 407 (1985).
- [134] S.Y. Chou, D.A. Antoniadis, and H.I. Smith, "Observation of electron velocity overshoot in sub-100-nm-channel MOSFETs," *IEEE Electron Dev. Lett.* **6**, 665 (1985).
- [135] D. Munteanu, G.L. Lecarval, and G. Guegan, "Impact of technological parameters on non-stationary transport in realistic 50 nm MOSFET," *Solid-State Electron.* **46**, 1045 (2002).
- [136] G. Baccarani, and S. Reggiani, "A compact double-gate MOSFET model comprising quantum-mechanical and nonstatic effects," *IEEE Trans. Electron Dev.* **46**, 232 (1999).
- [137] D. Chang, and J.G. Fossum, "Simplified energy-balance model for pragmatic multi-dimensional device simulation," *Solid-State Electron.* **41**, 1795 (1997).
- [138] K. Banoo, and M.S. Lundstrom, "Electron transport in a model silicon transistor," *Solid-State Electron.* **44**, 1689 (2000).
- [139] K. Natori, "Ballistic metal-oxide-semiconductor field-effect-transistor," *J. Appl. Phys.* **76**, 4879 (1994).
- [140] M. Lundstrom, "Elementary scattering theory of the Si MOSFET," *IEEE Electron Dev. Lett.* **18**, 361 (1997).
- [141] M.S. Lundstrom, and Z. Ren, "Essential physics of carrier transport in nanoscale MOSFETs," *IEEE Trans. Electron Dev.* **49**, 133 (2002).
- [142] A. Pirovano, A.L. Lacaita, and A.S. Spinnelli, "Two-dimensional quantum effects in nanoscale MOSFETs," *IEEE Trans. Electron Dev.* **49**, 25 (2002).

-
- [143] F.G. Picus, and K.K. Likharev, "Nanoscale field-effect transistors : an ultimate size analysis," *Appl. Phys. Lett.* **71**, 3661 (1997).
- [144] D. Munteanu, and J.L. Autran, "Two-dimensional modelling of quantum ballistic transport in ultimate double-gate SOI devices," *Solid-State Electron.* **47**, 1219 (2003).
- [145] M.G. Ancona, Z. Yu, R.W. Dutton, P.J. Vande Voorde, M. Cao, and D. Vook, "Density-gradient analysis of MOS tunneling," *IEEE Trans Electron Dev.* **47**, 2310 (2000).
- [146] A. Asenov, S. Kayas, and J.H. Davies, "Intrinsic Threshold voltage fluctuations in decanano MOSFETs due to local oxide thickness variations," *IEEE Trans. Electron Dev.* **49**, 112 (2002).
- [147] A. Asenov, A.R. Brown, and J.R. Watling, "Quantum corrections in the simulation of decanano MOSFETs," *Solid-State Electron.* **47**, 1141 (2003).
- [148] A. Asenov, R. Balasubramaniam, A.R. Brown, and J.H. Davies, "RTS amplitudes in decanometer MOSFETs : 3-D simulation study," *IEEE Trans. Electron Dev.* **50**, 839 (2003).
- [149] A. Asenov, S. Kayas, and A.R. Brown, "Intrinsic parameter fluctuations in decanometer MOSFETs introduced by gate line edge roughness," *IEEE Trans. Electron Dev.* **50**, 1254 (2003).
- [150] K. Matsuzawa, S.I. Takagi, M. Takayanagi, and H. Tanimoto, "Device simulation of surface quantization effect on MOSFETs with simplified density-gradient method," *Solid-State Electron.* **46**, 747 (2002).
- [151] Z. Yu, R.W. Dutton, and R.A. Kiehl, "Circuit/device modeling at the quantum level," *IEEE Trans. Electron Dev.* **47**, 1819 (2000).
- [152] S.E. Laux, A. Kumar, and M.V. Fischetti, "Analysis of quantum ballistic electron transport in ultrasmall silicon devices including space-charge and geometric effects," *J. Appl. Phys.* **95**, 5545 (2004).
- [153] M.V. Fischetti, "Theory of electron transport in small semiconductor devices using the Pauli master equation," *J. Appl. Phys.* **83**, 270 (1998).
- [154] M.V. Fischetti, "Master-equation approach to the study of electronic transport in small semiconductor devices," *Phys. Rev. B* **59**, 4901 (1999).
- [155] A. Svizhenko, M.P. Anantram, T.R. Govindan, B. Biegel, and R. Venugopal, "Nanotransistor modeling : two-dimensional Green's function method," *J. Appl. Phys.* **91**, 2343 (2002).
- [156] A. Svizhenko, and M.P. Anantram, "Role of scattering in nanotransistors," *IEEE Trans. Electron dev.* **50**, 1459 (2003).
- [157] R. Venugopal, M. Paulsson, S. Goasguen, S. Datta, and M. Lundstrom, "A simple quantum mechanical treatment of scattering in nanoscale transistors," *J. Appl. Phys.* **93**, 5613 (2003).
- [158] R. Venugopal, S. Goasguen, S. Datta, and M.S. Lundstrom, "Quantum mechanical analysis of channel access geometry and series resistance in nanoscale transistors," *J. Appl. Phys.* **95**, 292 (2004).

