

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.

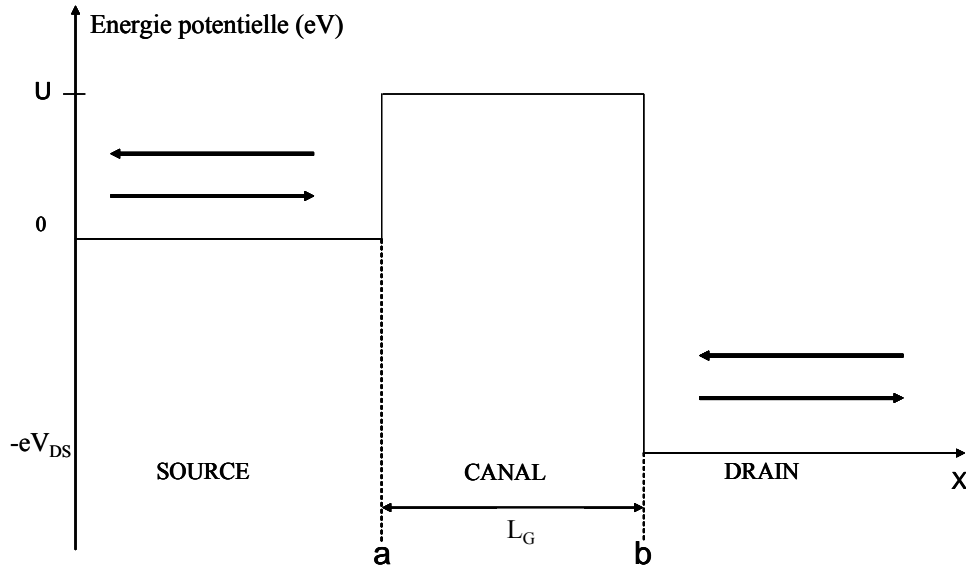


FIG. 4.1: Allure de l'énergie potentielle 1D le long de l'axe source-drain d'un transistor bloqué ($V_G \simeq 0$ V). Les solutions de l'équation de SCHRÖDINGER sont dans l'approximation de la masse effective des combinaisons linéaires d'ondes planes.

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.

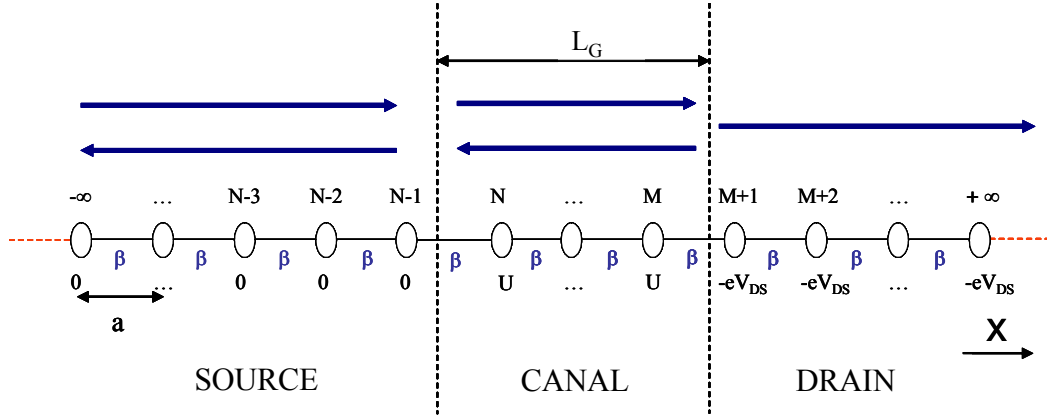


FIG. 4.2: Chaîne linéaire atomique représentant l'axe source-drain d'un transistor MOSFET. Les atomes sont numérotés de $-\infty$ à $+\infty$ et les indices sous chaque atome indiquent l'énergie potentielle intra-atomique. a est le paramètre de maille et β l'énergie de couplage.

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} .

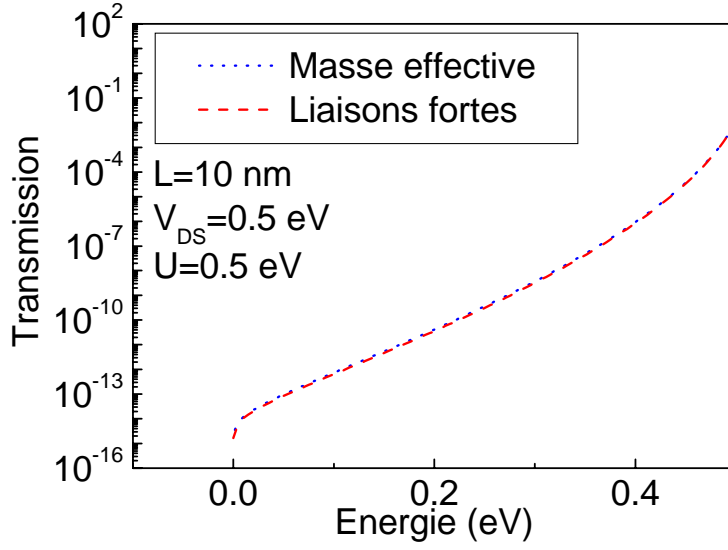


FIG. 4.3: Comparaison des coefficients de transmissions obtenus successivement dans l'approximation de la masse effective et la théorie des liaisons fortes.

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

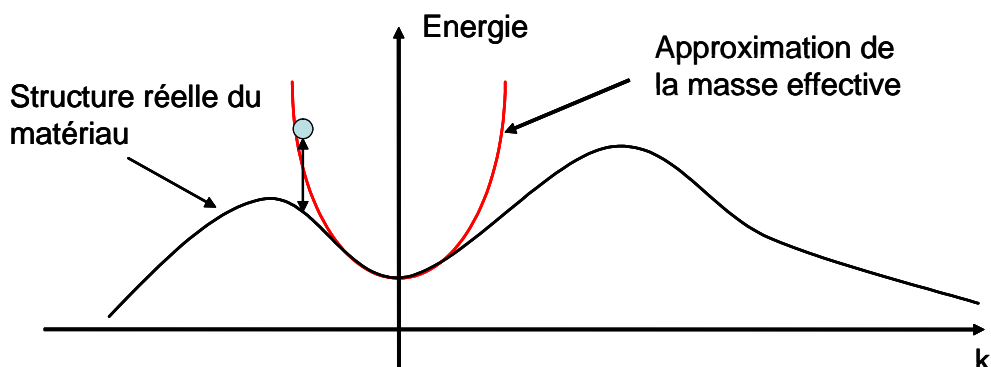


FIG. 4.4: Représentation schématique des limites de l'approximation de la masse effective pour des électrons hautement énergétiques.

