République Algérienne Démocratique et Populaire Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

UNIVERSITE MOULOUD MAMMERI DE TIZI-OUZOU



FACULTE DU GENIE ELECTRIQUE ET D' INFORMATIQUE DEPARTEMENT D'AUTOMATIQUE

Mémoire de Fin d'Etudes de MASTER ACADEMIQUE

Domaine : Sciences et Technologies Filière : Génie électrique Spécialité : Génie Microélectronique

> Présenté par BELMADI Mohamed

TECHNOLOGIE ET INTEGRATION EN SEMI-CONDUCTEURS III-V

Mémoire soutenu publiquement le 14/07/2016 devant le jury composé de :

M Hamid BOUZAR Professeur, UMMTO/FS, Président

M Arezki BENFDILA Professeur, UMMTO, Encadreur

M Ahcene LAKHLEF MCB, UMMTO, Examinateur

Melle Zedjiga HATEM MCB, UMMTO, Examinatrice

Remerciements

Ce travail n'aurait pas pu être mené à lieu, sans l'aide et le soutien de Monsieur BENFDILA.AREZKI, je tiens à lui exprimer ma profonde reconnaissance pour son soutien et les conseils qu'il m'a accordé tout au long de ce mémoire.

Je remercie monsieur LAKHELF qui a apporté de l'aide à ce travail.

Je remercie également

Mr.BELHIMER.LOUNES

Mr.KESSI.MOHAND

Mr.DJOUDER.MOHAMED

Enfin, j'exprime ma profonde gratitude à mes parents ; ma mère en particulier, pour leurs efforts depuis le début de mes études, sans oublier de remercier mes amis.

Dédicace

A ma chère Mère

A mon cher Père

Mon Frère

Je voudrais remercier toutes personnes que j'ai été amené à rencontrer au cours des années passées, pour leur sympathie, leur générosité et leur bonne humeur. Ils ont tous contribué à leur façon au bon déroulement de ce travail.

Table de matières

Introduction générale

Chapitre I : LE TRANSISTOR MOSFET ET SON EVOLUTION.

I.1 : Introduction
I.2 : La technologie MOSFET et son évolution
I.2.1 : Bref historique4
I.2.2 : Le transistor MOSFET architecture et principe de fonctionnement5
I.2.3 : Fonctionnement du transistor MOSFET6
I.3 : Grandeurs caractéristiques du MOSFET7
I.4 : Effet de la miniaturisation des MOSFETs8
I.4.1 : Origine des effets de canaux courts10
I.4.2 : Effet du champ électrique sur la mobilité11
I.4.3 : Effet de la résistance série source-drain12
I.4.4 : Influence des zones de déplétion de source et de drain14
I.4.4.a) Partage de charge14
I.4.4.b) Effet de canal étroit17
I.4.5 : Effet DIBL
I.4.6 : Perçage
I.4.7 : Effet de porteurs chauds22
I.5 : Conclusion

CHAPITRE II : LES MATERIAUX SEMI-CONDUCTEURS III-V

II.1 : Introduction
II.2 : Différence entre isolant conducteur et semi-conducteur25
II.3 : différents types de dopage26
II.3.1 : Semi-conducteur dopé n26
II.3.2 : Semi-conducteur dopé p26
II.4 : Semi-conducteurs usuels
II.5 : Définition des semi-conducteurs III-V
II.6 : Les composés binaires, ternaires et quaternaires des semi-conducteurs III-V29
II.6.1 : Les composés binaires29
II.6.2 : Les composés ternaires et quaternaires
II.7 : Les matériaux massifs
II.7.1 : Propriétés cristallines
II.7.2 : Propriétés électroniques
II.8 : Propriétés du GaAs
II.9 : structure des bandes d'énergies des semi-conducteurs III-V
II.9.1 : Structure de bande du GaAs
II.10 : Bandes interdites des composés III-V
II.10.1 : Bandes interdites des principaux composés III-V

CHAPITRE III : LA FABRICATION COLLECTIVE DES COMPOSANTS ET DES CIRCUITS INTEGRES

III.1 : Introduction	
III.2 : La lithographie	40

III.3 : Les procédés de retraits des matériaux45
III.3.1 : La gravure humide46
III.3.2 : La gravure sèche47
III.4 : Les procédés d'apports de matériaux
III.4.1 : Evaporation
III.4.2 : Epitaxie par jet moléculaires (MBE)50
III.4.3 : Dépôt par « sputtering »51
III.4.4 : Dépôt par laser pulsé (PLD)52
III.4.5 : Le dépôt en phase vapeur (CVD)52
III.4.6 : Dépôt de type MOCVD (Metal Organic Chimical Vapour Deposition)54
III.4.7 : Dépôt de type CSD (Chimical Solution Deposition)54
III.4.8 : croissance thermique55
III.4.9 : La croissance électrolytique56
III.4.10 : Implantation ionique57
III.4.11 : La diffusion thermique59
III.5 : Les procédés alternatifs60
III.5.1 : La nano-impression60
III.5.2 : Technique d'auto-assemblage62

CONCLUSION GENERALE	63
BIBLIOGRAPHIE	64

Liste des Abréviations

- C_{ox} Capacité de l'oxyde par unité de surface.
- C_{ds} Capacité Drain-Source.
- C_{gs} Capacité Grille-Source.
- C_{js} Capacité Source-Substrat.
- C_{gd} Capacité Grille-Drain.
- e Charge élémentaire de l'électron.
- ε Permittivité relative.
- ε_{ox} Permittivité de l'oxyde.
- E_g Energie du gap.
- E_F Energie de fermi.
- E_{v} Energie de valence.
- E_c Energie de conduction.
- ψ_s Potentiel de surface.
- ψ_F Potentiel de jonction.

 ϕ_{ms} La différence de travail de sortie entre le métal et le semiconducteur. $\phi_{\rm F}$ Potentiel de fermi.

 $2\phi_{Fi}$ La différence de potentiel entre la surface et le volume du semiconducteur en régime de forte inversion.

 G_{ds} Conductance drain-source.

 G_{gs} Conductance grille-source.

 G_{max} Conductance maximale.

 G_m Transconductance.

h Constante de Planck.

 I_{DS} Courant drain-source.

 I_{DSsat} Courant de saturation drain-source.

K Constante de Boltzmann.

KT Energie thermique.

 L_g Longueur de grille.

L Longueur de grille.

 μ Mobilité des porteurs.

 μ_0 Mobilité des porteurs dans le volume du semi-conducteur à champ faible.

 μ_n Mobilité des électrons.

 μ_p Mobilité des trous dans le semi-conducteur.

 μ_r Mobilité relative.

mos Métal Oxyde Semi-conducteur.

MOSFET Métal Oxyde Semi-conducteur Field Effect Transistor.

n Concentration des porteurs de charge.

 n_i Densité intrinsèque des porteurs à l'équilibre thermodynamique.

 N^+ Zone de type N fortement dopée (contact de drain et source).

 N^- Zone de type N faiblement dopée (zone de « drift »).

N_a Concentration relative aux impuretés « accepteurs ».

N_d Concentration relative aux impuretés « donneurs ».

 N_D Densité de dopage de la couche épitaxie.

q Charge électrique élémentaire.

 Δv_{th} Décalage de la tension de seuil.

T Température.

 T_{ox} Epaisseur d'oxyde.

 V_d Tension de diffusion.

 V_{gs} Tension grille-source.

 V_{GD} Tension grille-drain.

 V_{DS} Tension drain-source.

 V_{DSsat} Tension de saturation drain-source.

 V_{th} Tension de seuil.

 V_{FB} Tension Flat Band

w Largeur de grille

Introduction Générale

La miniaturisation des transistors CMOS permet d'améliorer les performances, la densité d'intégration et le coût des circuits électroniques. La loi de Moore s'est imposée comme un modèle économique décrivant l'évolution technologique des circuits intégrés. Elle prévoit (impose) une multiplication par 2 tous les 18 mois du nombre de transistors par puce. Quelles sont les limites de la technologie silicium ? Telle semble être la question qui motive la recherche technologique en microélectronique depuis 30 ans.

La technologie CMOS représente près de 90% du marché des semi-conducteurs et poursuit sa route dans la miniaturisation qui amènera l'utilisation de dispositifs MOS de longueur de grille de 40-50nm en 2007-2008 au plan industriel comme l'illustre la figure 1 :



Figure 1 Évolution de la longueur de grille des transistors (d'après la feuille de route ITRS 2001).

La réduction de la longueur de grille des transistors MOS est principalement motivée par la volonté d'augmenter la densité d'intégration des transistors sur une puce et par le désir d'augmenter leurs performances, par exemple en niveau de courant délivré ou bien en temps de propagation. Par loi d'échelle cette réduction de la longueur de grille entraîne des réductions de paramètres technologiques et électriques des transistors MOS comme l'épaisseur de l'oxyde de grille ou bien la tension nominale. Cette miniaturisation globale

entraîne aussi des effets néfastes sur certains paramètres électriques régissant le fonctionnement des transistors MOS. Pour contrecarrer ces effets, plusieurs solutions technologiques ont été proposées ces dernières années, soit en optimisant les architectures existantes, soit en proposant de nouvelles architectures. Afin de quantifier les avantages et les inconvénients de tels dispositifs ultracourts il est indispensable de pouvoir les

caractériser électriquement de façon efficace ainsi que de modéliser le comportement de leurs paramètres électriques avec la réduction des dimensions.

Au cours de ce travail nous nous sommes attaché à bien étudie et comprendre les effets indésirables qui sont dus à la miniaturisation des dispositifs CMOS, pour ensuite, proposer de nouveaux matériaux plus adéquat avec des propriétés physiques ou électriques plus confort pour essayer de minimiser les effets parasites dans les transistors CMOS ultrafin.

Dans le 1^{er} chapitre, après la présentation du transistor MOS et de son fonctionnement, nous décrirons les enjeux physiques régissant l'amélioration des performances. Toutefois, au cours de cette course effrénée vers la miniaturisation des composants électroniques, les problèmes technologiques ne se limitent plus aux seules difficultés de réalisation. Nous voyons apparaitre, des phénomènes parasites remettant en cause le bon fonctionnement du transistor L'objectif de cette lère partie est de présenter pourquoi il devient indispensable de faire évoluer le transistor MOS vers de nouvelles architectures et des nouveaux matériaux.

Dans le 2^{eme} chapitre, on présentera les semi-conducteurs III-V et leurs grand intérêt en raison de leurs propriétés, puis d'écrire les matériaux clés et leurs propriétés et enfin leurs structure cristallin.

Dans le dernier chapitre, les procédés d'apport de matériaux semi-conducteurs III-V, et dire quel est le procédés le plus utiliser dans cette technologies.

Chapitre I Le transistor MOS

I.1.INTRODUCTION

Avec l'évolution des circuits intègres à base de CMOS, et l'intégration de la microélectronique dans tous les domaines, de nouveaux défis sont parus à l'égard de l'exigence de performances à savoir la vitesse et la diminution d'énergie dissipée, ce qui nous amène vers la limite de la miniaturisation de MOS ou le maximum du SCALING (intégration du maximum de transistors sur un minimum d'espace) [1].

Pour atteindre cette objectif, il faut changer non seulement les procédés d'élaboration et l'architecture des dispositifs. Mais aussi introduire de nouveaux matériaux dans les filières microélectronique [2].

Dans ce but, l'industrie de la microélectronique essaie plusieurs matériaux en grille, en source et en drain des MOS.

Actuellement plusieurs architectures et plusieurs matériaux sont en compétition pour des composants de dimensions inférieurs à 15nm (longueur de grille) [2].

I.2. La Technologie MOS et son évolution

Depuis l'apparition en 1954 du transistor MOS, les technologies liées aux semi-conducteurs en général ont connu une révolution en termes de fonctionnalités basées sur la miniaturisation des dispositifs (cette miniaturisation est envisagée jusqu'à des dimensions de l'ordre d'une dizaine de nanomètres) avec l'accroissement de leur complexité.