- [159] M. Büttiker, "Four terminal phase coherent conductance," *Phys. Rev. Lett.* **57**, 1761 (1986).
- [160] T.S. Xia, L.F. Register, and S.K. Banerjee, "Quantum transport in double-gate MOSFETs with complex band structure," *IEEE Trans. Electron Dev.* **50**, 1511 (2003).
- [161] E.P.A.M. Bakkers, and Daniël Vanmaekelbergh, "Resonant electron tunneling through semiconducting nanocrystals in a symmetrical and an asymmetrical junction," *Phys. Rev. B* **62**, R7743 (2000).
- [162] M. Bescond, J.L. Autran, D. Muntenau, and M. Lannoo, "Atomic-scale modeling of double-gate MOSFETs using a tight-binding Green's function formalism," *Solid-State Electron.* **48**, 567 (2004).
- [163] J. Bardeen, "Tunneling from a many-particle point of view", *Phys. Rev. Lett.* **6**, 57 (1961).
- [164] W.A. Harrison, "Tunneling from an independent-particle point of view," *Phys. Rev.* **123**, 85 (1961).
- [165] J. Tersoff, and D.R. Hamann, "Theory and application for the scanning tunneling microscope," *Phys. Rev. Lett.* **50**, 1998 (1983).
- [166] J. Tersoff, "Role of tip electronic structure in scanning tunneling microscope images," *Phys. Rev. B* **41**, 1235 (1990).
- [167] I.O. Kulik and R.I. Shekhter, "Kinetic phenomena and charge discreteness effects in granular media," *Sov. Phys. JETP* **41**, 308 (1975).
- [168] D.V. Averin, A.N. Korotkov, and K.K. Likharev, "Theory of single-electron charging of quantum wells and dots," *Phys. Rev. B* **44**, 6199 (1991).
- [169] C.W.J. Beenakker, "Theory of Coulomb-blockage oscillations in the conductance of a quantum dot," *Phys. Rev. B* **44**, 1646 (1991).
- [170] S. Datta, "Electronic Transport in Mesoscopic Systems," (Cambridge University Press, Cambridge, 1997).
- [171] J. Guo, M. Lundstrom, and S. Datta, "Performance projections for ballistic carbon nanotube field-effect-transistors," *Appl. Phys. Lett.* **80**, 3192 (2002).
- [172] T. Yamada, "Analysis of submicron carbon nanotube field-effect-transistors," *Appl. Phys. Lett.* **76**, 628 (2000).
- [173] S. Datta, "Nanoscale modeling : the Green's function method," *Superlattices Microstruct.* **28**, 253 (2000).
- [174] E. bakkers, Z. Hens, A. Zunger, A. Franceschetti, L. Kouwenhoven, L. Gwevich, and D. Vanmaekelbergh, "Shell-tunneling spectroscopy of the single-particle energy levels of insulating quantum dots," *Nano Lett.* **1**, 551 (2001).
- [175] D.K. Ferry, and S.M. Goodnick, "Transport in Nanostructures," (Cambridge University Press, Cambridge, England, 1997).
- [176] R. Landauer, "Conductance determined by transmission : probes and quantized constriction resistance," *J. Phys. : Condens. Matter* **1**, 8099 (1989).
- [177] M. Büttiker, Y. Imry, R. Landauer, and S. Pinhas, "Generalised many-channel conductance formula with application to small rings," *Phys. Rev. B* **31**, 6207 (1985).