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)).

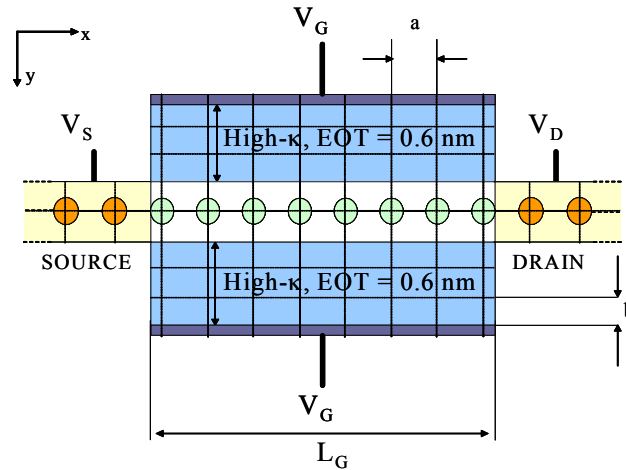


FIG. 4.5: MOSFET double-grille à canal de conduction confiné. L'équation de POISSON 2D est résolue en utilisant la méthode des différences finies : les deux directions \mathbf{x} et \mathbf{y} sont décrites par un maillage uniforme de pas respectivement égal à a et b .

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 FERMİ 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 FERMİ 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

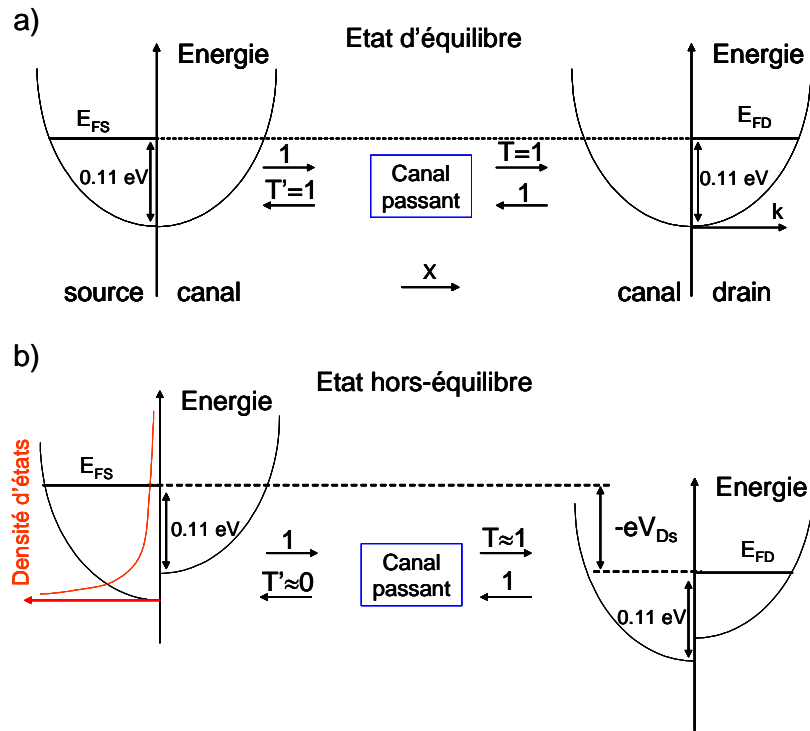


FIG. 4.6: Représentation du transport de charge dans un système balistique 1D a) à l'équilibre, b) dans un régime hors-équilibre. L'électro-neutralité des contacts est obtenue en grâce aux réflexions des interfaces.

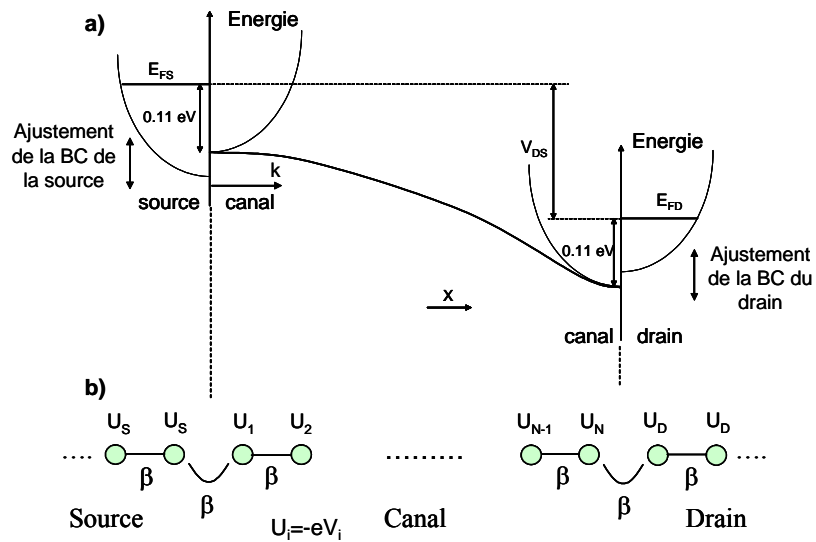


FIG. 4.7: a) Modélisation des bandes de conduction aux interfaces source/canal et canal/drain. b) Equivalence en terme de potentiel électrostatique et énergies de couplage. V_i représente le potentiel électrostatique le long de la chaîne.

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.

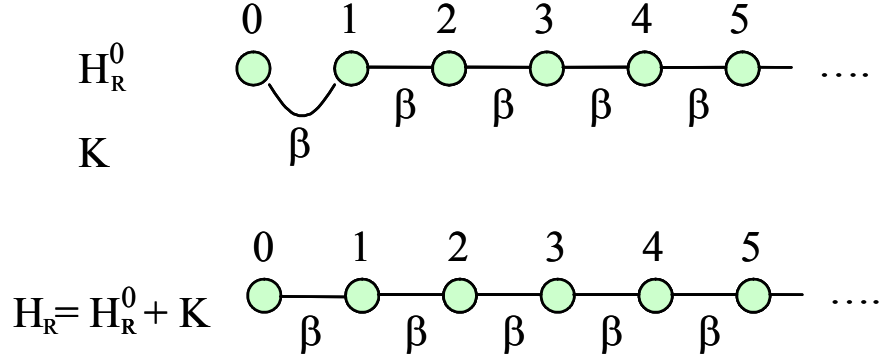


FIG. 4.8: Calcul de la fonction de GREEN de surface d'une chaîne linéaire semi-infinie. L'hamiltonien de départ H_R^0 décrit une chaîne linéaire plus un atome isolé. La perturbation K représente le couplage entre l'atome 0 et la chaîne semi-infinie commençant à l'atome 1.

