

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 dûes 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.

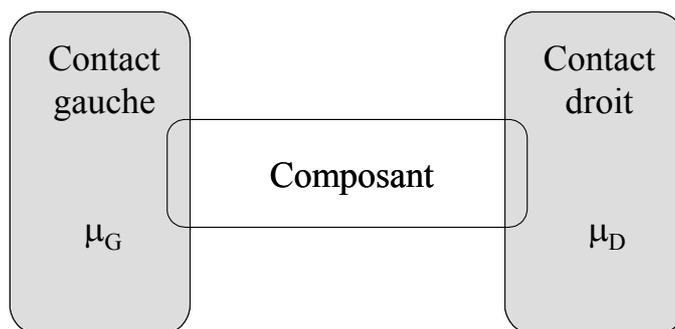


FIG. 3.1: Nano-dispositif couplé à deux électrodes semi-infinies (gauche et droite) avec des potentiels électrochimiques μ_G et μ_D .

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 :

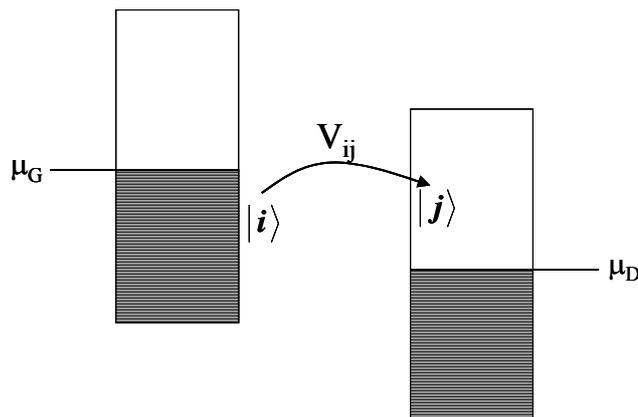


FIG. 3.2: Un électron a une faible probabilité d'être transféré via le couplage $V_{ij} = \langle i|V|j\rangle$, d'un état occupé $|i\rangle$ de l'électrode gauche vers un état vide $|j\rangle$ de l'électrode droite : la théorie des perturbations peut s'appliquer.

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

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

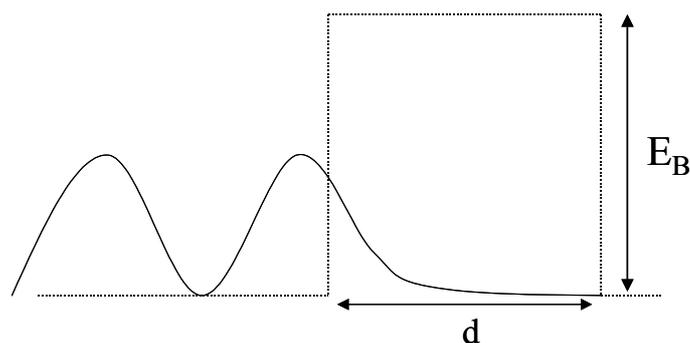


FIG. 3.3: Barrière d'énergie dans une jonction métal-isolant-métal.

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.

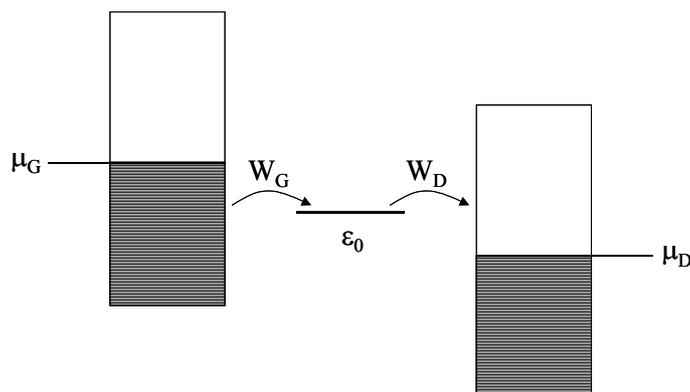


FIG. 3.4: *Transfert d'électrons à travers un îlot contenant un seul niveau ε_0 .*

$\varepsilon_n(q, V_{DG})$. 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