Cette amélioration des performances des circuits intègres nécessite, à la fois une diminution de la taille caractéristique des dispositifs et une augmentation de la densité d'intégration (SCALING), ainsi le nombre d'interconnections doit considérablement augmenté pour des milliards de transistor qui pose aussi de grave problèmes : il y a 8 à 10 couches de connexion métalliques les unes sur les autres ,mais le temps de transmission des signaux sur fils devient trop grand par rapport a la fréquence exigée par des microprocesseurs « ce qui conduit à une consommation considérable d'énergie électrique, qui est comparable avec la densité de

chaleurs sur une minuscule puce de microprocesseur qui atteint celle d'un réacteurs nucléaire »[3].

Le transistor et le dispositif le plus répandu dans la production actuelle de composants semiconducteurs car il est le composant de base de la technologie CMOS qui englobe seule plus de 70% de la production mondiale de circuits intégrés [1].

I.2.1. Aperçu historique sur les transistors MOS

En 1930, L.Lilienfeld de l'université de LEIPZIG à déposé un brevet dans le quel il a d'écrit un élément qui ressemble au transistor MOS actuel, mais son apparition ne sera qu'en 1954 par TEXAS instrument qui à fabriqué le premier transistor en silicium en 1958, les transistors avaient une taille de l'ordre du cm³, et avec « F.wanalass » en 1963, la technologie MOS évolue vers CMOS (Complementary Metal oxide Silicium), et en 1964 les transistors atteignait la taille d'un grain de sel.

Toutefois, les MOS trouvent leur applications dans des circuits logiques et des circuits radio et vidéo [4].

En 1971, Intel a élaboré le premier microprocesseur, le « 4004 » qui possède plus de 2250 transistors sur une surface de 6cm².

En 2004, Intel à sorti le « pentium4EE » qui possède 169 millions de transistors sur une puce de surface de 1.12cm² [3].

Au delà des dispositifs réalisés en microélectronique, les transistors MOS joue un rôle central dans la technologie de silicium son large succès est du à sa faible consommation statique (faible dissipation de puissance), de son potentiel de réduction à des dimensions très petites (forte densité) et de ces performances en fréquence (rapidité) [4].

Depuis 30ans, l'industrie de la microélectronique suit la loi de Moore énoncée en 1965 par Gorden Moore, co-fondateur d'Intel, selon la quelle le nombre de transistor doit doubler, tous les 18mois .cette loi à peu prés vérifie en 1973.



Figure I.1 La croissance exponentielle : le nombre de transistors sur puce double dans certain période tous les 18mois

La dimension critique des dispositifs correspond à la largeur de la grille du transistor MOS. Aujourd'hui, les fabriquant de semi-conducteurs commencent à produire des transistors de largeur 65nm, et les générations suivantes de largeur 45nm et 32nm sont en développement [5].

I.2.2 Architecture des transistors MOS

Le transistor MOS (Metal oxyde semi-conducteur) est un condensateur constitué d'un substrat semi-conducteur (n ou p) séparé d'une électrode métallique (la grille) par une fine couche de diélectrique (le plus souvent c'est un oxyde Sio₂) (**FigI.2**). Les deux extrémités (source et drain) de ce condensateur MOS sont des zones dopées différemment de la zone du substrat mais elles sont des parties intégrantes de celui-ci. Dont elles diffèrent par leurs types de conduction. **Source** et **drain**. (\mathbf{n}^+ ou \mathbf{p}^+)



Figure I.2 Schéma d'un transistor MOS [3].

I.2.3 Fonctionnement des transistors MOS

Le principe de fonctionnement du transistor MOS est comparable à celui d'un robinet ou bien d'un interrupteur. En effet, le principe de fonctionnement d'un transistor MOS repose sur l'effet de champ de la grille qui consiste à moduler de façon électrostatique une densité de charges mobiles dans un semi-conducteur (le canal),cette modulation et provoquée par un champ électrique perpendiculaire (la grille) à la direction de mouvement de ces charges dans le canal, et agissant entre deux électrodes séparés par un diélectrique, comme dans une capacité plane [3]-[6].

Ce champ électrique est géré par une électrode de commande (la grille) à travers une couche isolante (le diélectrique de grille). Les charges mises en jeu sont des électrons dans le cas du NMOS et des trous dans le cas du PMOS. Le transport des charges s'effectue entre la source fournissant les porteurs et le drain qui les collecte sous la forme d'un courant. À travers le long d'une couche semi-conducteur (le canal) possédant deux contacts entre la source 'S' et

le drain 'D', entre lesquelles est appliqué une différence de potentiel. Les régions source, drain et canal sont des parties intégrantes du substrat dont la source et le drain diffèrent par leur type de conduction [2].

Le canal conduit plus ou moins de courant en fonction de son niveau de remplissage en charges mobiles. De ce fait, le transistor MOS peut aussi être une résistance modulable électriquement et reliant deux contacts (source et drain) [3]-[4]-[6].

I.3 Grandeurs caractéristiques des MOS

Le fonctionnement du transistor MOS comporte deux régimes distincts dont la séparation est communément définie comme le seuil du transistor. Lorsqu'on applique une tension V_g sur l'électrode de la grille inférieure à la tension de seuil V_{th} , le transistor est régime sous le seuil : autrement il est bloqué. Dans le cas où cette tension est supérieure à V_{th} le transistor est alors en régime d'inversion ou passant. En réalité, Vth représente la Vg à appliquer pour la déplétion du substrat et créer ainsi la couche d'inversion forte (la création du canal). Ce dernier correspond à une concentration de porteurs minoritaires en surface (électrons dans le cas du NMOS) qui est égale à la concentration de porteurs majoritaires dans le substrat (trous dans le cas du PMOS). La Figure I.3, présente les schémas d'un transistor NMOS pour différentes valeurs de la tension Vg par rapport à Vth [2]-[5].



Figure I.3 : schéma de la structure d'un transistor NMOS à enrichissement [6]-[7].

(a)- Vg <0<Vth : le passage de courant entre la source et le drain et bloqué.

(b)-0<Vg<Vth : les porteurs majoritaires (trous) sont repoussés de la surface, créant entre la source et le drain une zones de charge d'espace (ZCE) sub-surfacique ne comportant plus que des charges fixes négatives (les atomes accepteurs ionisés du substrat P). La surface est dite en déplétion.

(c)-Vg >Vth : une zone d'inversion (canal) se forme entre la source et le drain, encore appelée canal induit. Le courant circule entre la source et le drain, et la conduction est assurée par les électrons (conduction de type n).

I.4 : Effets de la miniaturisation des MOSFETs

Figure I.4 Diminution de longueur de grille des technologies CMOS (source : ITRS2001).

Depuis des décennies, le monde de la microélectronique s'évertue à réduire de plus en plus la dimension des transistors MOS. La technologie d'aujourd'hui permet la réalisation des transistors MOS avec des canaux ultracourts allant jusqu'à des longueurs déca nanométriques (voir figure I.4). Parmi les principaux avantages induits par la réduction d'échelle ont peut citer l'augmentation de la densité d'intégration, la réduction de coûts de fabrication, la réduction du temps de transit des porteurs dans le canal, la réduction de la consommation. Mais la réduction de la géométrie des transistors MOS entraîne aussi des modifications néfastes de certains paramètres électriques parmi lesquels : la diminution de la mobilité, la dépendance de la tension de seuil avec la longueur de canal, l'augmentation de la conductance de sortie etc.

Des lois de réduction d'échelle ont été proposées afin de minimiser les effets de canaux courts. Leur but est de garder le même niveau de champ électrique interne quelque soit les dimensions du transistor[4].

On présente ci-dessous une méthode qui autorise des facteurs de réductions d'échelle différents (λ et κ). Cette loi a été proposée pour les applications sub-0.25 μ m. Ces relations sont résumées sur le tableau 1 ci-dessous :

Paramètre	Expression	Facteur d'échelle
physique		
Dimensions	W, L, Tox, xj	$1/\lambda$
Potentiels	Φ_{G}, Φ_{D}	1/κ
Concentrations	N _A , N _D	
d'impuretés		λ^2/κ
Champ	E	
Electrique		λ/κ
Capacités	Cox, Cj	$1/\lambda$
Puissance	$I_D.V_{DD}$	$1/\kappa^2$

Tableau 1 Lois de réduction d'échelle [5].

De façon générale, le tableau 1 montre qu'il ne suffit pas de diminuer les dimensions géométriques du transistor (L et W) mais aussi les dimensions des couches technologiques de celui-ci : oxyde de grille, jonctions source et drain ...etc. Par exemple, pour un transistor de 50nm de longueur de grille, l'oxyde de grille ne doit pas être plus épais que 1.2nm, ce qui ne signifie que quelques monocouches de silice. Or une si fine couche d'isolant va fuir par effet tunnel et ne jouera plus son rôle d'isolant. C'est pour cela notamment que pour des générations encore plus avancées, le changement de la nature de l'oxyde de grille est envisagé pour aller vers des matériaux à haute permittivité électriques (High K), comme l'oxyde d'Hafnium (HfO2) qui permettront de garantir la même capacité d'oxyde mais avec des épaisseurs de la couche d'oxyde plus grandes afin de limiter les fuites à travers l'oxyde. Ce type de transistors n'ont pas fait l'objet d'étude de notre part.

I.4.1 : Origine des effets de canaux courts

Afin de comprendre l'origine des effets de canaux courts, nous reprenons ici une étude proposée par T. Nguyen et J. Plummer [8].

L'équation de Poisson, qui donne l'évolution du potentiel dans la structure s'exprime par :

du champ électrique Ex provient principalement de l'électrode de grille. La composante latérale Ey du champ électrique est originaire des jonctions de source et de drain. On peut associer le premier terme de l'équation à une densité de charge de grille qNg et le deuxième à une densité de charges de jonction qNj. La somme de qNg et qNj peut être interprétée comme la densité de charge totale.

Dans le cas des transistors à canaux long, la composante transversale de champ Ey peut être négligée. L'équation de Poisson se résume dans ce cas à sa forme unidimensionnelle :

$$qN_g = \rho$$
.....I.2

En réduisant les dimensions des transistors, la composante latérale du champ n'est plus négligeable et le terme qNj ne peut plus être ignoré. La charge n'est alors plus contrôlée uniquement par la grille. Comme la densité de charges de jonction est dépendante des dimensions du transistor et des polarisations appliquées, la tension de seuil dépend également de ces grandeurs.

La miniaturisation du canal change la distribution du potentiel qui passe d'une distribution unidimensionnelle à une distribution bi-dimensionnelle. On voit donc que l'approximation du canal graduel faite dans le cas des canaux longs n'est plus valable pour les transistors MOS à canaux courts. Par la suite, nous allons décrire les principales conséquences de cette distribution de potentiel bidimensionnelle.

I.4.2 : Effet du champ électrique sur la mobilité

La mobilité dépend des champs électriques longitudinal et transversal. C'est à dire la mobilité dépend aussi de la polarisation appliquée entre le drain et la source ainsi que de celle appliquée sur la grille. Plus on réduit les composants, plus l'impact de la composante longitudinale du champ sur la mobilité augmente.

Pour des tensions Vds faibles on peut négliger l'effet du champ électrique longitudinal. Le champ électrique transversal confine les électrons vers l'interface Si/SiO2. Les collisions en

surface déterminent une réduction de la mobilité. Pour décrire ce phénomène on utilise une formule empirique :

 μ_n^* représente la mobilité corrigée seulement de l'effet de champ transversal, μ_{n0} la valeur pour des champs transversaux faibles (Vg=Vt) et θ_G un paramètre empirique.