-
- [178] J. Rammer, and H. Smith, "Quantum field-theoretical methods in transport theory of metals," *Rev. Mod. Phys.* **58**, 323 (1986).
- [179] L.P. Kadanoff, and G. Baym, "Quantum Statistical Mechanics" (Benjamin, Cummings, New-York, 1962).
- [180] C. Caroli, R. Combescot, P. Nozieres, and D. Saint-James, "Direct calculation of the tunneling current," *J. Phys. C* **4**, 916 (1971).
- [181] C. Caroli, R. Combescot, D. Lederer, P. Nozieres, and D. Saint-James, "A direct calculation of the tunneling current II. Free electron description," *J. Phys. C.* **4**, 2598 (1971).
- [182] Y. Meir, and N.S. Wingreen, "Landauer formula for the current through an interacting electron region," *Phys. Rev. Lett.* **68**, 2512 (1992).
- [183] A.P. Jauho, N.S. Wingreen, and Y. Meir, "Time-dependent transport in interacting and noninteracting resonant-tunneling systems", *Phys. Rev. B* **50**, 5528 (1994).
- [184] P.L. Pernas, A. Martin-Rodero, and F. Flores, "Electrochemical-potential variations across a constriction," *Phys. Rev. B* **41**, 8553 (1990).
- [185] V. Mujica, A.E. Roitberg, and M. Ratner, "Molecular wire conductance : Electrostatic potential spatial profile," *J. Chem. Phys.* **112**, 6834 (2000).
- [186] Z. Ren, R. Venugopal, S. Goasguen, S. Datta, and M.S. Lundstrom, "nanoMOS 2.5 : A two dimensional simulator for quantum transport in double-gate MOSFETs," *IEEE Trans. Electron Dev.* **50**, 1914 (2003).
- [187] C. Kittel, "Physique de l'état solide" (édition Dunod).
- [188] B.J. Van Wees, H. Van Houten, C.W.J. Beenakker, J.G. Williamson, L.P. Kouwenhoven, D. Van der Marel, and C.T. Foxon, "Quantized conductance of point contacts in a two-dimensional electron gas," *Phys. Rev. Lett.* **60**, 848 (1988).
- [189] A. Das, and M.S. Lundstrom, "A scattering matrix approach to device simulation," *Solid-State Electron.*, **33**, 1299 (1990).
- [190] E.N. Economou "Green's Functions in Quantum Physics" (Springer-Verlag, Berlin, 1979).
- [191] B.A. Lippmann, "Ehrenfest's theorem and scattering theory," *Phys. Rev. Lett.* **15**, 11 (1965).
- [192] D.S. Fisher, and P.A. Lee, "Relation between conductivity and transmission matrix," *Phys. Rev. B* **23**, 6851 (1981).
- [193] M. Lannoo, and P. Friedel, "Atomic and Electronic Structures of Surfaces," Springer Series in Surface Sciences **16** (Springer-Verlag, Berlin, 1991).
- [194] B. Doris *et al.*, "Extreme scaling with ultra-thin Si channel MOSFETs," *IEDM Tech. Digest*, p.267 (2002).
- [195] F.L. Yang *et al.*, "5 nm-gate nanowire Fin FET," *2004 Symposium on VLSI Technology*, Honolulu, HI, June 2004.
- [196] B.S. Doyle *et al.*, "High performance fully depleted Tri-gate CMOS transistors," *Electron Dev. Lett.* **24** (4), 263 (2003).

- [197] D. Jovanovic, "Computational techniques for the nonequilibrium quantum field theory simulation of MOSFETs," *7th International Workshop on Computational Electronics*, University of Glasgow, UK (2000).
- [198] A. Svizhenko, M. Anantram, and T. Govindan, "2D quantum simulation of MOSFET using the non-equilibrium Green's function method," *7th International Workshop on Computational Electronics*, University of Glasgow, UK (2000).
- [199] Z. Ren, *Ph.D Thesis*, Purdue University (2001).
- [200] P. Damle, A. Ghosh, and S. Datta, "Nanoscale device modeling," in *Molecular Electronics*, American Scientific Publishers, ed. Mark Reed and Takhee Lee, (2003).
- [201] J. Wang, E. Polizzi, and M. Lundstrom, "A computational study of ballistic silicon nanowire transistors," *IEDM Tech. Digest*, p.695 (2003).
- [202] J. Wang, E. Polizzi, and M. Lundstrom, "A three-dimensional quantum simulation of silicon nanowire transistors with the effective-mass approximation," *J. Appl. Phys.* **96**, 2192 (2004).
- [203] M.J. Gilbert, and D.K. Ferry, "Resonant tunneling behavior and discrete dopant effects in narrow ultrashort ballistic silicon-on-insulator metal-oxide-semiconductor field-effect transistors," *J. Vacuum Science & Technology B Microelectronics and nanometer Structures* **22**, 2039 (2004).
- [204] Y. Naveh, and K.K. Likharev, "Shrinking limits of silicon MOSFETs," *Superlattices Microstruct.* **27**, 111 (2000).
- [205] R.W. Keyes, "The effect of randomness in the distribution of impurity atoms on FET thresholds," *Appl. Phys.* **8**, 251 (1975).
- [206] T. Mizuno, and J. Okamura, "Experimental study of threshold voltage fluctuations due to statistical variation of channel dopant number in MOSFETs," *IEEE Trans. Electron Dev.* **41**, 2216 (1994).