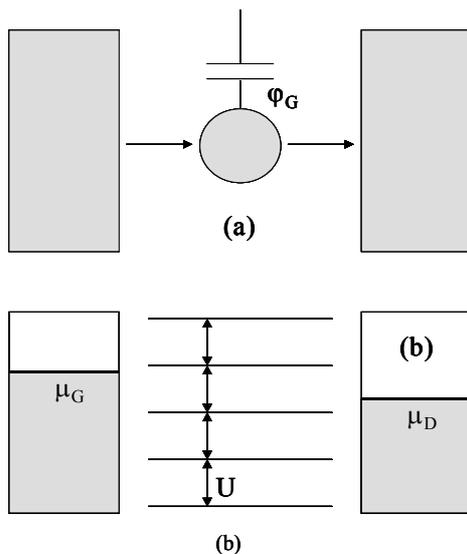


FIG. 3.5: (a) Transistor à un électron avec un îlot métallique. (b) Les niveaux ionisés de l'îlot sont régulièrement espacés par l'énergie de charge U .

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].

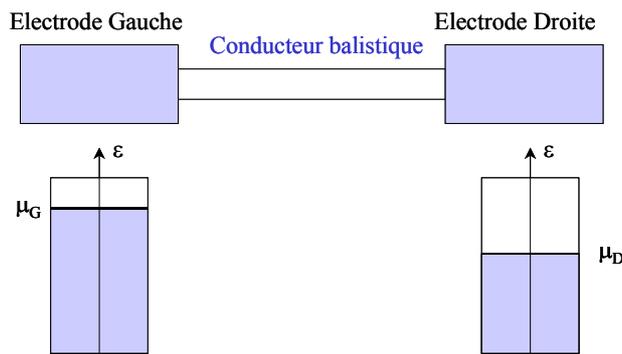


FIG. 3.6: Configuration d'un système mésoscopique balistique. Chaque réservoir, défini par son potentiel électrochimique, thermalise parfaitement tous les électrons incidents.

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].

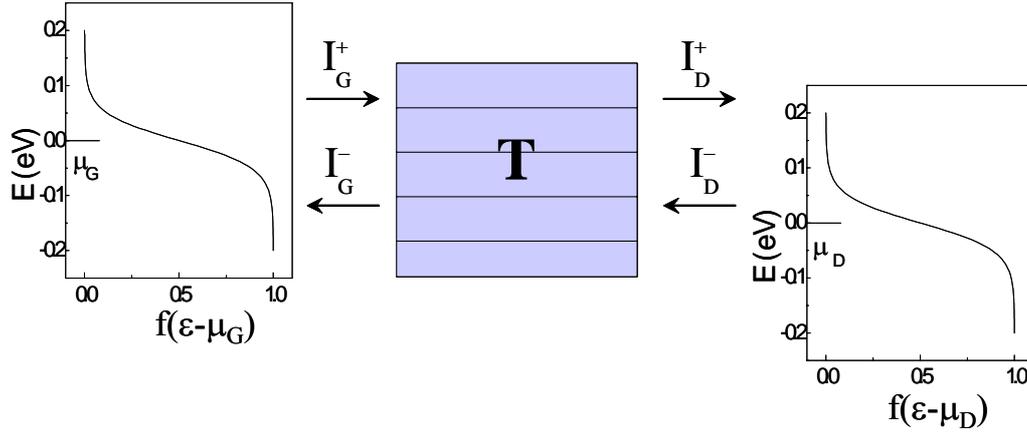


FIG. 3.7: Bilan des courants traversant un conducteur non-transparent.

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 ($D = I/V_{DG}$) :

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

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 FERMI-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 FERMI, 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 FERMI-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 FERMI-DIRAC de chaque réservoir (comme dans la théorie Orthodoxe [Eq.(3.34)]) :

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

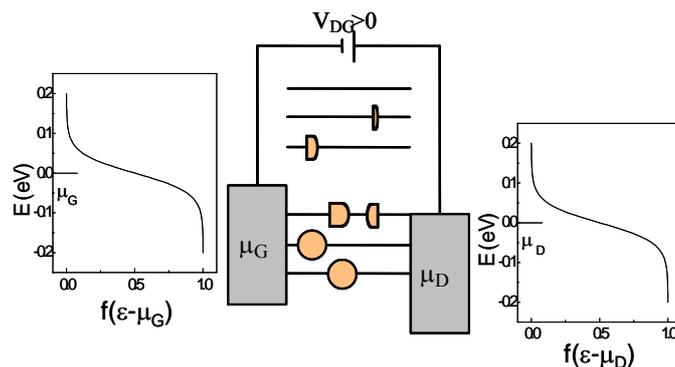


FIG. 3.8: Application d'une tension positive sur le réservoir de droite : les contacts gauche et droite veulent imposer leur distribution de FERMÍ-DIRAC aux électrons de la zone active.