Le champ électrique longitudinal détermine une autre réduction de la mobilité. Pour des valeurs importantes de champ longitudinal on obtient la saturation de vitesse moyenne des porteurs. Cet effet apparaît vers le drain, où le champ atteint les plus grandes valeurs et augmente avec la réduction de longueur du canal. Pour tenir compte de cet effet on utilise une relation approchée de la vitesse des porteurs avec le champ électrique longitudinal.

$$\mu_{n} = \frac{\mu_{n}^{*}}{1 + \frac{\mu_{n}^{*}}{V_{s}}E_{y}}....I.4$$

Où μ_n est la mobilité corrigée des effets de champ transversal et longitudinal, V_s est la vitesse moyenne de saturation des porteurs.

Pour des valeurs faibles de E_y on obtient $\mu_n \approx \mu_n^*$ et pour Ey important on obtient la Saturation de vitesse des électrons, c'est à dire $\mu_n \approx V_s / E_y$.

$$\mu_n = \frac{\mu_{n_0}}{1 + \theta_G \left(V_g - V_t \right) + \theta_D \left(V_d - V_s \right)} \dots I.5$$

Où $\theta_D = \frac{\mu_{n_0}}{(V_s L)}$ considéré comme un paramètre empirique de la relation.

I.4.3 : Effet de la résistance série source – drain

La résistance de canal diminue avec la diminution de la longueur de canal. On ne peut plus négliger l'effet de résistance série côté source et drain Rs et Rd respectivement. Cela conduit en régime ohmique à une diminution de la tension effective appliquée entre la source et le drain du transistor intrinsèque (Vd's') par rapport à la tension appliquée aux électrodes :

Donc le courant de drain devient :

$$I_{d} = \left[\frac{G_{m}\left(V_{g} - V_{t}\right)}{1 + G_{m}\left(V_{g} - V_{t}\right)\left(R_{s} + R_{d}\right)}\right] V_{ds} \dots I.7$$

Où
$$G_m = \frac{\left(W\mu_0 C_{ox}\right)}{L}$$
.....I.8

Figure I.5 Schéma électrique équivalent du MOSFET - influence des résistances séries de source et de drain.

Par comparaison à un transistor idéal dans les mêmes conditions de polarisation, il y a une diminution de courant de drain. Cet effet est mieux mis en évidence par la diminution de la conductance de canal par rapport à celle du transistor idéal :

$$g_{0} = \frac{g_{0}}{1 + g_{0}(R_{s} + R_{d})}....I.9$$

La conductance de canal est définie par :

$$g_0 = \frac{dI_d}{dV_{ds}} \dots I.10$$

Il est donc possible, en régime ohmique, si on connaît la résistance série source-drain de calculer un courant de drain corrigé de l'influence de cette résistance série :

I.4.4 : Influence des zones de déplétion de source et de drain

L'approximation du canal graduel néglige les zones de déplétion source et drain Ws et Wd. Cette condition qui peut s'écrire Ws+Wd << Lg, n'est plus valable lorsque la longueur de grille Lg diminue et devient du même ordre de grandeur que Ws et Wd qui sont données par :

Où Vbi est la tension interne de jonction donnée par:

$$V_{bi} = \frac{KT}{q} \ln\left(\frac{N_a \cdot N_d}{n_i^2}\right).$$
....I.13

I.4.4 a) : Partage de charge

Une conséquence majeure du rapprochement des jonctions source et drain est la perte par la grille du contrôle d'une partie des charges situées à sa verticale. C'est ce que l'on nomme « partage de charge », ses répercussions sont importantes sur certains paramètres électriques, nous discuterons principalement de la tension de seuil de charge.

Un transistor MOS est constitué de deux jonctions de part et d'autre de la grille: la jonction source-substrat et drain-substrat. On discute le cas d'un transistor à canal de type n. Les jonctions mentionnées sont maintenues bloquées par les polarisations appliquées. Soit le substrat est au même potentiel que la source et seul le potentiel de diffusion Vbi de la jonction en assure le blocage, soit Vb est négative par rapport à la source (jonction en inverse) pour améliorer l'isolation électrique. La tension de drain est positive pour collecter les électrons du canal. La jonction drain-substrat voit une chute de potentiel égale à Vd-Vb-Vbi. Toute polarisation en inverse crée une zone de charge d'espace où la densité des porteurs libres est négligeable. Plus la polarisation est élevée, plus la zone de désertion s'étend du côté le moins dopé c'est-à-dire dans le substrat et sous la grille pour ce qui concerne la surface du silicium. Lorsque les extensions des zones de désertion de jonction sont de l'ordre de grandeur de la longueur de grille, la part de la surface du semi-conducteur contrôlée par la grille diminue. La charge positive du substrat P avant de former la couche d'inversion diminue ce qui se traduit par une réduction de la tension de seuil.

En résumé, avec la réduction de la longueur de canal, la charge Q n'est plus contrôlée seulement par la tension de grille, mais une bonne partie est générée par les tensions appliquées sur le drain et la source.

Figure I.6 L'effet de canal court sur la charge de substrat Q_{BT}.

La partie de charge Q_{BT} contrôlée par la tension de grille est notée Q_{BG} et les deux parties contrôlées par le drain et la source sont égales et notées Q_{BJ} (voir figure I.6).

On peut alors écrire :

$$\frac{Q_{BG}}{Q_{BT}} = \frac{1}{2} \left(1 + \frac{L}{L} \right) \dots I.14$$

Pour des dispositifs à canal long, L' \approx L et $Q_{BG} = Q_{BT}$. A la limite L'/ L << 10n obtient :

$$Q_{BG} = \frac{Q_{BT}}{2} \dots I.15$$

Ce qui montre la réduction d'un facteur deux du facteur de substrat (lois d'échelle). Dans le cas général, on utilise le modèle trapézoïdal et la tension de seuil s'écrit [9] :

Où $V_{FB} = \Phi_{ms} - Q_{tot}/C_{ox}$ est la tension de bandes plates,

$$W = \left[\left(\frac{2\varepsilon_s}{qN_a} \right) (V_s + 2\phi_f) \right]^{1/2} \dots I.17$$

$$K = \frac{\left(2\varepsilon_s q N_a\right)^{1/2}}{C_0}....I.18$$

Et Φ_f est le potentiel de Fermi dont l'expression est :

$$\phi_f = \frac{E_i - E_f}{q} = \frac{KT}{q} \ln\left(\frac{N_a}{n_i}\right)....I.19$$

En développant en série la racine carrée de l'équation (I.16), la simplification suivante est obtenue [5] :

$$V_{t} = V_{FB} + 2\phi_{f} + K \left(2\phi_{f} + V_{s}\right)^{1/2} \left(1 - \theta_{B} \left(V_{s} + 2\phi_{f}\right)^{1/2}\right) \dots I.20$$

Avec
$$\theta_B = \frac{1}{L} \left(\frac{2\varepsilon_s}{qN_A}\right)^{1/2}$$
.....I.21

En diminuant la longueur de grille, la partie de la charge contrôlée par les jonctions de source et de drain devient non négligeable devant celle effectivement contrôlée par la grille, ce qui induit une réduction de la tension de seuil. L'influence de Wd peut également être amplifiée quand Vd augmente (voir relation (I.12)).

Figure I.7 Exemple d'effet de canal court sur la tension de seuil.

La figure I.7 donne un exemple typique de l'effet du partage de charge sur la tension de seuil. Par la suite nous parlerons plus générale d'effet de canal court ou SCE (Short Channel Effect) lorsque nous étudierons la variation de la tension de seuil avec la longueur de grille.

I.4.4 b) : Effet de canal étroit

En tenant compte de l'effet de la diminution de largeur de canal on obtient une augmentation de la tension de seuil par l'effet d'accroissement de la charge QB qui se trouve sous les zones d'isolation en bord du canal (voir figure I.8). L'accroissement de la tension de seuil est directement proportionnel à l'augmentation de l'aire transversale de la zone désertée $(W \times L)$. Donc l'équation (I.16) devient [9] :

Figure. I.8 L'effet de canal étroit sur la charge contrôlée par la grille.

Donc, en diminuant la largeur de grille, nous obtenons une augmentation progressive de

la tension de seuil (voir figure I.9).

Figure I.9 Exemple d'effet de canal étroit sur la tension de seuil.

I.4.5 : Effet DIBL :

A forte polarisation de drain (Vd>Vd,sat), un autre phénomène devient important : c'est l'effet DIBL (Drain Induced Barrier Lowering). Il se traduit par une réduction de la hauteur de barrière source/substrat à fort Vd induisant également une diminution de la tension de seuil. Il en résulte une augmentation du courant de drain avec la tension de drain en régime de saturation. L'effet d'abaissement de la barrière de potentiel induit par le drain a été largement étudié durant les dernières décennies [7], [9]. Cependant, il est toujours d'actualité en raison de la réduction constante des dimensions des dispositifs.

Dans les MOSFETs à canal court, les zones de diffusion de source et de drain sont proches ce qui entraîne une pénétration importante du champ électrique du drain vers la source. La barrière de potentiel à la source peut donc être réduite en raison de cette influence du drain. La figure I.10 illustre ce phénomène [8]. L'importance de cet effet dépend, bien sûr, de la longueur de canal mais également de la profondeur de jonction ou encore du dopage. La conséquence de l'abaissement de la barrière de potentiel de la source est une injection d'électrons de la source entraînant une augmentation du courant de drain.

Dans le modèle de Grotjohn[10], il a été établi que l'augmentation du potentiel de surface, au premier ordre, peut être reliée à la polarisation de drain par la relation $\Delta \Psi_s = BV_{ds}$ où B est le coefficient de DIBL donné par :

Où ε_{si} et ε_{ox} sont les permittivités respectives du silicium et de l'oxyde de grille, η est un paramètre géométrique, L est la longueur du canal, et L* une longueur du canal en dessous de laquelle le perçage devient apparent.

Figure I.10 Schéma de la barrière de potentiel le long du canal [7].

Le DIBL peut être aussi modélisé au niveau électrique par une réduction de la tension de seuil en fonction de la tension appliquée sur le drain. La relation courante a été proposée par Grotjohn [10] :

Où V_{t_0} est la tension de seuil pour Vd proche de zéro.

Le paramètre λ est le paramètre de DIBL sont relié au coefficient B par [8] :

$$\lambda = B \frac{C_{ox} + C_d}{C_{ox}}....I.25$$

Ce modèle présente l'avantage de déterminer le paramètre DIBL λ sans avoir à ne mesurer une tension de seuil, évitant de la sorte les incertitudes dues à la définition de Vt. Ainsi, le DIBL est alors caractérisé par la variation de la tension de seuil selon la relation (I.24) et pour le mesurer, il suffit donc de mesurer le décalage Δ Vt.

Le courant Id est une fonction de Vd et de (Vg–Vt) de la faible jusqu'à la forte inversion. Il est facile de retrouver la relation suivante [8] :

$$g_{d} = g_{d_{0}} + \frac{\partial I_{d}}{\partial V_{t}} \frac{dV_{t}}{dV_{d}} = g_{d_{0}} + \lambda g_{m} \dots I.26$$

Où g_{d0} est la conductance de sortie en l'absence de DIBL et g_m est la transconductance. Dans la région de saturation g_{d0} s'annule, et l'équation (I.26) devient :

Où g_{msat} est la transconductance en régime de saturation.

L'équation précédente montre que le rapport de la conductance de sortie à la transconductance en régime de saturation doit mettre en évidence un plateau, qui permet d'extraire la valeur de λ .

Pour illustrer l'effet du DIBL sur les caractéristiques Id-Vg, il suffit de faire deux mesures, l'une en régime ohmique (Vd<<Vd,sat) et l'autre en saturation (Vd>Vd,sat) :

Figure I.11 Courant de drain en fonction de la tension de grille pour un transistor long (L=1µm) et un transistor ultracourt (L=50nm) à faible (50mV) et forte (1V) polarisation de drain.

La figure I.11 montre bien que pour un transistor court, la tension de seuil est plus faible à forte qu'à faible polarisation de drain, donc si on se place à une valeur de Vg fixe, le DIBL entraîne une hausse du courant.