Appendice 1. Les fonctions de Green

Définitions et relations utiles

Soit le système d'équations :

$$AX = 0, \quad (\text{A1.1})$$

où A est une matrice $n \times n$ et X est un vecteur de dimensions $1 \times n$. Ce système possède des solutions quand $\det A = 0$, c'est-à-dire quand l'application linéaire associée à A n'est pas bijective et que son noyau (ensemble des vecteurs qui appliqués à A donnent le vecteur nul) ne se limite pas à la solution triviale du vecteur nul.

Etudions maintenant le cas d'états stationnaires de l'hamiltonien indépendant du temps H :

$$H |\psi\rangle = \varepsilon |\psi\rangle, \quad (\text{A1.2})$$

où $|\psi\rangle$ et ε sont respectivement les vecteurs propres et les valeurs propres de l'opérateur. L'équation (A1.2) peut se réécrire sous la forme :

$$(H - \varepsilon) |\psi\rangle = 0, \quad (\text{A1.3})$$

qui possède des solutions si $\det (H - \varepsilon)$ s'annule pour plusieurs valeurs de l'énergie. Inversement, pour ces valeurs d'énergie $(H - \varepsilon)^{-1}$ est infini, et l'on doit lui ajouter un terme complexe afin d'éviter la divergence. Nous venons ainsi de définir l'opérateur de la fonction de GREEN. Il s'agit d'une fonction de GREEN particulière (dite "retardée") qui possède une grande quantité d'informations. La définition rigoureuse de l'opérateur de la fonction de GREEN est la suivante :

$$G(\varepsilon) = \lim_{\eta \rightarrow 0^+} \frac{1}{\varepsilon - H + i\eta}. \quad (\text{A1.4})$$

Nous pouvons également exprimer G dans une base de vecteurs propres $|\psi_k\rangle$ de l'hamiltonien :

$$G(\varepsilon) = \lim_{\eta \rightarrow 0^+} \sum_k \frac{|\psi_k\rangle \langle \psi_k|}{\varepsilon - \varepsilon_k + i\eta}, \quad (\text{A1.5})$$

où ε_k est la valeur propre associée à $|\psi_k\rangle$. La fonction de GREEN, définie à une énergie donnée, possède donc deux entrées (une côté ket et une côté bra) et permet de connecter deux positions de l'espace réel $|r\rangle$ et $|r'\rangle$. La fonction de Green prend alors une forme matricielle dont l'élément s'écrit :

$$G_{r,r'}(\varepsilon) = \lim_{\eta \rightarrow 0^+} \sum_k \frac{\langle r | \psi_k \rangle \langle \psi_k | r' \rangle}{\varepsilon - \varepsilon_k + i\eta} = \lim_{\eta \rightarrow 0^+} \sum_k \frac{\psi_k(\mathbf{r}) \psi_k(\mathbf{r}')}{\varepsilon - \varepsilon_k + i\eta}, \quad (\text{A1.6})$$

où $\psi_k(\mathbf{r})$ est la fonction propre associée à $\langle r | \psi_k \rangle$. Généralement, les vecteurs propres $|\psi_k\rangle$ sont projetés sur une base d'orbitales atomiques comme en théorie des liaisons fortes.

Une fois la fonction de GREEN du système calculée nous avons accès à toutes sortes de paramètres physiques comme nous le montrons au chapitre 3. Un des paramètres auquel nous faisons appel

dans le cas d'un système à l'équilibre est la densité d'états dont la définition usuelle est la suivante :

$$N(\varepsilon) = \sum_k \delta(\varepsilon - \varepsilon_k), \quad (\text{A1.7})$$

où δ symbolise la fonction de DIRAC. Or d'après l'expression (A1.5), la partie imaginaire de la trace de $G(\varepsilon)$ donne :

$$\text{Im}(TrG(\varepsilon)) = - \sum_k \lim_{\eta \rightarrow 0^+} \frac{\eta}{(\varepsilon - \varepsilon_k)^2 + \eta^2}. \quad (\text{A1.8})$$

Et,

$$\int_{-\infty}^{+\infty} \frac{\eta}{(\varepsilon - \varepsilon_k)^2 + \eta^2} d\varepsilon = \left[\text{Arc tan} \left(\frac{\varepsilon - \varepsilon_k}{\eta} \right) \right]_{-\infty}^{+\infty} = \pi, \quad (\text{6.48})$$

De plus, $\frac{\eta}{(\varepsilon - \varepsilon_k)^2 + \eta^2}$ est infinie si $\varepsilon = \varepsilon_k$ et nulle autrement, de sorte que :

$$\lim_{\eta \rightarrow 0^+} \frac{\eta}{(\varepsilon - \varepsilon_k)^2 + \eta^2} = \pi \times \delta(\varepsilon - \varepsilon_k). \quad (\text{6.49})$$