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

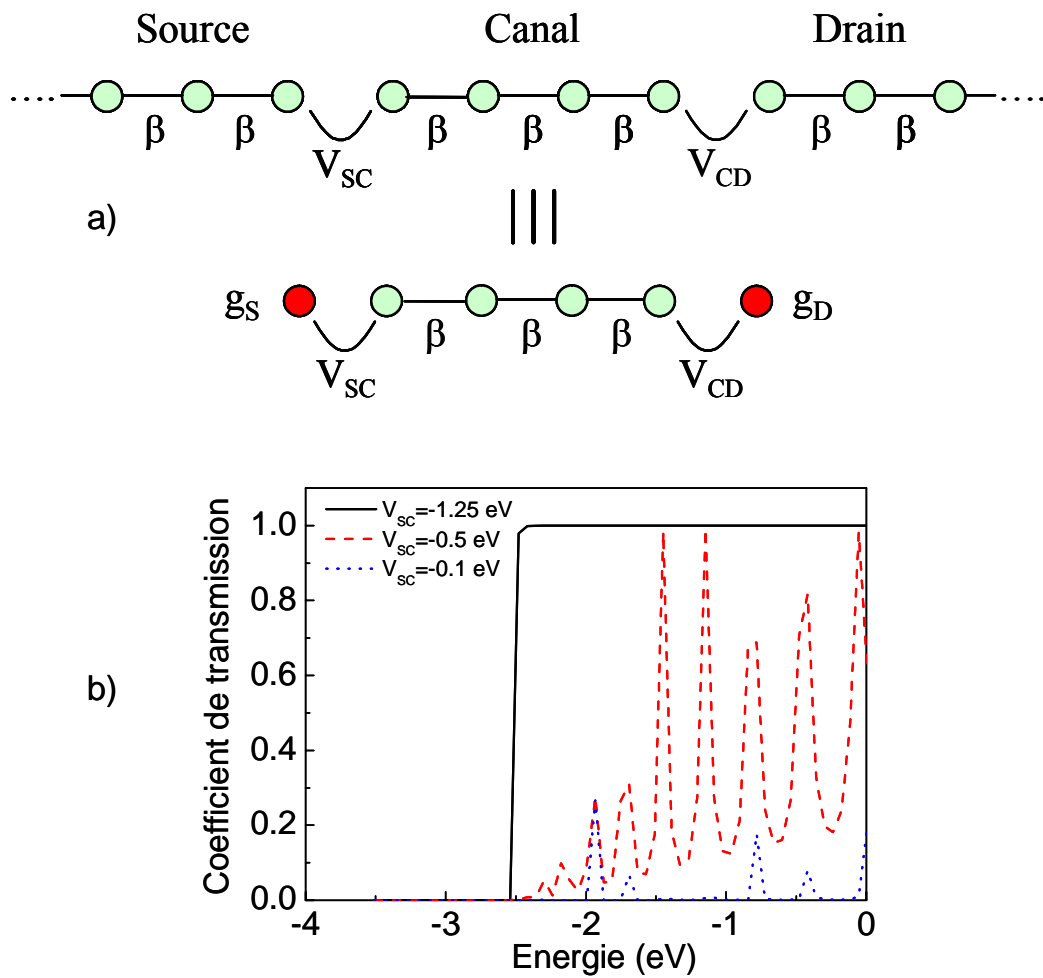


FIG. 4.9: (a) Le concept de la self-énergie permet de modéliser les réservoirs d'électrons en élargissant les niveaux discrets de la zone active isolée. (b) Evolution du coefficient de transmission en fonction du couplage de la zone active aux réservoirs. $\beta = -1.25$ eV, $V_{SC} = V_{CD}$ et $g_{S(D)}$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

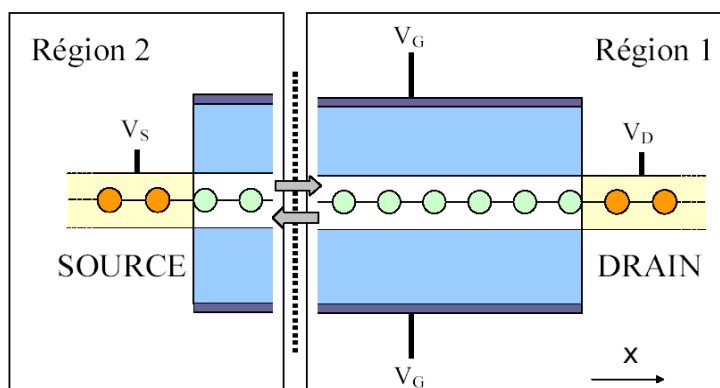


FIG. 4.10: Calcul du courant : le système est virtuellement clivé en deux régions. Cette approche, dite théorie de la diffusion, exprime la diffusion des états électroniques d'une première région en états de la seconde.