I.4.6 : Perçage

Pour des tensions de drain élevées, les zones de déplétion de part et d'autre du canal peuvent se toucher, dans ce cas Ws + Wd = Lg. Cette situation extrême porte le nom de perçage (*punchthrough* en anglais). Les porteurs majoritaires de la source (les électrons dans le cas d'un transistor MOS de type N) peuvent être injectés directement dans le canal entièrement déplété et collectés par le drain.

Le phénomène est essentiellement lié à la hauteur de la barrière de potentiel entre la source et le drain à travers le volume du substrat. Il est fortement dépendant de l'extension des régions de déplétion sous le canal. Le punchthrough peut être minimisé par un dopage adéquat du canal susceptible de favoriser l'augmentation de la barrière de potentiel source substrat, par exemple par une implantation de canal rétrograde, ou en élaborant une technologie ground plane [10]. Le punchthrough est essentiellement un courant de diffusion qui prend la forme générale suivante [10] :

$$I_{PT} = I_0 \exp\left[\frac{q\left(\phi_b + V_b\right)}{KT}\right] \left[1 - \exp\left(-\frac{qV_d}{KT}\right)\right].$$
....I.28

Où Φ_b est la barrière de potentiel interne source substrat, lo une constante caractéristique du courant inverse de saturation de la jonction source substrat.

I.4.7 : Effets de porteurs chauds [8]

Les porteurs acquièrent de l'énergie grâce au champ électrique et en dissipent une partie au réseau cristallin par des collisions sur les phonons acoustiques et optiques. Lorsque le champ électrique longitudinal augmente, les porteurs gagnent plus d'énergie qu'ils n'en dissipent. En utilisant une distribution Maxwellienne, on peut dire que $T_p > T_r$, où T_p est la température correspondant aux porteurs et T_r la température thermique du réseau. Donc, sous l'action du champ électrique longitudinal les porteurs peuvent devenir des porteurs "chauds".

Le champ électrique et ainsi les effets des porteurs chauds seraient plus grands avec l'augmentation de potentiel entre drain et source et/ou avec la réduction du canal.

Une des conséquences principales des effets de porteurs chauds est la génération de paires électron-trou (voir figure I.12). Ce phénomène se produit lorsque les électrons ou les trous ont une énergie suffisante pour ioniser par impact les atomes du réseau. Nous pouvons distinguer deux stades d'ionisation par impact. L'ionisation primaire dont les porteurs du canal sont responsables : les trous générés vont constituer le courant de substrat alors que les électrons seront collectés par le drain (dans le cas d'un nMOSFET). Et l'ionisation secondaire pour laquelle une partie des porteurs créés par ionisation primaire constitue la source de cette ionisation secondaire. Un courant de grille en est issu.

Figure I.12 Diagramme illustrant les mécanismes d'ionisation par impact dans un MOSFET.

La compréhension des mécanismes d'ionisation par impact est nécessaire pour évaluer les situations de dégradation maximale due aux porteurs chauds selon deux principaux critères : la structure technologique et la polarisation de cette structure[11].
I.5 Conclusion

La miniaturisation des transistors MOS et plus particulièrement la diminution de la longueur de canal a permis d'augmenter la densité d'intégration et la vitesse de fonctionnement des circuits. Cette réduction des dimensions a engendré des phénomènes parasites (DIBL, modification de la tension de seuil, augmentation du phénomène de porteurs chauds,...) qui détériorent les caractéristiques courant-tension. Toutefois, les technologues ont imaginé des procédés de fabrication particuliers en vue de conserver ces caractéristiques, et d'introduire de nouveaux matériaux à haute mobilité comme les composés semi-conducteurs III-V qui feras l'étude dans le deuxième chapitre.

Chapitre II Les matériaux semiconducteurs III-V

II-1 Introduction

Après les spectaculaires résultats enregistrés dans la technologie Silicium dans le domaine de la microélectronique, la recherche de composants électroniques présentant des propriétés de transport supérieures à celles du Silicium, conduit la plupart des grands laboratoires de recherches à s'intéresser à des nouveaux matériaux et en particulier les composés semiconducteurs III-V. C'est alors qu'une nouvelle classe de matériaux fait son apparition, basés sur la création d'un potentiel dans un cristal par:

- \cdot La modulation de dopage.
- · La variation de composition d'un alliage.
- · La combinaison des deux méthodes précédentes.

La découverte d'hétérostructures à base de semi-conducteurs III-V a permis le développement de l'optoélectronique. Les nitrures d'éléments III-V (GaN, AlN, InN et leur alliages) sont des semi-conducteurs aux propriétés remarquables. La plus importante est leur bande interdite directe qui varie de 1,9 eV pour InN à 3,4 eV pour GaN (de 0.365 μ m à 0.653 μ m). Ils ont une large bande interdite, ce qui est particulièrement attrayant pour la physique et les applications optoélectroniques.

II-2 Différence entre isolant conducteur et semi-conducteur

Isolants : les bandes d'énergie les plus faibles sont entièrement pleines. La hauteur de la bande interdite est grande (\geq 5eV).il n'y a pas de niveaux d'énergie accessibles et pas de conduction[12].

Conducteurs : la dernière bande occupée est partiellement remplie, il existe beaucoup de niveaux disponibles et la conduction est grande.

Semi-conducteurs : le taux de remplissage de la dernière bande occupée est soit très faible soit très important. La hauteur de la bande interdite est faible. La conduction est faible est varie beaucoup avec la température.

II-3. Différents types de dopage II.3.1. Semi- conducteur dopé n

Matériau dopé n est un semi-conducteur dont la concentration en électrons est largement supérieure à la concentration en trous. On y a introduit généralement du phosphore, de l'Arsenic ou encore de l'Antimoine. Prenons par exemple le cas du Silicium dans lequel on introduit de l'Arsenic (possède 5 électrons sur la couche extérieure)[12].

On associe à l'Arsenic un niveau donneur Ed dans la bande interdite très proche de la bande de conduction. L'intérêt est que pour les températures supérieure à 0°K, les atomes d'arsenic sont ionisés ce qui revient à faire passer l'électron du niveau donneur à la bande de conduction. La concentration de donneurs sera donc supérieure à la concentration d'accepteurs (*Nd-Na* >0) ce qui correspond à la définition d'un semi-conducteur dopé n.

A la température ambiante pratiquement tous les donneurs sont ionisés et si la concentration en atomes donneurs est *Nd*, la densité de porteurs libres du semi-conducteur sera :

$$n = n0 + Nd \tag{II-1}$$

Où *n0* est la densité d'électrons générés par le processus de rupture de liaisons de covalence qui engendre des paires électrons-trous.

II.3.2. Semi-conducteur dopé p

Il s'agit d'un semi-conducteur dont la concentration en trous est largement supérieure à la concentration en électrons. On y a introduit généralement du Bore, de l'Aluminium, du Gallium ou encore de l'Indium, prenons le cas du Silicium dans lequel on introduit du Bore. On associe au Bore un niveau accepteur *Ea* dans la bande interdite très proche de la bande de valence [12].

De la même façon que pour le semi-conducteur dopé *n*, la concentration en trous du semiconducteur sera :

$$p = p0 + Na \tag{II-2}$$

II.4. Semi-conducteurs usuels

Les semi-conducteurs usuels sont essentiellement constitués d'éléments de la colonne IV et des colonnes voisines du tableau périodique de Mendeleïev. On en distingue plusieurs types : ·Les semi-conducteurs élémentaires sont des cristaux constitués d'un seul élément chimique. On rencontre des structures cristallines dites "simples" si l'élément constituant est de la colonne IV1. Les deux éléments les plus importants pour l'électronique sont le silicium (Si) et le germanium (Ge), qui se cristallisent en engageant des liaisons covalentes. Le premier est l'élément le plus utilisé dans l'industrie des composants.

Le germanium, quant à lui, à été largement utilise lors de fabrication des premières diodes et des premiers transistors, mais a été ensuite remplacé par le silicium. Il est néanmoins utilisé dans quelques applications (détection infrarouge, hétéro structures, . . .)[13].

On rencontre également des semi-conducteurs dits "élémentaires complexes" comme le sélénium (Se, colonne VI), notamment employé pour ses propriétés photovoltaïques. Cet élément se cristallise selon une structure différente de celle du Si, mais toujours avec des liaisons à prédominance covalente.

Les semi-conducteurs composés sont constitue de plusieurs éléments par exemple, les composés binaires peuvent être constitues de deux éléments distincts de la colonne IV (SiC, SiGe), d'éléments des colonnes III et V (composés III-V tels que le (GaAs et le GaN) ou encore des éléments de colonnes II et VI (composés II-VI tels que le ZnS et le CdS). Les composés GaAs, GaP et GaN sont fréquemment utilisé dans la fabrication de diodes électroluminescentes. Enfin, on peut également trouver des alliages de types ternaire (AlGaAs) et quaternaire (GaInAsP), employés par exemple dans la fabrication de diodes lasers[13].

Ces matériaux sont des semi-conducteurs solides inorganiques. On peut également rencontrer des semi-conducteurs organiques, notamment utilisés pour la fabrication d'écrans souples.

II.5. Définition des semi-conducteurs III-V

Les matériaux semi-conducteurs III-V sont des corps composés formés à partir d'un élément de la III (ème) colonne et d'un élément de la V (éme) colonne de la classification périodique de Mendeleïev. Le tableau II-1 regroupe un extrait de cette classification (les chiffres en haut et bas représentent respectivement le nombre atomique et la masse atomique). Ainsi de nombreux composés binaires peuvent être réalisés [17].

On peut obtenir un semi-conducteur si la somme des électrons des deux espèces est égale à 8 électrons.

III	IV	V
⁵ _{10,81} B	⁶ _{12,01} C	⁷ _{14,01} <i>N</i>
$^{13}_{26,98}Al$	¹⁴ _{28,09} Si	$^{15}_{30,97}P$
³¹ _{69,74} Ga	³² _{72,59} Ge	³³ _{74,92} As
⁴⁹ _{114.82} In	⁵⁰ _{118,69} Sn	$^{51}_{121,75}Sb$

 Tableau II-1 Extrait de la classification périodique des éléments

Le tableau II-1 montre une représentation simplifie du tableau périodique de Mendeleïev permettant de connaître les composés et alliages possible.

Par exemple :

- Ga : le nombre d'électron de valence est 3 électrons.

- As : le nombre d'électron de valence est 5 électrons.

Donc :

Ga + As = GaAs: est un semi-conducteur car la somme des électrons des deux espèces (Ga et As) est égale à 08 électrons. Telle que (GaAs) il satisfait à la condition du Bragg (2d sinq = n l). Qui représente la loi de déplacement des électrons.

II.6. Les composés binaires, ternaires et quaternaires des S/C III/V

II.6.1. Les composés binaires

Parmi tous les composés binaires possibles, tous n'ont pas le même intérêt potentiel. L'étude de leurs propriétés, et en particulier de la structure de bandes montre que les éléments les plus légers donnent des composés dont laquelle la bande interdite est large et indirecte, et dans laquelle la masse effective des électrons est élevée.

Les composés contenant du bore, de l'aluminium ou de l'azote entrent dans cette catégorie; ils ont en général peu d'intérêt pour l'électronique rapide [14], qui demande des semi-conducteurs à forte mobilité de porteurs ou pour l'optoélectronique ou une structure de bande directe est nécessaire pour que les transitions optiques soient efficaces [17]. A l'autre extrémité, les éléments lourds donnent des composés à base de Galium (GaAs, GaSb) ou d'indium (InP, InAs, InSb) dont les propriétés sont les plus intéressantes. Le tableau II-2 résume quelques paramètres pour différents matériaux de la famille III-V.