Finalement, nous obtenons la relation suivante :

$$N(\varepsilon) = -\frac{1}{\pi} \text{Im} \{Tr(G(\varepsilon))\}. \quad (\text{6.50})$$

De même, la densité locale d'état à un point \mathbf{r} est définie par :

$$N(\mathbf{r}, \varepsilon) = \sum_k |\psi_k(\mathbf{r})|^2 \delta(\varepsilon - \varepsilon_k), \quad (\text{A1.9})$$

qui peut s'exprimer à partir de l'équation (A1.6) sous la forme :

$$N(\mathbf{r}, \varepsilon) = -\frac{1}{\pi} \text{Im} \{G_{r,r}(\varepsilon)\}. \quad (\text{A1.10})$$

Nous voyons donc que la connaissance de G permet la détermination immédiate de la densité d'état, grandeur essentielle au le formalisme de transport des électrons et de manière générale à tout problème relatif à la matière condensée.

De la fonction de Green à la fonction d'onde

Nous avons démontré au paragraphe (4.2) l'expression suivante de la fonction de GREEN de surface d'une chaîne atomique semi-infinie au paramètre de maille a (figure (6.21)) :

$$g_{00}(\varepsilon) = \frac{1}{2\beta^2} \left(\varepsilon - i\sqrt{4\beta^2 - \varepsilon^2} \right), \quad \text{si } |\varepsilon| < 2|\beta|, \quad (\text{6.51})$$

où β représente la constante de couplage et $\varepsilon = 2\beta \cos(ka)$. Il en résulte une nouvelle expression de la fonction de GREEN de surface :

$$g_{00}(\varepsilon) = \frac{1}{2\beta^2} (2\beta \cos(ka) - i2|\beta| \sin(ka)), \quad (\text{6.52})$$

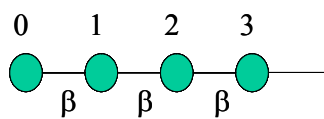


FIG. 6.21: *Chaîne atomique semi-infinie avec une constante de couplage entre premiers voisins β .*

qui donne finalement :

$$g_{00}(\varepsilon) = \frac{1}{\beta} e^{ika}, \quad (6.53)$$

et qui correspond à injecter une onde plane dans la zone active.

Appendice 2. Equivalence entre les deux formules de courant

Nous avons démontré dans le chapitre 3 que le formalisme des fonctions de GREEN pouvait conduire à deux expressions différentes du courant. La première formule résulte d'une division du dispositif en 2 régions quelconques (figure(6.22)) et s'exprime à l'aide de la fonction de GREEN retardée du système découplé G_0 :

$$I = -\frac{2\pi e}{\hbar} \int \{f(\varepsilon - \mu_G) - f(\varepsilon - \mu_D)\} Tr_1(n_{11}\Lambda_{11}^\dagger V_{12}n_{22}V_{21}\Lambda_{11})d\varepsilon, \quad (A2.1)$$

où Tr_1 représente la trace restreinte à la région 1, V est la matrice de couplage, $\Lambda = (I - G_0VG_0V)^{-1}$ et $G_0 - G_0^* = -2\pi i n(\varepsilon)$.

La seconde formule consiste à utiliser l'architecture du dispositif en divisant le système en 3 parties prédéfinies : les 2 réservoirs et la zone active. Son expression fait alors intervenir la fonction de GREEN retardée du système couplé G :

$$I = -\frac{e}{\hbar} \int Tr \left[\Gamma^G G \Gamma^D G^\dagger \right] \{f(\varepsilon_i - \mu_G) - f(\varepsilon_i - \mu_D)\} d\varepsilon, \quad (A2.2)$$

où

$$\begin{aligned} \Gamma_{n,m}^G(\varepsilon) &= 2\pi \sum_{i \in G} V_{ni} V_{im} \delta(\varepsilon - \varepsilon_i), \\ \Gamma_{n,m}^D(\varepsilon) &= 2\pi \sum_{j \in D} V_{nj} V_{jm} \delta(\varepsilon - \varepsilon_j). \end{aligned} \quad (A2.3)$$

Dans les équations (A2.3) les indices n et m réfèrent aux états propres ($|n\rangle$ et $|m\rangle$) de H_0 dans la zone active découplée(région M de la figure (6.22)b). Nous nous proposons dans cet appendice, de prouver l'équivalence formelle entre les équations (A2.1) et (A2.2). Partons de l'équation (A2.1). Cette dernière peut se réécrire sous la forme :

$$I = -\frac{2\pi e}{\hbar} \int d\varepsilon \{f(\varepsilon - \mu_G) - f(\varepsilon - \mu_D)\} \sum_{i \in 1, j \in 2} n_{ii} n_{jj} \left| (V\Lambda)_{ji} \right|^2, \quad (A2.4)$$