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 :

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

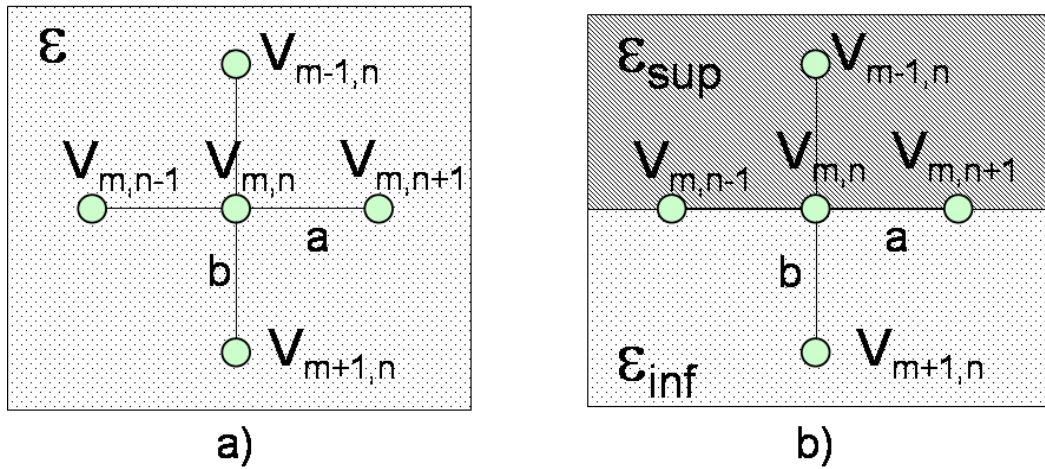


FIG. 4.11: *Discretisation de l'équation de POISSON en différences finies. a) Dans un matériau homogène de constante diélectrique ϵ . b) A l'interface entre deux matériaux aux constantes diélectriques distinctes ϵ_{sup} et ϵ_{inf} .*

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"

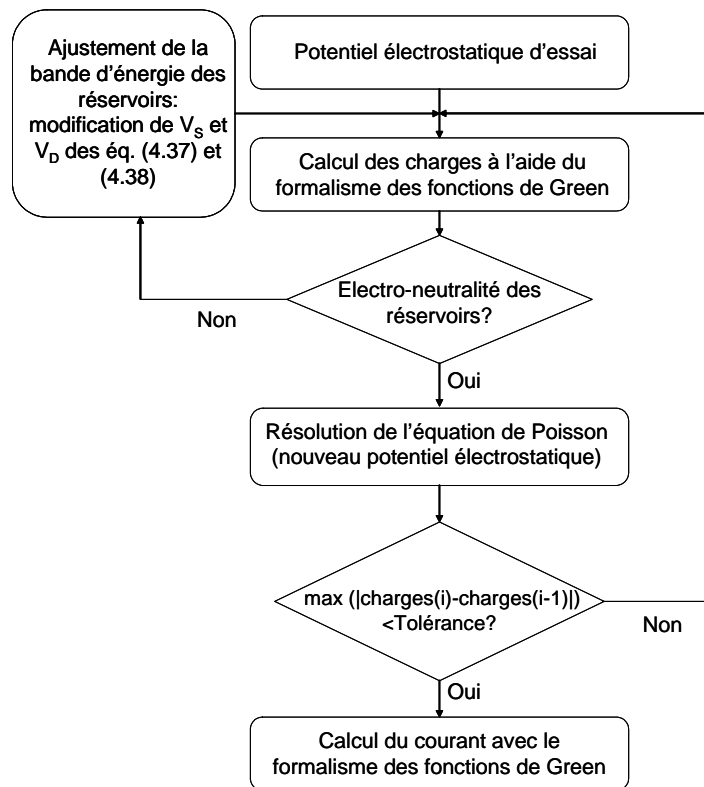


FIG. 4.12: *Etapes numériques de la simulation. L'indice i représente le nombre d'itérations de la boucle SCHRÖDINGER-POISSON.*

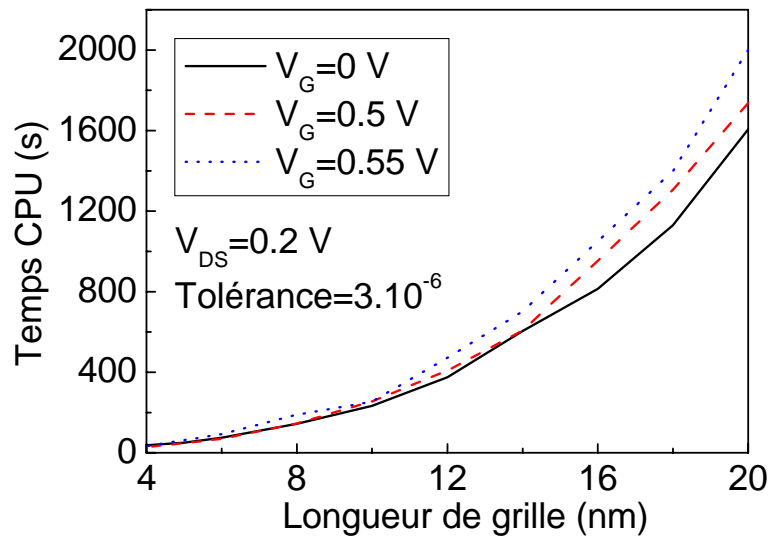


FIG. 4.13: *Temps de calcul CPU d'un point de courant en fonction de la longueur de grille, pour différentes tensions V_G .*

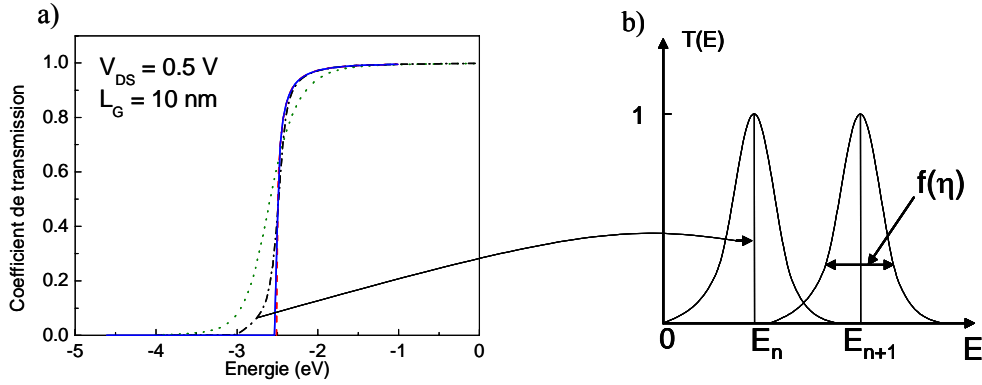


FIG. 4.14: a) Evolution du coefficient de transmission d'une chaîne linéaire en fonction de l'élargissement complexe η . Ligne discontinue : chaîne infinie avec un élargissement nul. Ligne continue : chaîne finie avec un élargissement artificiel $\eta = 0.1$. Ligne tiret-pointillés : chaîne finie avec 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