Composé III-V	Eg (eV)	m*/m0	μ (cm ² /V.s)	a (Å)
Al P	2,45			5,4510
Al As	2,16			5,6605
Al Sb	1,58	0,12	200	6,1355
BP	2,0			4,5380
Ga P	2,26	0,82	110	5,4512
Ga As	1,42	0,067	8500	5,6533
In P	1,35	0,077	4600	5,8686
In As	0,36	0,023	33000	6,0584
In Sb	0,17	0,0145	80000	6,4794

Tableau II-2 Propriété des principaux composés binaires III-V à 300 K [14,17]

II.6.2. Les composés ternaires et quaternaires

L'intérêt pratique des semi-conducteurs III-V est encore considérablement renforcé par la possibilité de réaliser des alliages par substitution partielle de l'un des éléments par un autre élément de la même colonne. On sait par exemple obtenir des alliages ternaires, ou quaternaires qui sont identifié de la façon suivante :

Ternaires : S'il y a substitution de 2 atomes sur l'un des sous réseaux, soit, $A_x A'_{(1-x)} B$ Exemple: $Ga_x In_{(1-x)} P$, lorsque la composition exacte compte peu, on écrit tout court GaInP.

Quaternaires 1+3 : S'il y a substitution de 3 atomes sur des sous réseaux soit :

 $A_x A_y A_{(1-x-y)} B$. Exemple. $Ga_x In_y Al_{(1-x-y)} As$

Quaternaires 2+2 : S'il y a substitution de 2 atomes sur chacun des deux sous réseaux, soit $A_x A_{(1-x)} B_y B_{(1-y)}$. Exemple $Ga_x In_{(1-x)} P_y As_{(1-y)}$.

La plupart des solutions solides ainsi réalisées sont complètes, la loi de Vegard (relation linéaire entre le paramètre de réseau et la composition) est approximativement suivie, et on observe une évolution progressive et régulière des propriétés (dont la bande interdite et les paramètres cristallins) en fonction du taux de substitution.

II.7. Les matériaux massifs

II.7.1. Propriétés cristallines

Les matériaux semi-conducteurs sont exclusivement des solides cristallins. Leurs propriétés électroniques, optiques, mécaniques et thermiques sont donc liées à la géométrie particulière qui représente l'arrangement des atomes au sein du matériau : le réseau cristallin.

Les semi-conducteurs III-V tels que GaAs (arséniure de gallium), AlAs, AlSb, etc... Ont une structure cristallographique de type blende de zinc. Cette structure, qui s'apparente à celle du diamant (Si, Ge, etc...), est constituée de deux sous-réseaux cubique à faces centrées, l'un d'éléments III, l'autre d'éléments V, décalé l'un par rapport à l'autre du quart de la diagonale principale (voir la figure I.1(a), l'atome en rouge représente un élément III par exemple, l'atome en jaune représente un élément V). La maille élémentaire comporte deux atomes, le premier (Ga) à l'origine et l'autre (As) à (a/4,a/4,a/4), où a représente le paramètre de maille du matériau.

Du fait que les deux sous-réseaux cubiques à faces centrées sont décalés, le cristal n'est pas centrosymétrique. Il en découle des propriétés physiques différentes suivant les directions cristallographiques considérées.

De ce fait, les liaisons atomiques dans les matériaux III-V ne sont pas simplement covalentes comme dans le cas du silicium. Elles reposent sur le transfert d'électrons des atomes du groupe V (As) sur ceux du groupe III (Ga). Dans un cristal composé de l'arséniure de gallium, chaque atome d'arséniure est entouré de quatre atomes de gallium, et chaque atome de

gallium est entouré de quatre atomes d'arséniure. Il se produit alors un échange d'électrons, le cristal se construit avec les ions Ga- et As+, qui ont tous quatre électrons périphériques. Cette répartition est à l'origine du caractère partiellement ionique et partiellement covalent des liaisons (semi-conducteurs polaires) qui sont orientées dans l'espace suivant les axes de symétrie d'un tétraèdre régulier. Cette propriété est en particulier responsable de l'anisotropie optique dans les hétéro structures semi-conductrices III-V qui a été mise en évidence expérimentalement par Krebs et Voisin [14]-[16].



Figure II-1 (a) La structure blende de zinc et (b) sa première zone de Brillouin.

Le réseau réciproque du réseau de Bravais correspondant à la structure blende de zinc est un réseau cubique centré. La première zone de Brillouin du réseau réciproque a la forme octaèdre tronqué (voir figure II.1(b)) par les six faces d'un cube. Elle présente un centre de symétrie à l'origine noté Γ et les axes de symétrie Δ , λ , et γ .Les points de croisement de chacun de ces axes avec les frontières de la zone de Brillouin sont les points de haute symétrie, et ils jouent un rôle primordial dans la structure de bande.

II.7.2. Propriétés électroniques

Les matériaux de semi-conducteurs III-V ont huit électrons par cellule unitaire contribuant aux liaisons chimiques. Les autres électrons n'interviennent pas dans les propriétés optiques des hétéro structures. Les orbitales de type \mathbf{s} et de type \mathbf{p} de chaque atome de gallium Ga s'hybrident avec les orbitales des atomes d'arséniure As, et forment des liaisons covalentes tétraédriques de type **sp3 :** 4 orbitales liantes et 4 orbitales antiliantes. Les quatre orbitales liantes donnent lieu à quatre bandes d'énergie, chacune deux fois dégénérée de spin, et forment la bande de valence. Cette bande est pleinement occupée par des électrons à T = 0K, pour un semi-conducteur parfait. Les quatre autres orbitales antiliantes donnent naissance à quatre bandes supérieures, et forment la bande de conduction qui est inoccupée et est séparée de celle précédente par une bande d'énergie interdite de largeur Eg (band gap en anglais). Pour les semi-conducteurs à gap direct, le maximum de la bande de valence et le minimum de la bande de conduction sont au point Γ .

II.8. Propriétés du GaAs

Comme tous les semi-conducteurs composés III-V, le GaAs cristallise dans une structure de type zinc-blende. Il est formé de deux réseaux cubiques à faces centrées identiques qui s'interpénètrent, l'un contenant les atomes de gallium 'Ga', l'autre les atomes de l'arsenic 'As'. Le décalage entre les deux réseaux est d'un quart de la diagonale principale du cube. On a donc quatre paires d'atomes par cellule conventionnelle.

Les liaisons tétraédriques covalentes telles que les liaisons Si-Si par exemple ou partiellement ioniques et partiellement covalentes telles que les liaisons Ga-As dans le cas qui nous concerne sont la base de la structure zinc-blende. En effet, chaque atome est l'origine de quatre liaisons orientées dans l'espace suivant les axes de symétrie d'un tétraèdre régulier.

Une autre propriété chimique importante du GaAs est que si l'on clive (coupe) le cristal selon la direction [001], on aura une succession de plans cristallins comportant des atomes de l'élément de type III, puis les atomes de l'élément de type V, et ainsi de suite. Cela permet une croissance couche à couche du semi-conducteur.

D'un point de vue énergétique, la relation de dispersion E(k), présentant l'énergie en fonction du vecteur d'onde, montre que contrairement au silicium et au germanium où la transition est indirecte, le GaAs est un matériau à transition directe (Figure II.2)

Eg = 1,424 eV à 300 K [12] [15] [16]. Par comparaison, les valeurs des bandes interdites correspondant au Si et au Ge sont de 1,12 eV et 0,66 eV respectivement.



Figure II-2 Bandes interdites à transition directe et indirecte dans les semi conducteurs. À gauche transition directe (cas du GaAs). À droite transition indirecte (cas du Si).

II.9. Structure des bandes d'énergie des semi-conducteurs III - V

Les bandes d'énergie donnent les états d'énergie possibles pour les élections en fonction de leurs vecteurs d'onde. On les représente donc dans l'espace réciproque, et pour simplifier, dans les directions de plus hautes symétries de la première zone de Brillouin.

Elles ne décomposent en bandes de valence la plus haute, la bande de conduction la plus basse, et la bande interdite qui les sépare qui détermine principalement les propriétés de transport du semi-conducteur.

II.9.1. Structure de bande de GaAs

Pour le GaAs, le maximum de la bande de valence est en K=0 et celle-ci est doublement dégénérée avec deux branches E=f(K) confondues au centre de la zone de Brillouin.

Les deux bandes décrivent la disposition de porteurs de masses effectives différentes, la masse des trous lourds est de 0.68m0, celle des trous légers de 0.12m0 où m0 est la masse de l'électron au repos.

Une troisième branche de la bande de valence est abaissée par rapport aux deux premières d'une énergie Δ , égale à 0,33 ev.

La bande de conduction présente, au centre de la zone de Brillouin (K=0) un minimum non dégénérée, qui est une vallée à symétrie (G) appelée vallée centrale. Il existe également d'autres minima à des niveaux d'énergie plus élevée, les 4 vallées à symétrie L dans la direction <111> et les directions équivalentes.

D'après la courbure des différentes vallées, à la quelle est reliée la masse effective des léger et donc beaucoup plus légers et donc beaucoup plus mobiles que dans les vallées X et L. Cette propriété est en partie responsable de certaines caractéristiques des phénomènes de transport dans le semi-conducteur GaAs.

L'arséniure de gallium étant un composé à gap direct, les transitions électriques les moins énergétiques entre la bande de valence de la bande de conduction peuvent se faire avec conservation du vecteur d'onde K (sans l'intervention des phonons).

L'énergie correspondante à la largeur de bande interdite ΔEg au centre de la zone de Brillouin varie avec la température.

Eg = 1.43ev à 300K et 1.51ev à 77K.



Figure II-3 Structure de bande d'énergie directe pour le GaAs [12]

II.10. Bandes interdites des composés III-V

II.10.1. Bandes interdites des principaux composés III-V

La figure II-4 montre les bandes interdites Eg des principaux composés III-V en fonction des paramètres cristallins a_0 [12]-[16]. En continu bandes interdites directes et en tirets bandes interdites indirectes; les discontinuités sont dues au changement de structures.



Figure II-4 Bandes interdites des principaux composés III-V en fonction des paramètres cristallins [12].

Le diagramme de la figure II-4 montre l'évolution de l'énergie de bande interdite des alliages de composés III-V en fonction du paramètre cristallin "a" qui varie lui même avec la composition. Les points du graphe montrent la position des composés binaires, et les lignes représentent l'évolution du gap "Eg" et du paramètre cristallin "a" en fonction de la composition des alliages ternaires et quaternaires. Certaines lignes présentent un point anguleux qui dénote une transition entre un gap direct et un gap indirect. Ce diagramme est donc très important parce qu'il permet de connaître la composition de tout alliage ternaire ou quaternaire susceptible d'être déposé en couche mince, par épitaxie, sur un substrat binaire comme GaAs ou InP, afin d'obtenir le gap désiré. Les matériaux III-V offrent donc une grande variété d'alliages permettant de moduler leurs propriétés électroniques.

Ce diagramme montre qu'il est possible d'obtenir des matériaux dont la largeur de bande interdite est variable, et donc les propriétés optiques, varient dans une large gamme.

Il existe cependant une contrainte importante pour la fabrication de ces matériaux, qui sont réalisés en couches minces par croissance épitaxie sur un substrat binaire, le paramètre cristallin doit être très proche de celui du substrat. Le diagramme de la figure II-4 permet de connaître la composition de tout alliage ternaire susceptible d'être épitaxie en couche mince sur des substrats binaires.