où n_{ii} représente la moyenne de l'opérateur n lorsque le système est préparé dans l'état $|i\rangle$:

$$n_{ii} = \langle i | n | i \rangle. \quad (A2.5)$$

A l'aide de l'équation de DYSON, nous pouvons montrer que :

$$(V\Lambda)_{ji} = (G_0^{-1})_{jj} G_{ji} (G_0^{-1})_{ii}. \quad (A2.6)$$

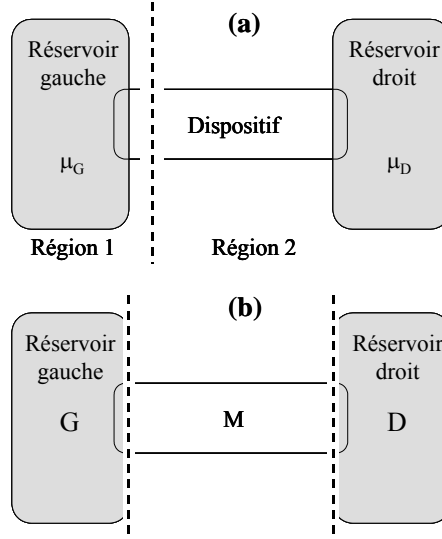


FIG. 6.22: a) Première expression de la formule de LANDAUER exprimée dans un formalisme de fonctions de GREEN : le système est divisé en 2 régions. b) Seconde expression de la formule de LANDAUER. Utilisant la structure du nano-composant, le device est divisé en 3 parties prédéfinies : les 2 réservoirs et la zone active.

En appelant K la somme discrète de l'équation (A2.4), on obtient :

$$K = \sum_{i \in 1, j \in 2} n_{ii} n_{jj} (G_0^{-1})_{jj} G_{ji} (G_0^{-1})_{ii} (G_0^{-1})_{jj}^\dagger G_{ji}^* (G_0^{-1})_{ii}^\dagger. \quad (\text{A2.7})$$

Or $G_{ji}^* = G_{ij}^\dagger$ et la nouvelle expression de K s'exprime :

$$K = \sum_{i \in 1, j \in 2} G_{ji} (G_0^{-1})_{ii} n_{ii} (G_0^{-1})_{jj}^\dagger G_{ij}^\dagger (G_0^{-1})_{jj} n_{jj} (G_0^{-1})_{jj}, \quad (\text{A2.8})$$

qui peut se condenser en :

$$K = \text{Tr} \left[G_{21} (G_0^{-1})_{11} n_{11} (G_0^{-1})_1^\dagger G_{12}^\dagger (G_0^{-1})_{22} n_{22} (G_0^{-1})_{22}^\dagger \right], \quad (\text{A2.9})$$

où les indices 1 et 2 définissent les matrices des opérateurs dans les régions 1 et 2 respectivement. La fonction de GREEN du système non-couplé est indépendante de l'endroit où la division de la nano-structure s'opère. Nous pouvons en conséquence diviser le système en trois parties et écrire (figure (6.22)b) :

$$\begin{aligned} (G_0^{-1})_{11} n_{11} (G_0^{-1})_{11}^\dagger &= (G_0^{-1})_{GG} n_{GG} (G_0^{-1})_{GG}^\dagger, \\ (G_0^{-1})_{22} n_{22} (G_0^{-1})_{22}^\dagger &= (G_0^{-1})_{DD} n_{DD} (G_0^{-1})_{DD}^\dagger. \end{aligned} \quad (\text{A2.10})$$

Conjuguant les équations (A2.10) et (A2.9), il vient :

$$K = G_{DG} (G_0^{-1})_{GG} n_{GG} (G_0^{-1})_{GG}^\dagger G_{GD}^\dagger (G_0^{-1})_{DD} n_{DD} (G_0^{-1})_{DD}. \quad (\text{A2.11})$$

L'équation de DYSON fournit les égalités suivantes :

$$\begin{aligned}
 G_{DG} (G_0^{-1})_{GG} &= G_{DM} V_{MG}, \\
 (G_0^{-1})_{GG}^\dagger G_{GD}^\dagger &= V_{GM}^\dagger G_{MD}^\dagger, \\
 (G_0^{-1})_{DD} G_{DM} &= V_{DM} G_{MM}, \\
 G_{MD}^\dagger (G_0^{-1})_{DD}^\dagger &= G_{MM}^\dagger V_{MD}^\dagger.
 \end{aligned} \tag{A2.12}$$

Injectant les équations (A2.12) dans (A2.11), et en posant :

$$\begin{aligned}
 \Gamma^G &= 2\pi V_{MG} n_{GG} V_{GM}^\dagger, \\
 \Gamma^D &= 2\pi V_{MD} n_{DD} V_{DM}^\dagger,
 \end{aligned} \tag{A2.13}$$