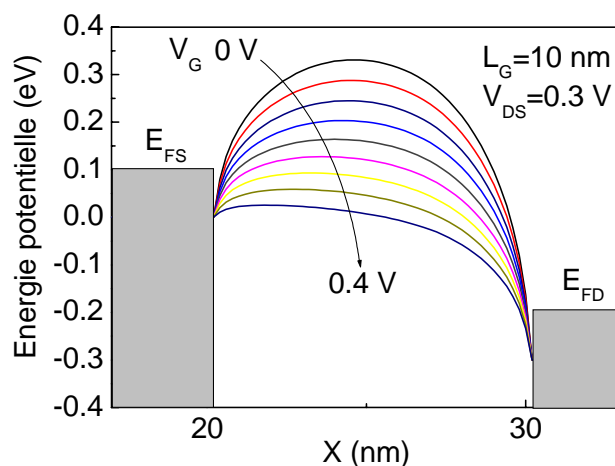


FIG. 4.15: Profils de l'énergie potentielle le long du canal sous une tension de drain $V_{DS} = 0.3\text{V}$. Les régions grisées représentent la source et le drain dont les niveaux de FERMÍ sont respectivement E_{FS} et E_{FD} . V_G varie de 0 à 0.4 V par pas de 0.05.

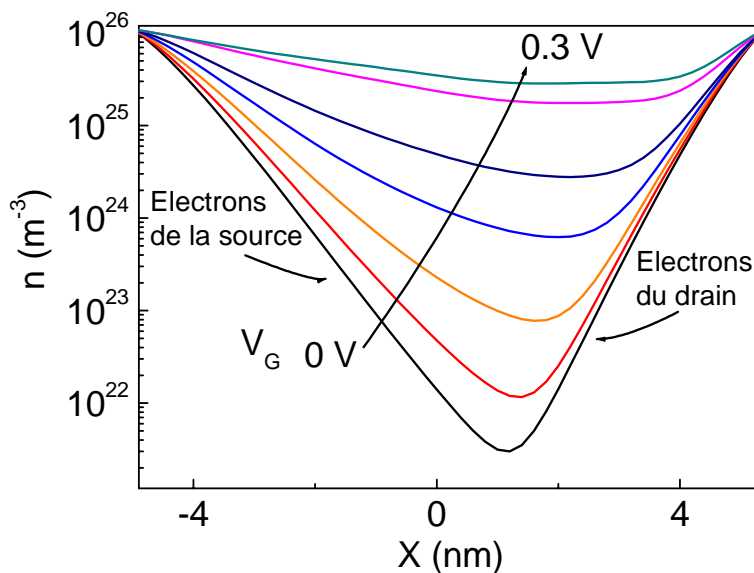


FIG. 4.16: Densité électronique le long du canal obtenue par le formalisme des fonctions de GREEN. A faible V_G l'évolution exponentielle du nombre de porteurs dans le canal traduit une domination de l'effet tunnel dans le transport. L'augmentation de V_G fait commuter le transistor et accroît la concentration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

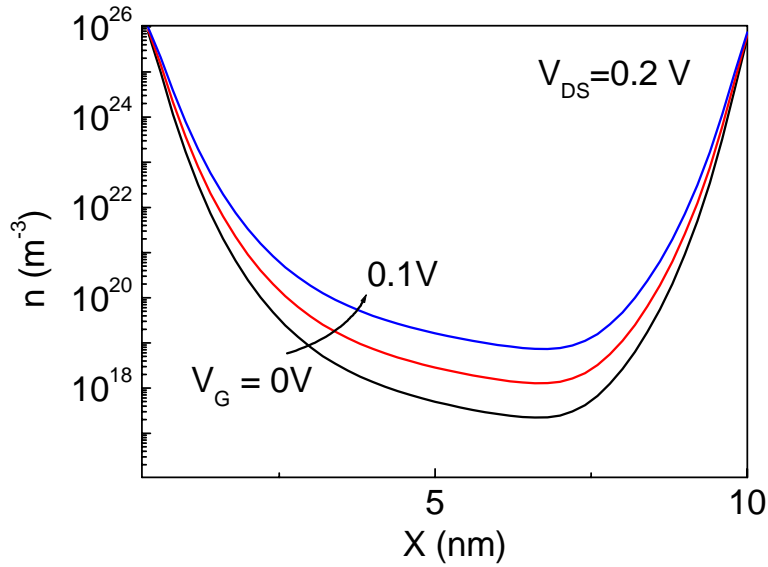


FIG. 4.17: Densité électronique le long du canal proposée par une approche classique de dérivation-diffusion (cf § (3.1.1)). A tension et dimensions équivalentes, la quantité d'électrons en régime sous le seuil chute d'un facteur 10^4 par rapport à l'approche quantique des fonctions de GREEN (figure (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.

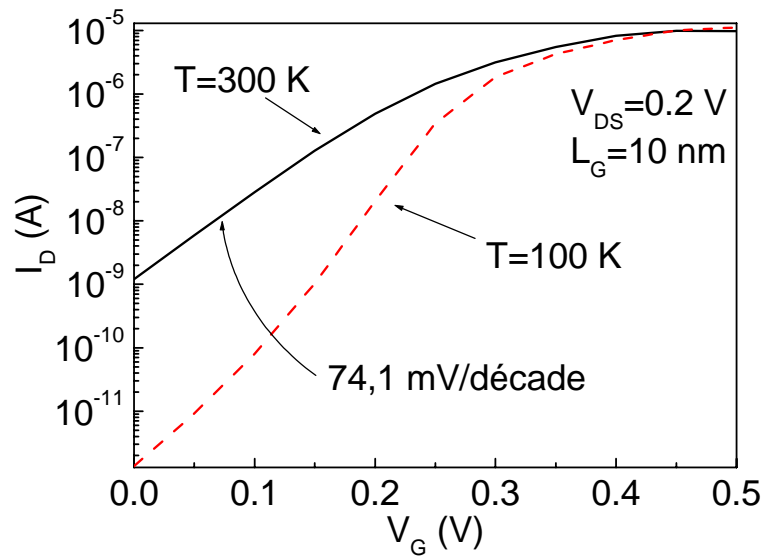


FIG. 4.18: Caractéristiques I_D en fonction de V_G simulées à deux températures, 100 et 300 K ($V_{DS} = 0.2$ V et $L_G = 10$ nm).