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.

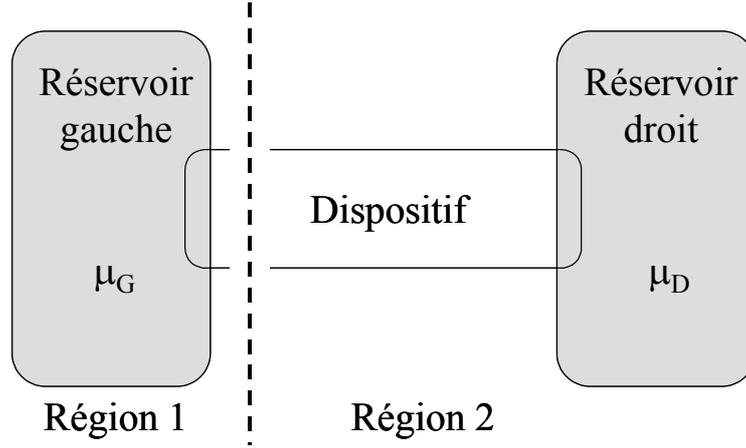


FIG. 3.9: Nano-structure virtuellement clivée en deux régions.

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 ($\eta \rightarrow 0^+$) :

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

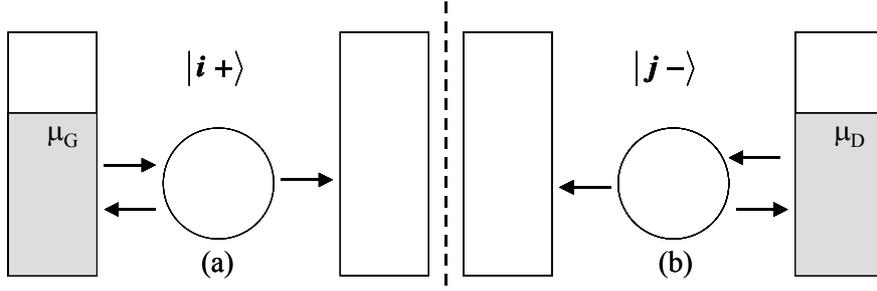


FIG. 3.10: Dans un système hors-équilibre, les états propres de l'hamiltonien total sont divisés en deux catégories : (a) ceux injectés depuis le contact gauche et occupés jusqu'au potentiel électrochimique μ_G et (b) ceux injectés depuis le contact droit et occupés jusqu'au potentiel électrochimique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

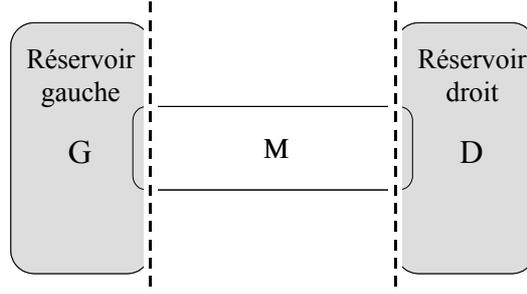


FIG. 3.11: Afin de tenir compte de l'architecture du dispositif, le système est divisé en trois parties : les deux réservoirs et la partie centrale M .

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

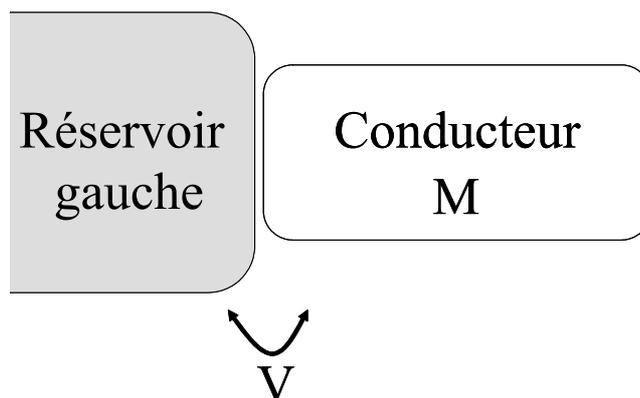


FIG. 3.12: Conducteur connecté à un réservoir via la matrice de couplage V .