On peut écrire tout d'abord :

$$K = \frac{1}{2\pi} \text{Tr} \left[G_{DM} \Gamma_{MM}^G G_{MD}^\dagger (G_0^{-1})_{DD}^\dagger n_{DD} (G_0^{-1})_{DD}^{-1} \right], \tag{A2.14}$$

puis :

$$K = \frac{1}{(2\pi)^2} \text{Tr} \left[(\Gamma^G)_{MM} G_{MM}^\dagger (\Gamma^D)_{MM} G_{MM} \right]. \tag{A2.15}$$

La nouvelle expression du courant est alors :

$$I = -\frac{e}{\pi\hbar} \int d\varepsilon \{ f(\varepsilon - \mu_G) - f(\varepsilon - \mu_D) \} \text{Tr}[\Gamma^G G^\dagger \Gamma^D G]. \tag{A2.16}$$

L'invariance de la trace par permutation circulaire, nous permet d'obtenir l'équation (A2.2).

Communications

PUBLICATIONS DANS DES REVUES AVEC COMITÉ DE LECTURE

1. N. Cavassilas, M. Bescond, F. Payet, and J.L. Autran, "Simulation of a free SiGe strained-Si double-gate nanotransistor," soumis à *J. Appl. Phys.* (2004).
2. M. Bescond, J.L. Autran, D. Munteanu, N. Cavassilas, and M. Lannoo, "Influence of point defects on multiple-gate MOSFET characteristics using the nonequilibrium Green's function formalism," accepté dans *J. Comp. Electron.* (2005).
3. M. Bescond, J.L. Autran, D. Munteanu, and M. Lannoo "Atomic-scale modeling of Double-Gate MOSFETs using a tight-binding Green's function formalism," *Solid-State Electron.* **48**, 567 (2004).
4. M. Bescond, M. Lannoo, D. Goguenheim and J.L. Autran, "Towards a full microscopic approach to the modeling of transistors with nanometer dimensions," *J. Non-Crys. Sol.*, **322**, 160 (2003).

CONFÉRENCES INTERNATIONALES

1. M. Bescond, K. Nehari, J.L. Autran, N. Cavassilas, D. Munteanu, and M. Lannoo, "3D quantum modeling and simulation of multiple-gate nanowire MOSFETs," *IEDM Tech. Digest*, p.617, San Francisco (USA), Décembre 2004.
2. M. Bescond, J.L. Autran, N. Cavassilas, D. Munteanu, and M. Lannoo, "Treatment of point defects in nanowire MOSFETs using the nonequilibrium Green's function formalism," *Proceedings 10th International Workshop on Computational Electronics (IWCE10)*, p.84, Purdue University (USA), Octobre 2004.
3. M. Bescond, K. Nehari, N. Cavassilas, D. Munteanu, M. Lannoo, and J.L. Autran, "Quantum transport in modeling of nanoscale MOSFETs," *Proceedings of Ultimate Lithography and Nanodevice Engineering conference*, Agelonde (France), Juin 2004 (Invited paper).
4. M. Bescond, J.L. Autran, D. Munteanu, N. Cavassilas, and M. Lannoo, "Theoretical influence of discrete oxide charge on MOSFET operations," (Poster) *5th symposium SiO₂, Advanced dielectrics and related devices*, Chamonix Mont-Blanc (France), Juin 2004.
5. J.L. Autran, D. Munteanu, M. Houssa, M. Bescond, X. Garros, and C. Leroux, "Electrical modeling and simulations of nanoscale MOS devices with a high-permittivity dielectric gate stack," *MRS Proceedings volume 811, symposium D* (Spring 2004).

-
6. D. Munteanu, J.L. Autran, M. Bescond, and M. Houssa, "Impact of high- κ gate dielectric on decanometer double-gate MOSFETs : gate-fringing field and parasitic charge effects," *Proceedings of the 5th European Workshop on Ultimate Integration of Silicon* (ULIS 2004), Leuven (Belgium), Mars 2004.
 7. M. Bescond, J.L. Autran, D. Munteanu, N. Cavassilas, and M. Lannoo, "Atomic-scale modeling of source-to-drain tunneling in ultimate schottky barrier double-gate MOSFET's," *Proceedings of ESSDERC Conferences*, p.395, Estoril (Portugal), Septembre 2003.
 8. M. Bescond, J.L. Autran, and M. Lannoo, "Quantum transport simulation in DG MOSFETs using a tight binding Greens's function formalism," *Proceedings of the 4th European Workshop on Ultimate Integration of Silicon* (ULIS 2003), p.121, Udine (Italy), Mars 2003.
 9. M. Bescond, M. Lannoo, D. Goguenheim, and J.L. Autran, "Towards a full microscopic approach to the modeling of transistors with nanometer dimensions", *Proceedings 4th symposium SiO2 , Advanced dielectrics and related devices*, Trento (Italy), Septembre 2003.