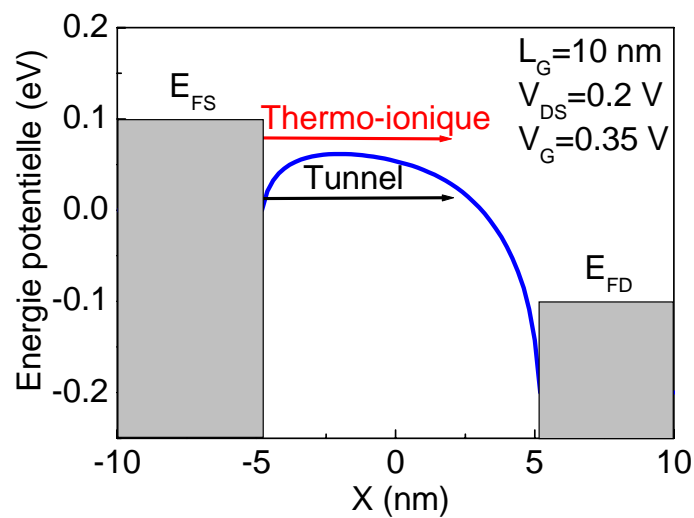


FIG. 4.19: Le courant total d'un nano-transistor peut se décomposer en une composante tunnel et une composante thermo-ionique.

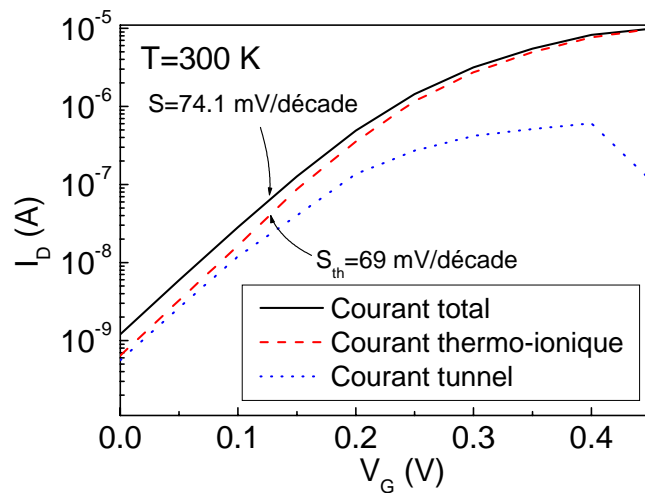


FIG. 4.20: Caractéristiques I_D-V_G calculées pour une longueur de grille $L_G = 10$ nm. Tous les autres paramètres sont identiques à ceux de la figure (4.18). Le courant total est la somme des composantes thermo-ionique et tunnel ($T = 300$ K).

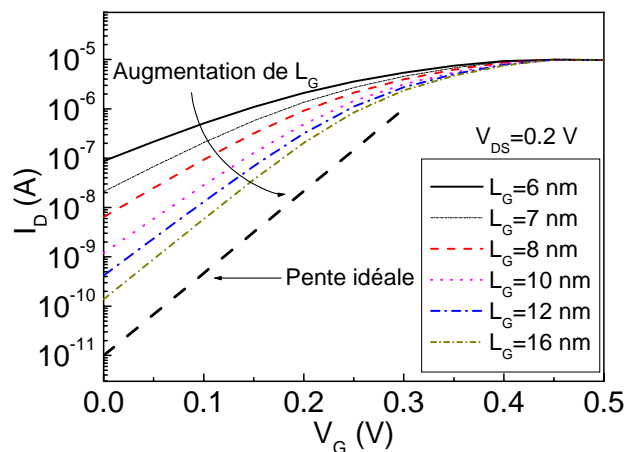


FIG. 4.21: Courbes I_D-V_G pour des composants aux dimensions comprises entre 6 à 16 nm. Le courant tunnel dégrade fortement la pente sous le seuil des petits composants.

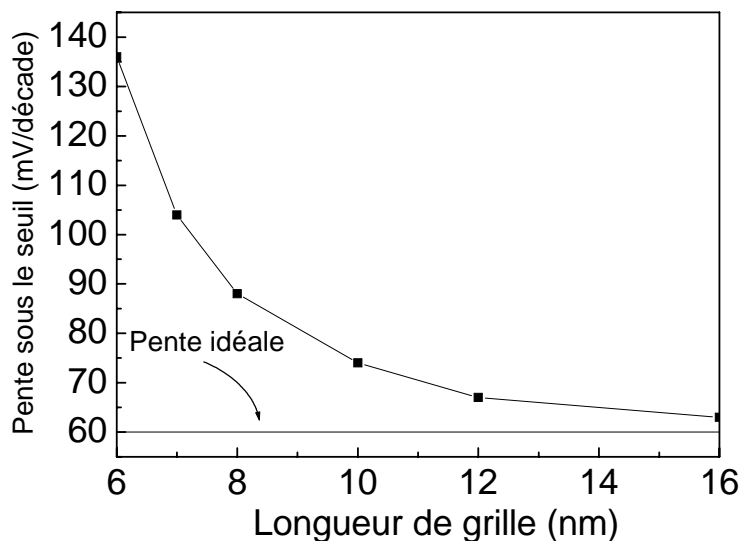


FIG. 4.22: *Pente sous le seuil en fonction de la longueur de grille.*

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).

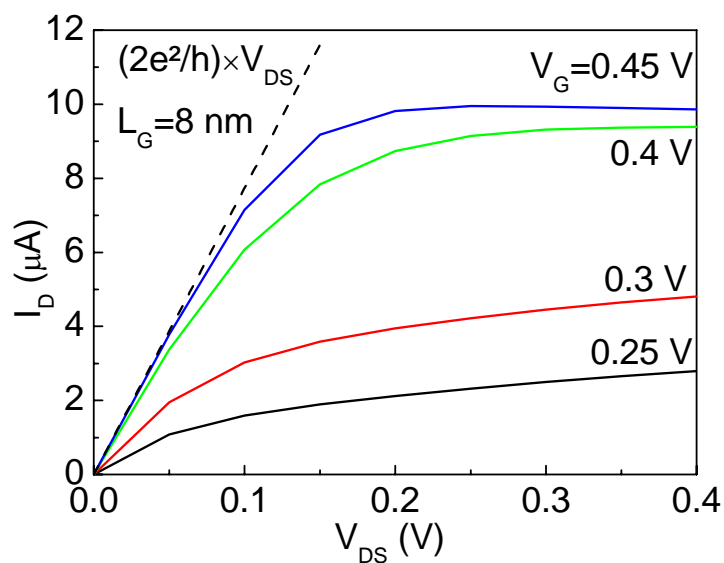


FIG. 4.23: Caractéristiques $I_D - V_{DS}$ sous différentes tensions de grille ($V_G = 0.25 - 0.45 \text{ V}$). La ligne en pointillés représente le courant fourni par le quantum de conductance $G_0 = 2e^2/h$ ($T = 300 \text{ K}$).