Chapitre III La fabrication collective des composants et des circuits intégrés

III.1 Introduction

Le but de ce chapitre est d'expliquer de manière très simplifiée les procédés utilisés classiquement pour fabriquer les composants semi-conducteurs et d'indiquer les évolutions possibles. L'industrie de la micro-électronique est arrivée à un degré de maturité élevé avec la fabrication de circuits intégrés comportant des millions de transistors de tailles sub-micrométriques tout en garantissant des rendements de fabrication supérieurs à 80%. Des ouvrages entiers sont consacrés à la description des procédés de fabrication qui sont très nombreux et mettent en jeu un grand nombre de paramètres. Ce chapitre est une simple introduction permettant de classer et de définir les procédés principaux. Le flot de fabrication de taille du transistor avec des longueurs de grille attendues très inférieures au micron conduit à une complexité croissante des équipements de fabrication et à des exigences draconiennes en terme de propreté et de contrôle des impuretés. Ces contraintes entraînent une croissance exponentielle du coût des équipements. De nouvelles techniques sont donc investiguées pour limiter cette explosion des coûts comme les techniques de nano lithographie et les techniques d'auto-assemblage. Le chapitre est organisé de la manière suivante :

III.2. La lithographie

- III.3. Les procédés de retrait de matériaux
- III.4. Les procédés d'apport de matériaux
- III.5. Les procédés alternatifs

III.2 La lithographie

Le principe général de fabrication collective des composants consiste à créer sur une tranche de semi-conducteur les zones dopées, les zones d'oxyde et les zones de contact nécessaires pour faire un composant donné. Un grand nombre de composants peuvent être fabriqués en même temps puisque la surface d'un composant élémentaire est très petite devant la surface de la tranche de semi-conducteur. Cette tranche épaisse de 1 mm environ mais de diamètre pouvant atteindre 300 mm est appelée wafer. On peut donc fabriquer simultanément des composants élémentaires (diodes, transistors, diode laser...). Il est également possible de fabriquer des ensembles de composants interconnectés appelés circuits. En pratique l'industrie fabrique sur un wafer de manière collective des milliers de circuits, chaque circuit pouvant comporter des millions de composants élémentaires.

Pour fabriquer des composants à des endroits précis sur le wafer, il faut construire des régions présentant des propriétés différentes dans le semi-conducteur. Certaines sont fortement dopées, d'autres faiblement dopées et d'autres isolantes. De plus, il faudra réaliser des connexions conductrices entre zones. Ces régions seront réalisées sur le support silicium de base en ajoutant des dopants, en déposant du métal (aluminium ou cuivre) ou de l'oxyde. Dans tous les cas l'opération devra être strictement limitée aux régions choisies sur le wafer. La lithographie permet de faire ce choix. C'est la technique de base de la fabrication collective des composants.



Figure III-1 Principe de la fabrication par lithographie

Imaginons le cas simple de la fabrication d'une zone dopée n sur un wafer dopé p. La figure III.1 représente les étapes. Reprenons les étapes du procédé. Dans une première phase, une résine est déposée sur le wafer. Dans une seconde étape, un rayonnement est envoyé sur la résine avec une modulation spatiale représentant le motif à réaliser. La façon de réaliser cette modulation sera vue dans la suite. Dans les procédés industriels, le rayonnement est généralement de la lumière et principalement des ultraviolets. Les régions exposées de la résine voient une modification de leurs propriétés chimiques. Elles seront par exemple plus résistantes à l'action d'un solvant. La quatrième étape est la dissolution des régions non exposées de la résine. L'inverse est également possible et dans ce cas, les régions insolées sont dissoutes plus facilement. Dans une cinquième étape, on fait diffuser les ions devant être inclus pour doper le silicium (du phosphore pour un dopage de type n) dans les régions non recouvertes de résine. Les techniques permettant cette opération seront étudiées ultérieurement. Enfin, et c'est la sixième étape, on enlève la résine pour obtenir le wafer structuré demandé. Les régions grisées sont dopées n et peuvent, par exemple, constituer les puits des transistors PMOS.

Il nous faut maintenant étudier comment moduler la lumière. Deux méthodes sont utilisées : les méthodes parallèles et les méthodes séquentielles. Le principe des méthodes séquentielles est de graver tous les motifs en même temps en se servant d'un masque. Le masque est un objet interposé entre la source de lumière et le wafer qui laisse passer la lumière de manière sélective. Le principe des méthodes séquentielles est d'inscrire point par point le motif à reproduire en guidant le faisceau lumineux. Les deux méthodes sont illustrées figure III .2. Quand le faisceau est un faisceau d'électrons et non pas de lumière, la lithographie est dite à faisceau d'électrons et est appelée lithographie « e-beam » [20].



Figure III-2 Méthodes parallèle et séquentielle.

Dans la lithographie parallèle, un masque interpose entre la source de lumière et le Wafer arrête les rayons lumineux aux endroits définis par le design. Il est constitue d'une plaque de quartz transparente aux photons recouverte d'une couche de chrome gravée en fonction des motifs à réaliser. Ce masque très précis est en général fabrique par une méthode de gravure faisant usage de la lithographie a faisceau d'électrons permettant des précisions bien inferieures au micron. Cela explique cependant le cout élève d'un jeu de masques.

Dans sa réalisation la plus simple, la lithographie par contact consiste à placer le masque au contact avec le wafer comme le montre la figure III.3.



Méthode parallèle Figure III.3 Lithographie par contact

Cette technique très simple a été utilisée jusqu'au milieu des années 70. Elle est limitée par la diffraction de Fresnel qui prédit que le motif le plus fin (a_{min}) réalisable est donne par :

$$a_{\min} = \frac{3}{2} \sqrt{\lambda(s + \frac{e}{2})}$$
III.1

Dans cette relation, λ est la longueur d'onde de la lumière utilisée pour insoler et e est l'épaisseur le la résine. La distance s entre le wafer et le masque est minimisée mais non nulle.

Avec une longueur d'onde de 400 nm et une distance de 10 microns entre wafer et masque, on atteint une résolution de 3 microns environ. Cette technique est limitée fondamentalement par la contrainte de planéité du wafer ce qui fixe la distance s minimale a 10 microns. La lithographie par contact n'est plus utilisée de nos jours de manière importante car sa résolution est insuffisante.

La lithographie par projection s'est imposée au fil du temps pour résoudre les problèmes de planéité évoques précédemment. Elle consiste à interposer un objectif photographique entre le masque et le wafer comme le montre la figure III-4.

Elle est très contraignante pour l'optique qui doit être de grande qualité. On considère alors le masque comme une source étendue. L'image de cette source est formée sur le wafer par

l'optique avec un facteur de réduction dépendant de la distance focale de l'optique. La séparation minimale de deux objets est donnée par la relation classique.

$$a_{\min} = \frac{K\lambda}{n\sin i}$$
 III.2

Dans cette relation de base, λ est la longueur d'onde de la lumière. L'indice du milieu est n et i est l'angle maximum de collection de la lumière défini sur la figure III.4. Le coefficient k a une valeur théorique de 0.61 mais est en pratique égal à 0.8 pour des installations classiques. Le produit (n.sini), appelé ouverture numérique, est passé de 0.3 à 0.9 avec des progrès constants dans la conception des optiques. Pour améliorer la précision de la lithographie, il est donc possible de diminuer la longueur d'onde. C'est le sens de la lithographie UV et à plus long terme de la lithographie à base de rayons X. Il est également possible avec une source UV d'utiliser des techniques plus sophistiquées comme les masques à contraste de phase ou la lithographie à immersion.



Méthode parallèle

Figure III.4 Lithographie par projection.

Il ne faudrait pas croire qu'un wafer entier puisse être insolé de cette manière. La zone utile appelée champ est bien plus petite que la surface du wafer. La solution généralement mise en œuvre est d'insoler une zone du wafer correspondant au champ puis de déplacer le wafer pour

insoler une autre zone et ainsi de suite. Cette technique permet d'appliquer un facteur de grandissement G variant entre 5 et 20. Le motif sur le masque peut être G fois plus grand que le motif gravé sur le wafer. Le masque est donc plus facile à fabriquer que s'il était à la même échelle. Si le facteur de grandissement est élevé, il faut déplacer le wafer un grand nombre de fois pour une insolation totale du wafer. Le temps d'insolation sera donc plus important. Il y a un choix optimal à faire en fonction de la résolution nécessaire et du coût de l'opération.

En réalité, quand on fabrique un circuit on utilise non pas un masque mais un jeu de masques car les opérations à effectuer sont nombreuses. Le coût d'un jeu de masques dans une technologie avancée est un véritable problème pour l'industrie micro-électronique. Dans une technologie 50 nm des coûts de 4 millions d'euros sont annoncés pour réaliser un circuit intégré.

La précision de l'alignement des masques est un facteur décisif pour la fabrication collective de composants miniaturisés. La figure III-5 montre les erreurs possibles en cas de mauvais alignement. En pratique les masques sont positionnés les uns par rapport aux autres à l'aide de motifs spéciaux.



Figure III.5 Alignement des masques[20].

III.3. Les procédés de retrait de matériaux

Ils permettent d'enlever de la matière dans des zones définies par la lithographie. Trois procédés sont possibles : la gravure humide, la gravure sèche et la gravure ionique réactive.

III.3.1. La gravure humide

Prenons l'exemple de la gravure d'une tranchée dans le silicium. Les autres zones sont protégées par la résine. Le substrat est trempé dans un bain chimique attaquant le silicium mais pas la résine. Une zone gravée se forme d'autant plus profonde que le temps d'immersion est élevé. Des vitesses de gravure de plusieurs microns par minute sont possibles en fonction de la concentration de la solution d'attaque. La figure III.6 illustre ce procédé de base. La gravure est un procédé simple mais présente l'inconvénient de la sous-gravure c'est-à-dire la gravure sous la résine. Ce procédé est donc peu adapté à l'obtention de motifs fins et à l'obtention de flancs de gravure raides.

L'orientation du réseau cristallin change les vitesses de gravure. Les plans les plus denses du silicium sont par exemple gravés beaucoup plus lentement que les autres régions. Il est donc possible en orientant convenablement le wafer de réaliser des gravures non isotropes comme le montre la figure III.6.

On peut graver le silicium mais aussi les oxydes et les métaux comme le montre le tableau cidessous.

Matériaux à graver	Agents de gravure
	Soude
Silicium	
	(KOH)
Dioxyde de silicium	Acide fluoridrique
	(HF)
résine	$H_2SO_4+H_2O_2$



Figure III.6 Les procédés de gravure humide[21].

III.3.2. La gravure sèche

Le principe n'est plus de contrôler une réaction chimique dans un bain mais de bombarder les zones non protégées de la surface avec des ions. L'opération se passe dans une enceinte sous vide comme le montre la figure III.7. Trois techniques sont alors possibles : la gravure dite par « sputtering », la gravure par plasma, la gravure ionique réactive (RIE).

Ces trois techniques font usage de deux principes physiques : le premier est l'arrachement d'un atome de la surface par collision élastique avec un ion incident (sputtering), le second est l'activité chimique en surface du wafer quand un plasma est créé dans l'enceinte à vide. Ce plasma est créé par un champ radiofréquence à 13,56 MHz.

La gravure par sputtering utilise le premier effet ; elle est purement mécanique. La gravure par plasma utilise le deuxième effet ; elle est purement chimique. La gravure ionique réactive combine les deux effets et permet ainsi de décaper la surface avec une grande efficacité. Elle est donc très largement mise en œuvre dans les procédés actuels de fabrication. Elle permet

également d'atteindre d'excellentes résolutions de gravure (quelques dizaines de nm) et de réaliser des flancs quasi-verticaux. Les trois méthodes sont illustrées figure III.7.



Technologie RIE

Figure III-7 Les trois méthodes de gravure sèche.

Certains réacteurs sont configurés en mode « remote plasma ». Le plasma est généré à une distance assez grande de la plaquette. La gravure est alors essentiellement chimique, par les radicaux réactifs qui diffusent jusqu'à la plaquette.

III.4. Les procédés d'apport de matériaux

Il s'agit non plus d'enlever un matériau existant mais d'amener dans des régions définies par la lithographie les matériaux nécessaires à la fabrication des composants. Les matériaux sont du silicium, des isolants (dioxyde de silicium ou nitrure de silicium) et des métaux (aluminium, cuivre, cobalt, titane et tungstène). Ce sont aussi des ions (Bore, arsenic et phosphore) qui sont ajoutés au silicium pour modifier le dopage.