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

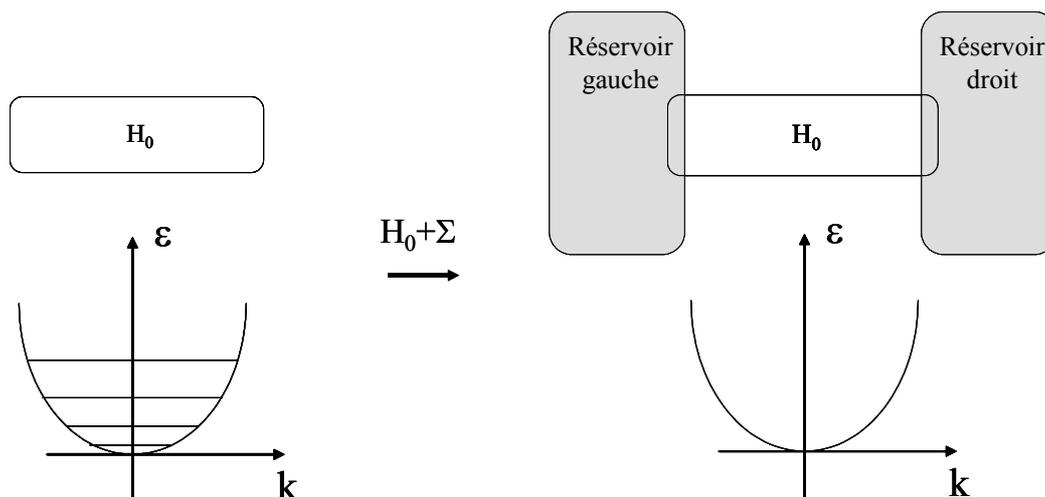


FIG. 3.13: L'interaction du nano-système avec des contacts entraîne une renormalisation de son hamiltonien par une self-énergie Σ et la création d'un continuum d'énergie.

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.

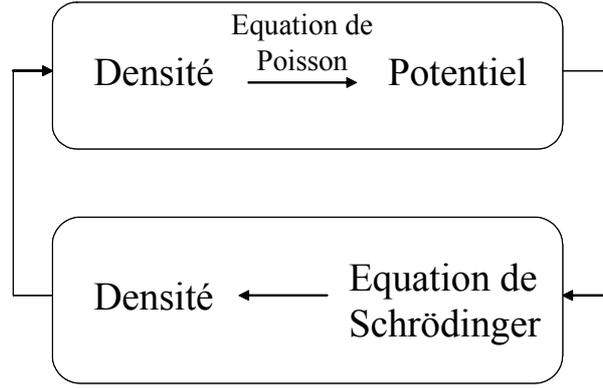


FIG. 3.14: La densité électronique et le potentiel doivent être calculés de manière auto-cohérente afin de tenir compte de l'occupation hors-équilibre des niveaux d'énergies.

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

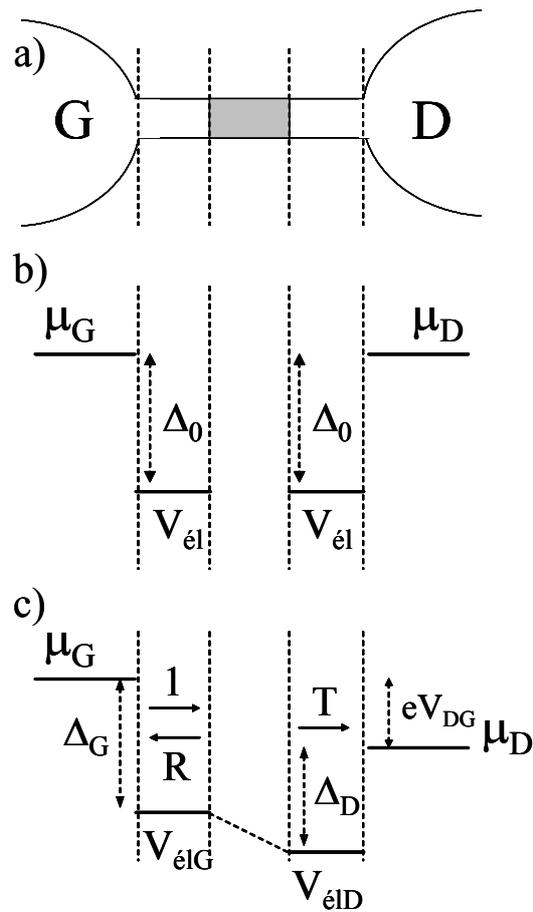


FIG. 3.15: a) Schéma d'une constriction (contenant un centre diffusif représenté en gris) connecté à deux réservoirs métalliques et leur diagramme énergétique associé (b) lorsque le système est à l'équilibre, (c) lorsqu'une tension V_{DG} est appliquée.