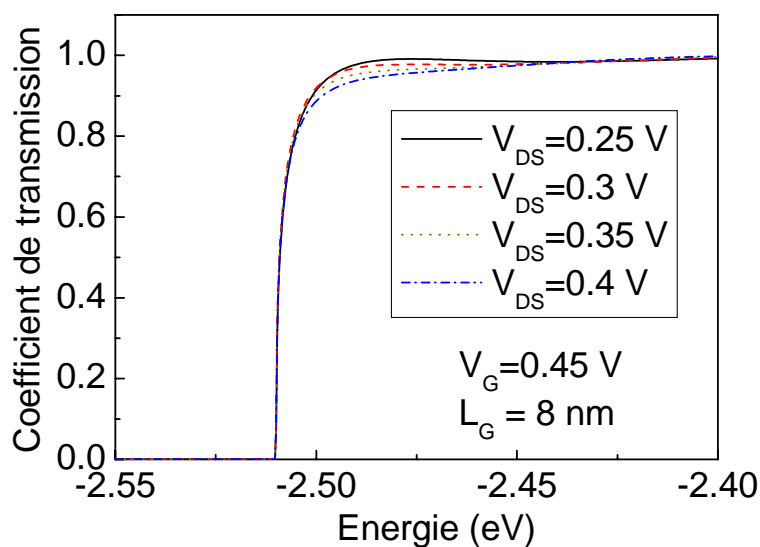


FIG. 4.24: Coefficients de transmission en fonction de l'énergie à $V_G = 0.45 \text{ V}$ pour différents V_{DS} . Le bas de bande de conduction du canal côté source est situé à 2β où $\beta = -1.25 \text{ eV}$.

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

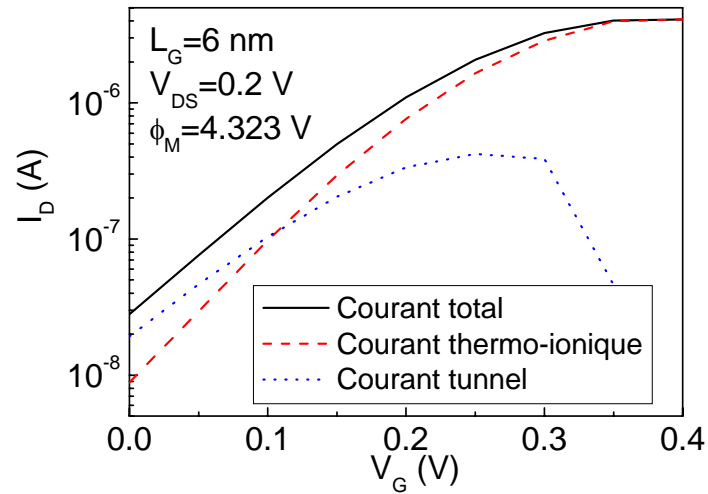


FIG. 4.25: Caractéristique I_D-V_G calculée pour une structure double-grille de longueur $L_G = 6 \text{ nm}$ et un travail de sortie identique à celui du composant de longueur $L_G = 10 \text{ nm}$: $\phi_M = 4.323 \text{ V}$. Le courant I_{OFF} est fortement dégradé.

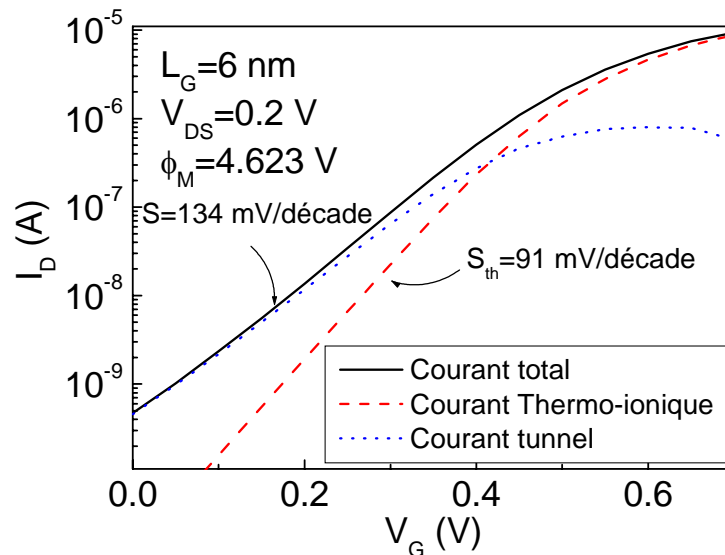


FIG. 4.26: Caractéristique I_D-V_G du dispositif double-grille de longueur $L_G = 6 \text{ nm}$, avec un travail de sortie $\phi_M = 4.623 \text{ V}$. Les autres paramètres sont identiques à ceux de la figure (4.25). Le courant I_{OFF} s'est amélioré ($< 10^{-9} \text{ A}$), au détriment de la tension 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}).

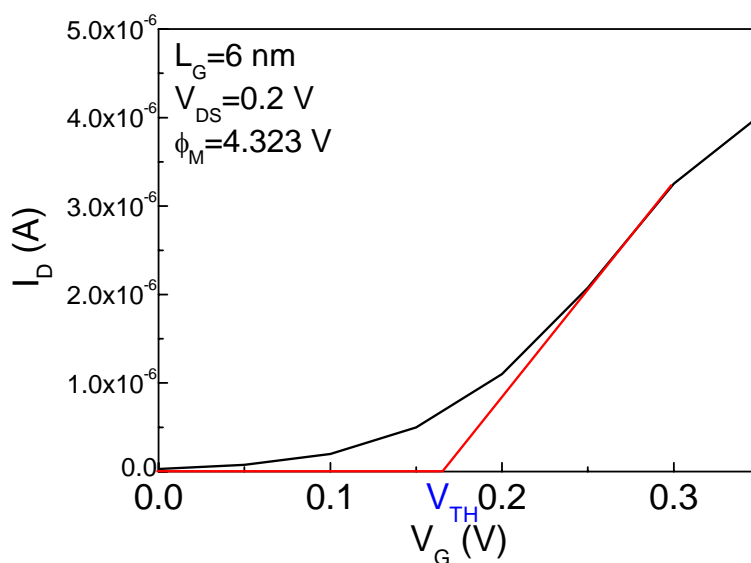


FIG. 4.27: Calcul géométrique de la tension de seuil.