Les procédés sont assez nombreux : implantation ionique, PVD, CVD, croissance thermique, croissance électrolytique. Ils sont classés dans la figure III.8 selon trois types d'apport. Les procédés de dépôt consistent à amener sur le wafer des matériaux existants. Les dépôts sont alors des couches minces (épaisseur inférieure au micron). Les procédés de croissance consistent à déclencher la croissance du matériau à partir de ses constituants chimiques. Enfin, les procédés d'apport en profondeur consistent à introduire des atomes dans le wafer pour changer les propriétés électriques.



Figure III.8 Les procédés d'apport de matériaux [20].

III.4.1. Evaporation

C'est un procédé très simple qui consiste à chauffer un matériau dans un creuset, à le transformer en phase vapeur puis à provoquer la croissance du même matériau en phase solide sur les zones non protégées du wafer. Ce sont les métaux qui sont principalement déposés par cette technique car le point de fusion est relativement bas. Le matériau peut être chauffé par une résistance mais aussi par un faisceau d'électrons quand il est nécessaire d'atteindre des températures plus importantes. Les taux de déposition sont élevés, jusqu'à

5000 nm par minute. Ce procédé assez simple est mis en œuvre dans de nombreux laboratoires de recherche.



Figure III.9 Évaporation.

III.4.2. Epitaxie par jets moléculaires (MBE)

C'est une amélioration de la technique d'évaporation qui bénéficie des techniques d'ultra vide. Différentes sources sont présentes dans l'enceinte sous vide. Des obturateurs rapides sont placés devant chaque source et des équipements de caractérisation sont en général ajoutés pour contrôler les dépôts. Un équipement type est symbolisé figure III.10



Figure III.10 Épitaxie par jets moléculaires.

Cette technique permet en particulier de réaliser des dépôts multicouches. Des couches très fines (quelques couches atomiques) et monocristallines peuvent être fabriquées par cette technique. La méthode est donc toute indiquée pour réaliser des structures multicouches. Le procédé est assez long (un micron par heure environ) et est principalement utilisé dans les laboratoires de recherche. Des systèmes chauffant ou des canons à électrons peuvent être utilisés pour réaliser les dépôts [22].

III.4.3. Dépôt par « sputtering »

Cette technique est principalement utilisée pour déposer des contacts métalliques. Elle a remplacé d'année en année la méthode par évaporation car les taux de contamination peuvent être très faibles. L'équipement est le même que celui de la gravure par « sputtering » à la différence près que l'anode est le wafer et que la cathode est une cible faite dans le matériau qu'il faut déposer. Un gaz inerte comme l'argon est ionisé et les ions sont accélérés vers la cible. Des atomes de la cible sont éjectés et vont se déposer sur le wafer. Le sputtering RF peut également être utilisé pour déposer des isolants [21].



Figure III.11 Dépôt par sputtering.

III.4.4. Dépôt par laser pulsé (PLD)

Cette technique plus récente est mise en œuvre pour déposer des isolants en multicouches. Un laser pulsé de puissance (un Joule par impulsion environ) forme un plasma au niveau de la cible.



Figure III.12 Dépôt par ablation laser[20].

Ce plasma contient des atomes, des ions et des molécules de la cible. Ces composés se déposent sur le wafer. Le dispositif est représenté figure III.12.

III.4.5. Le dépôt en phase vapeur (CVD)

Cette méthode est très utilisée pour fabriquer des composants et des circuits Puisqu'elle permet de dépose des semi-conducteurs, des oxydes et des métaux. Le principe est de faire croitre sur un substrat une couche relativement mince à partir de composants en phase vapeur appelés précurseurs. Le substrat est chauffe dans un dispositif comme celui de la figure III.13.



Figure III.13 Réacteur CVD.

Différentes techniques sont possibles :

- Pression atmosphérique : APCVD,
- Basse pression : LPCVD,
- Haute pression : HPCVD.

Le dépôt à basse pression se fait à plus haute température. Le dépôt à haute pression peut se faire a plus basse température. Il est possible de déposer des semi-conducteurs et des isolants en exploitant les réactions chimiques suivantes :

 Dépôt de silicium poly cristallin : 	
$SiH_4 \rightarrow Si + 2H_2$	580-650 °C pour une pression de 1 mbar
 Dépôt de nitrure de silicium : 	
$3 \operatorname{SiH}_4 + 4\operatorname{NH}_3 \rightarrow \operatorname{Si}_3\operatorname{N}_4 + 12 \operatorname{H}_2$	700-900 °C a la pression atmosphérique

- Dépôt de silice : $SiH_4 + O_2 \rightarrow SiO_2 + 2H_2$ 450 °C

Ce sont les réactions les plus classiques. Elles utilisent un gaz appelé silane (SiH₄).

Le dioxyde de silicium produit par CVD ne peut cependant pas remplacer celui forme par oxydation thermique dans un four soumis a un flux d'oxygène ou de vapeur d'eau. L'oxyde ainsi forme à d'excellentes propriétés électriques. La CVD est également utilisée pour déposer

des métaux comme le tungstène, l'aluminium et le titane. Le cuivre pourrait aussi être dépose par cette méthode mais une méthode spécifique appelée « damascène» est généralement mise en œuvre.

III.4.6. Dépôt de type MOCVD

Le dépôt MOCVD (Metal Organic Chimical Vapour Deposition) est une évolution de la méthode CVD adaptée au dépôt de composes métalliques. Les précurseurs contiennent le métal à déposer et des composes organiques. Cette technique permet en particulier de déposer les matériaux a haute permittivité nécessaires à la fabrication des DRAMs et autres composants de la microélectronique. Prenons comme exemple le dépôt de PZT (compose à base de plomb, de titane et de zirconium). Les précurseurs sont les liquides suivants : $Pb(C_2H_5)_4$; $Ti(OC_3H_7)_4$; $Zr(OC_4H_9)_4$.

Comme exemple de réaction, prenons la formation de l'arséniure de gallium a partit du trimemthylgallium.

AsH $_3$ + Ga(CH₃) $_3$ \rightarrow GaAs + 3CH₄

III.4.7. Dépôt de type CSD (Chemical Solution Deposition)

Ce sont des méthodes chimiques qui consistent a partir de précurseurs généralement en phase liquide pour arriver a un film poly cristallin ou cristallin en passant par une phase amorphe. Il faut également citer les méthodes de type Langmuir-Blodgett pour déposer des composes organiques. Des molécules organiques présentant une extrémité hydrophobe et une autre extrémité hydrophile peuvent être déposées sur un volume d'eau de la même manière qu'un film d'huile se forme au dessus d'un volume liquide. Le film mince ainsi formé peut alors être transfère sur un substrat. Cette technique simple permet en particulier la fabrication des diodes électroluminescentes organiques et laisse espérer le développement d'une électronique grande surface.

III.4.8. Croissance thermique

Ce sont les méthodes qui permettent en particulier de fabriquer le wafer lui-même. Mais la méthode CVD est limitée a la fabrication d'une couche mince. Dans ce paragraphe, on étudie comment réaliser une tranche de 1 mm d'épaisseur. Les wafers sont découpées dans un lingot de silicium. Un lingot est un cylindre de diamètre élève (300 mm dans les fabrications les plus avancées) qui présente la propriété importante d'être monocristallin. D'autres semi-conducteurs peuvent être fabriques comme le germanium, l'arséniure de gallium ou le tellurure de cadmium. Le silicium a cependant pris une place largement majoritaire. L'obtention de cristaux monocristallins d'arséniure de gallium ou de tellurure de cadmium reste une opération difficile. Le procédé de croissance du lingot est le procédé Czochralski. Il est représenté de manière simplifiée figure III.14



Figure III.14 Croissance thermique du Silicium[21].

La méthode utilise le matériau EGS silicium poly cristallin fondu de haute pureté. Ce matériau est obtenu à partir de sable de haute pureté après un certain nombre de réactions chimiques. A partir d'un germe de silicium monocristallin, le silicium fondu apporté dans le creuset se solidifie autour du germe au fur et à mesure que le solide ainsi formé se déplace selon un mouvement combinant rotation autour de l'axe du cylindre et translation vers le haut. Le silicium dans le creuset est formé à partir de la décomposition de la silice, matériau très abondant dans la nature et d'un processus de purification permettant de contrôler les impuretés. Les déplacements du lingot (rotation et translation) sont très lents si bien que la

fabrication de lingots de haute pureté avec un contrôle précis de la structure cristallographique reste une opération complexe et lente justifiant le prix relativement élevé des wafers. Le lingot de silicium est ensuite découpé en tranches de 1 mm d'épaisseur environ pour assurer un minimum de rigidité mécanique. Du Bore et du Phosphore peuvent être apportés pendant la croissance pour obtenir un lingot de résistivité donnée. On peut également fabriquer par cette méthode des lingots d'Arséniure de Gallium ou de Tellurure de Cadmium. Pour ces deux matériaux la méthode « Bridgman » utilisant deux zones de cristallisation est également mise en œuvre.

Après la croissance, des tranches d'épaisseur inférieure au mm sont coupées dans le lingot ce qui permet d'obtenir les « wafers ». Il est ensuite nécessaire de polir les faces pour garantir un minimum de planéité puis d'éliminer les zones externes comportant des défauts suite au polissage. Une attaque chimique est donc une étape indispensable. A la fin du processus, il est nécessaire de polir les faces pour être compatible avec les très faibles dimensions des composants.

Les tranches obtenues ne sont pas des cristaux parfaits. Des impuretés peuvent occuper des sites substitutionnels (ils remplacent un atome du cristal) ou interstitiels (ils sont insérés dans le cristal). Enfin la régularité du réseau cristallin peut être perturbée localement, on parle alors de dislocations. Ces défauts auront une grande influence sur les propriétés électriques des composants fabriqués ultérieurement et un grand défi de la technologie est de réduire au maximum la densité de défauts par wafer.

III.4.9. La croissance électrolytique

Cette technique permet d'obtenir des couches relativement épaisses de métal sur un substrat. Prenons l'exemple du dépôt de nickel représente figure III.15.



Figure III.15 Dépôt par électrolyse.

La solution est dans ce cas du chlorure de nickel mélange a du chlorure de potassium. A la cathode, il y a une réaction de réduction :

 $Ni^{2+}+2e^{-}\rightarrow Ni.$

Au niveau de l'anode, il y a une réaction d'oxydation :

 $2Cl^{-}\rightarrow Cl^{2}+2e^{-}$.

Il y a donc dépôt de nickel sur le wafer jouant le rôle de cathode. La réaction peut se faire de manière sélective au travers d'ouvertures gravées dans la couche de résine déposée sur le substrat. Si le substrat n'est pas un bon conducteur, il faut déposer par une autre méthode une mince couche de métal qui sert de base de déclenchement de la réaction électrochimique [22].

III.4.10. Implantation ionique

Nous abordons maintenant les techniques qui permettent de modifier des matériaux en profondeur. La notion de profondeur est toute relative puisque les régions sont créées en général à moins d'un micron de la surface du wafer. La zone active du dispositif est donc située dans une région d'épaisseur négligeable devant l'épaisseur du wafer.
La première technique étudiée est l'implantation ionique. Ce n'est pas réellement une technique permettant de fabriquer en profondeur un matériau mais une technique qui permet de changer les propriétés électriques du matériau de base. Le principe est d'envoyer perpendiculairement à la surface du wafer un flux d'ions de haute énergie. L'énergie de ces ions est suffisamment élevée pour qu'ils pénètrent à l'intérieur de la matière avant d'être arrêtés sous l'effet des interactions avec les électrons des atomes de silicium du wafer. Le principe d'un implanteur ionique est illustré figure III.16.



Figure III.16 Implantation ionique.

Des atomes sont ionisés dans la source puis extraits vers l'aimant de déflexion. Cet aimant est utilisé en spectromètre de masse. La valeur du champ magnétique créé permet de dévier les ions en fonction de leur masse. Les ions choisis seront les seuls à pouvoir être extraits de cette zone de champ magnétique pour être ensuite accélérés par la colonne d'accélération. L'énergie des ions est donc réglable ce qui permet d'ajuster la profondeur de pénétration dans la matière. Des dispositifs de focalisation et de balayage complètent le dispositif. En résumé, L'implanteur permet les réglages suivants :

- choix des dopants par le réglage du champ magnétique,

- choix de la pénétration des ions par le réglage de la tension d'accélération,

- choix de la dose implantée par le réglage de l'intensité du faisceau et du temps d'implantation.