CONFÉRENCES NATIONALES

1. M. Bescond, J.L. Autran, N. Cavassilas, K. Nehari, L. Raymond, D. Munteanu, et M. Lannoo, "Simulation quantique 3D d'un MOSFET Gate-al-around," (Poster) *Proceedings Journées Nationales de Microélectronique et d'Optoélectronique* (JNMO), p.259, Juin 2004.
2. N. Cavassilas, M. Bescond, J.L. Autran, "Transport quantique dans un transistor contraint" (poster) *Proceedings Journées Nationales de Microélectronique et d'Optoélectronique* (JNMO), p.257, Juin 2004.
3. M. Bescond, J.L. Autran, et M. Lannoo, "Modélisation atomique d'un transistor MOSFET double grille," *Proceedings GDR Nano-électronique* (CNRS 2054), p.97, Grenoble (France), Janvier 2003.
4. M. Bescond, J.L. Autran et M. Lannoo "Modélisation atomique d'un transistor MOSFET double-grilles à canal confiné" *Proceedings VI^{èmes} Journées Nationales du Réseau Doctoral de Microélectronique* (JNRDM), Toulouse, 2003.

Titre : Modélisation et simulation du transport quantique dans les MOSFET nanométriques

Résumé :

La réduction régulière de la taille des transistors conduit aujourd'hui à des dispositifs nanométriques dans lesquels les effets quantiques sont de plus en plus prédominants. Ce travail modélise des transistors MOSFET ultimes et détermine l'impact des effets quantiques dans les architectures multi-grilles émergentes. Nous utilisons le formalisme auto-cohérent des fonctions de Green hors-équilibre exprimé dans la théorie des liaisons fortes. Nous simulons tout d'abord un transistor double-grille 2D confiné, dans lequel l'axe source-drain est représenté par une chaîne atomique. Nous étudions l'amplitude du courant tunnel source-drain en fonction de la longueur de grille et montrons que les transistors conservent des caractéristiques électriques acceptables jusqu'à une longueur de grille de 7 nm. Nous développons ensuite un modèle 3D pour décrire les architectures à nano-fil de silicium (*Tri-gate*, *Pi-gate*, *Omega-gate*, *Gate-all-around*). Une étude détaillée illustre plusieurs concepts de la théorie de transport de Landauer (quantum de conductance, résistance des réservoirs) et compare les performances électriques de chaque configuration de grille. Nous discutons l'influence du contrôle électrostatique en fonction de la longueur de grille et des dimensions de la section transversale. Enfin, nous proposons un modèle capable de traiter la présence de défauts ponctuels dans de tels composants 3D et analysons l'impact de leur type et de leur position.

Mots clés : Nanotechnologies, Théorie de transport de Landauer, Fonctions de Green hors-équilibre, MOSFET nano-métriques, Liaisons fortes auto-cohérentes, Effet tunnel, Confinement quantique, Défauts ponctuels.

Ecole Doctorale : Physique, Modélisation et Sciences pour l'Ingénieur

Spécialité : Physique et modélisation des systèmes complexes

Title : Modeling and simulation of quantum transport in nanoscale MOSFETs

Abstract :

The aggressive downscaling of the transistors leads to nanoscale devices in which quantum effects are increasingly prevalent. This work models MOSFETs transistors and determines the influence of quantum effects on multiple-gate emerging architectures. We use the self-consistent non-equilibrium Green's function formalism expressed in the tight-binding theory. We first model a two-dimension double-gate MOSFET, in which source-drain axis is represented by an atomic chain. We study the magnitude of the tunneling current between source and drain according to the channel length and show that the transistor still has acceptable electrical characteristics down to a 7 nm channel length. We then develop a three-dimension model to describe the emerging MOSFET architectures *i.e.* *Tri-gate*, *Pi-gate*, *Omega-gate*, and *Gate-all-around*. A detailed study illustrates several concepts of the theory of Landauer transport (quantum of conductance, resistance of the reservoirs) and compares the electrical performances associated to each gate configuration. We discuss the influence of electrostatic control according to the channel length and dimensions of the cross-section. Lastly, we propose a model able to treat point defects in such components and investigate the impact of their type and their position.

Key words : Nanotechnology, Theory of Landauer transport, Non-equilibrium Green function, Nanoscale MOSFETs, Self-consistent tight-binding, Tunneling effect, Quantum confinement, Point defects.

Thèse effectuée au Laboratoire Matériaux et Microélectronique de Provence, UMR CNRS 6137, Faculté de sciences de St Jérôme, case 142, 13397 Marseille cedex 20.