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} .

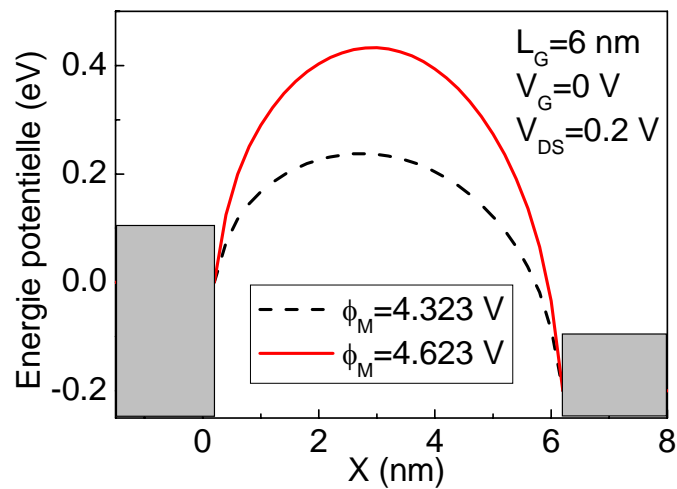


FIG. 4.28: *Energie potentielle le long du canal de longueur $L_G = 6$ nm pour deux travaux de sortie de grille ($\phi_M = 4.323$ et 4.623 V). $T = 300$ K.*

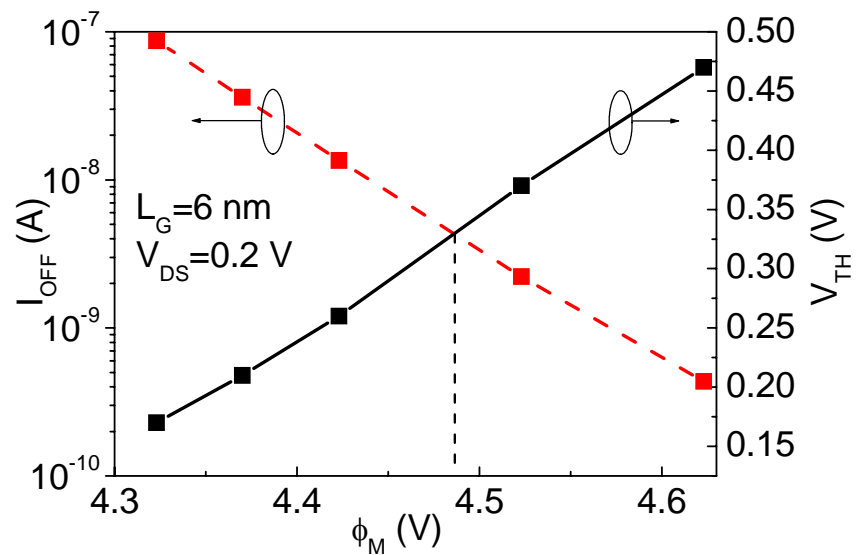


FIG. 4.29: *Courant I_{OFF} et tension de seuil V_{TH} en fonction du travail de sortie du métal de grille ϕ_M . Un bon compromis entre performances et vitesse de commutation semble atteint pour $\phi_M = 4.47$ V.*

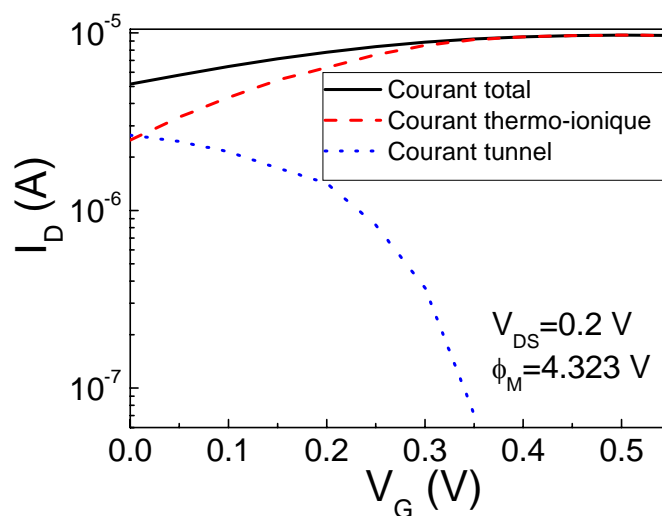


FIG. 4.30: Caractéristique I_D-V_G calculée pour une structure double-grille de longueur de grille $L_G = 3 \text{ nm}$ et $\phi_M = 4.323 \text{ V}$. Étonnamment, la composante tunnel du courant n'est plus dominante $T = 100 \text{ K}$.

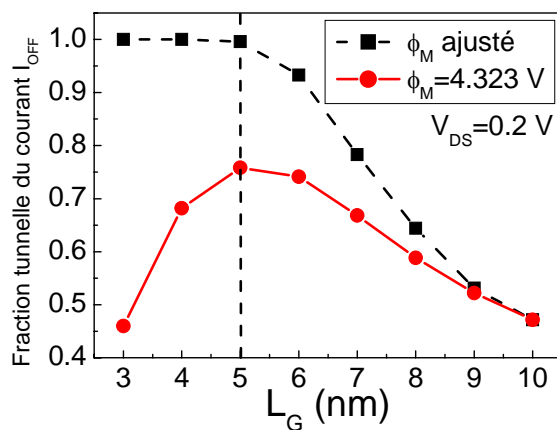


FIG. 4.31: Fraction du courant de drain tunnel @ $V_G = 0 \text{ V}$ en fonction de la longueur de grille, i) pour un travail de sortie de grille fixe $\phi_M = 4.323 \text{ V}$, ii) pour un travail de sortie de grille ajusté afin de maintenir le même courant $I_{OFF} \sim 10^{-9} \text{ A}$.

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.