Les énergies d'implantation sont comprises entre quelques keV et quelque MeV. Si on examine le wafer dans sa profondeur après implantation, on obtient une répartition des ions implantés du type de la courbe III.17. La profondeur d'implantation ne peut être définie qu'en moyenne car tous les ions ne subissent pas le même nombre de collisions avec les électrons du wafer. La profondeur moyenne dépend de l'énergie des ions et est inférieure au micron. Pour les faibles énergies, les ions sont implantés au voisinage de la surface du wafer [20].

Il ne suffit pas d'implanter des ions pour créer en profondeur des zones de dopage donné, il faut que les dopants (Bore, Arsenic, Phosphore) occupent des positions substitutionnelles dans le réseau cristallin. C'est à cette condition que les dopants sont électriquement actifs. De plus, l'implantation créée de nombreux défauts dans le réseau cristallin en déplaçant les atomes. Il est donc nécessaire pour ces deux raisons de chauffer le substrat implanté pour guérir les défauts et pour rendre les dopants électriquement actifs et modifier les propriétés électriques. Cette opération s'appelle le recuit et se pratique vers 600 °C.



Figure III.17 Répartition des dopants après implantation.

III.4.11. La diffusion thermique

La diffusion thermique est une opération complémentaire de l'implantation ionique qui permet de distribuer des dopants dans un substrat. Si le matériau est chauffé à haute température (au dessus de la température du recuit), les dopants peuvent avoir assez d'énergie pour se déplacer dans la matière en quittant leurs sites initiaux. Ces déplacements dépendent de la température et du temps pendant lequel la haute température est appliquée. La diffusion

élargit la distribution initialement créée par l'implantation ionique. Il est également possible de faire diffuser dans le wafer des atomes déposés en surface.

III.5. Les procédés alternatifs

La lithographie a fait la force de la micro-électronique en permettant la fabrication collective des circuits. Elle est cependant la principale source de difficultés pour l'avenir. En effet, les dispositifs de lithographie sont d'une complexité croissante au fur et à mesure que la résolution demandée s'affine. Le coût des équipements devient très élevé. Un dispositif d'insolation avec photo répétition est acheté autour de 10 millions d'euros et un jeu de masques pour réaliser un circuit complexe peut coûter quelques millions d'euros. Les masques sont fabriqués à l'aide d'un appareil générant un faisceau d'électrons capable de balayer la surface du masque avec une précision extrême. Le procédé n'est plus limité par la longueur d'onde puisque la longueur d'onde associée à l'électron est très faible. La résolution possible de gravure est de quelques nm mais le temps d'insolation est très long à cause du caractère séquentiel de l'opération ce qui explique en partie le coût du jeu de masques fabriqué par cette méthode.

On pourrait imaginer des procédés de fabrication permettant de se passer des masques. On pourrait par exemple utiliser la lithographie par faisceau d'électrons non pas pour fabriquer les masques mais pour produire directement les circuits intégrés. Cela est effectivement pratiqué pour réaliser des prototypes. La lithographie par faisceau d'électrons ne peut cependant être mise en œuvre dans un procédé industriel car le temps d'insolation est long par principe puisqu'il est séquentiel. Il est donc naturel que la micro-électronique investigue des techniques alternatives à la lithographie optique [20].

III.5.1. La nano-impression

Ce procédé a été proposé par S.Y. Chou en 1995. Il est devenu une méthode de référence pour fabriquer des nano dispositifs car très simple de mise en œuvre. Son principe est hérité des technologies de l'impression. Il est illustré figure III.18.



Figure III.18 Nano-impression.

Le procédé s'explique simplement, Un moule réalisé par lithographie électronique sur un support de silicium écrase un polymère déposé sur le substrat. La pression est de quelques dizaines de bars. La gravure ionique réactive permet ensuite de graver le polymère en atteignant le substrat. Un film mince est ensuite déposé sur le polymère résiduel, un film métallique par exemple. L'étape suivante appelée « lift-off » consiste à enlever le polymère par dissolution chimique.

On obtient alors le motif choisi imprimé dans le film déposé. Ce film peut par la suite servir de masque de gravure. Cette technique permet de réaliser des motifs avec une résolution de l'ordre de 10 nm. Le moule peut en effet être fabriqué par une technique mettant en œuvre la lithographie e-beam puisqu'il est fabriqué une seule fois. La nano impression est appliquée pour étudier des dispositifs de très faible dimension et pour réaliser des matériaux magnétiques ou optiques nanostructures.

Des techniques dérivées sont également étudiées comme la nano impression sous irradiation, la nano compression et la lithographie molle. La nano impression sous irradiation permet de travailler à température ambiante. La nano compression utilise le polymère sous forme de granulés. La lithographie molle est basée sur l'utilisation d'un élastomère pour réaliser une sorte de tampon encreur capable de déposer des molécules organiques. On peut également citer la lithographie en champ proche, technique de lithographie par contact mais qui utilise un masque souple en contact parfait avec le substrat. Toutes ces techniques sont encore au stade de la recherche et ne sont pas encore appliquées dans l'industrie.

III.5.2. Techniques d'auto-assemblage

Pour éviter les coûts de la lithographie, il est séduisant d'imaginer des structures organisées qui seraient réalisées par des procédés purement chimiques. Il est difficile d'imaginer que de telles méthodes soient capables de générer des systèmes non réguliers comme les unités de calcul mais il est envisageable de les appliquer à la fabrication de systèmes réguliers tels que les mémoires.

La fabrication atome par atome est possible en appliquant les principes de la microscopie en champ proche mais elle n'est pas applicable pour la réalisation de systèmes complexes. Il est donc nécessaire d'imaginer des méthodes collectives conduisant à la création de structures régulières de nano-objets. Ces méthodes sont très nombreuses mais peuvent se répartir en deux grandes familles. La première s'appuie sur les propriétés cristallographiques du support et la seconde intègre différentes techniques issues de la synthèse chimique.

Il est intéressant d'envisager les méthodes de croissance de nano-objets sur des surfaces préstructurées. Prenons l'exemple de la surface obtenue par section d'un cristal. La surface de coupe peut, si la coupe ne se fait pas dans la direction d'un plan cristallin, présenter une succession de marches à l'échelle nanométrique. Ce support structuré peut alors être la base pour faire croître des nano fils le long des marches.

Trois grandes techniques sont mises en œuvre pour former les surfaces pré-structurées :

- Profiter des propriétés intrinsèques des surfaces (réseaux de défauts, réseau de marches, ...),

 Structurer la surface par des techniques de gravure particulières (géométrie de la surface ou création de réseaux de dislocations),

- Coupler les deux méthodes précédentes.

Les méthodes de synthèse chimique s'inspirent des assemblages moléculaires qui ont conduit aux organismes vivants. Les assemblages peuvent alors se faire en volume ou en surface. Ces méthodes permettent en particulier de créer des réseaux de nano particules magnétiques mais aussi des monocouches organisées sur des surfaces.

CONCLUSION GENERALE

L'objectif de ce travail est l'étude des matériaux semi-conducteurs III-V pour d'avantage remédie aux effets qu'entraine la miniaturisation d'un transistor MOS, Grace à leurs hautes mobilités et leurs structure de bande d'énergie pour des applications Didier à la nano électronique et la nano photonique.

Le premier chapitre a expliqué ce qu'est un transistor MOS à effet de champ, puis a présenté son principe de fonctionnement ainsi que les équations de base régissant en termes de courants les différents modes de fonctionnement dans lesquels sont utilisés les transistors MOS. Ensuite nous nous sommes penchés sur les effets qu'entraîne la miniaturisation d'un transistor MOS sur les principaux paramètres électriques régissant son fonctionnement.

Le second chapitre à présenté les matériaux semi-conducteur III.V, les différents composés binaires, ternaires et quaternaires leurs propriétés cristallines ainsi que leurs structures de bande d'énergie qui fait d'eux des matériaux plus compatible pour des transistors plus submicronique voir nanométriques.

Enfin dans le dernier chapitre en à repris la fabrication collective des circuits intégrés, la lithographie, les différents apports de matériaux (dépôt, épitaxie,...Etc.) puis les procédés alternatifs comme la nano-impression et les techniques d'auto-assemblage

BIBLIOGRAPHIE :

[1] BENLATRECHE Mohamed Salah « Caractérisation des dispositifs électroniques dans les technologies MOS » Universités de Boumerdes, 2012.

[2] F.Hofman,W.H Krautschneider « A simple technique for determing the interface trap distribution of submicron Metal-oxide-semiconducteur transistor by the charge pumping method »

[3] DJEFFAL Fayçal « Modélisation et simulation prédictive du transistors MOSFET fortement submicronique, application à la conception des dispositifs intégrés » ;04/2006.

[4] KRUNOSLAV Romanjek « Caractérisation et modélisation des transistors CMOS des technologies 50 nm et en deçà » Institue National Polytechnique de Grenoble ;Novembre 2004.

[5] G.Baccarani ,M.wordman et R .Dennard « Generalized Scaling theory and its application to a ¹/₄ MOSFET design » IEEE PDF,vol 31.Page(s) 452-462 ;1984

[6] TAMOUM Mohammed « Caractérisation fine et modélisation non linéaire des transistors MOSFET » Février 2013.

[7] M.J Deenand .Z.X.Yan « DIBL in short-channel Nmos devices at 77k »

IEEE TED PDF, vol 39, page(s) 908, 1992.

[8] BENHAMIDA yahia « Etude des caractérisation physiques et électrique d'un MOSFET nanométrique »

[9] Xie,Q ;XU ,J ; Taur,y « Review and critique of analytic models of MOSFET short-channel effects in subthreshold » Electron Devices ,IEEE PDF ,transaction on ,on page(s) 1569-1579 , Volume 59 ;juin 2012.

64

[10] T.Grotjohn et B. Hoefflinger « A parametric short-channel MOS transistor model for subthreshold and strong inversion currebt ,IEEE Transactions on Electron Devices ,vol31,page(s) 234-246 , 1984.

[11] GUEDDA Hayat « Etude des effets indésirables dans les transistors MOSFET à canaux courts » ;2012.

[12] BENATMANE Kaled «Etude de laser à base de semi-conducteurs III-V » Decembre 2011.

[13] BENAICHA Mouna «Etude des matériaux Semi-conducteurs III-V ;Application à la conversion photovoltaïque »2007.

[14] MOUATSI Abdelmalek «Composants à hétéro structures : application en nanoélectronique et nano photonique »2013.

[15] ZHE-CHUAN-FENG « Semiconductor Materials».

[16] <u>WWW.Techniques-ingenieur.fr</u> « Technologies matériaux composés (Transistors et circuits intégrés à hétéro structures (III-V)) »; REF INTERNET E2450.

[17] JUN Cheng « Intégration monolithique de matériaux III-V » ;Ecole centrale de Lyon ,2011.

[18] CORRADO Sciancalepore «Intégration hétérogène III-V sur silicium de micro laser à émission par la surface à base de cristaux photoniques » ;Ecole centrale de Lyon ,2012.

[19] AZZA Chettaoui « Croissance d'hétérostructures III-V sur des couches tampons de SrTio₃ /Silicium» ;ecole centrale de Lyon,2013.

[20] HENRY Mathieu 6^{eme}Edition « Physiques des semi-conducteurs et des composants électroniques».

[21] CHELOUCHE Salim « Propriétés des fenêtres optiques ZnO :Al pour cellules solaires en couches minces à base de CIGS ».dec 2012

[22] ZERIG Alkhadhir, BENAMOR Rakbia « Etude de procédé bain chimique et les réactions amenant au matériau cds en couche minces ».juin 2015.