

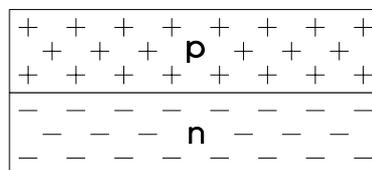
LE TRANSISTOR

Ce chapitre décrit la constitution, le fonctionnement et la fabrication du transistor. Nous étudierons tout d'abord le fonctionnement de la diode, qui permet d'expliquer simplement les rudiments du comportement des semi-conducteurs en présence de tensions, puis nous passerons au transistor à jonction, aux principes très proches de la diode, et qui nous permettra d'aborder sereinement le transistor MOS, lequel constitue la brique élémentaire de la plupart des circuits numériques.

LA DIODE

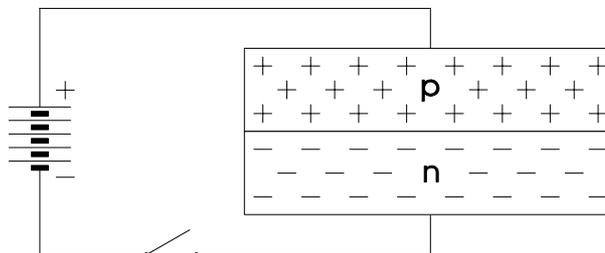
Nous n'évoquerons plus la diode par la suite, mais il s'agit d'une étape importante pour comprendre le transistor.

Réalisons une diode. Au sein d'un même cristal, on dope deux zones jointives : l'une dopée p, l'autre dopée n (la surface de liaison entre ces deux zones est appelée *jonction*).



Sens passant

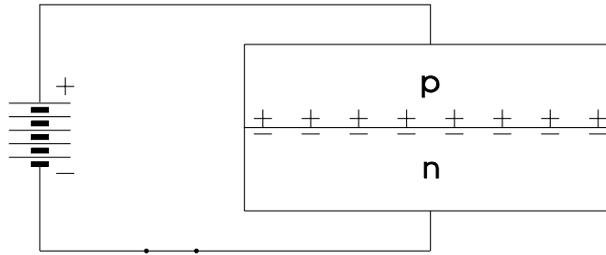
Relions maintenant ce semi-conducteur à un générateur de tension (via un interrupteur), en connectant la zone p au pôle positif du générateur, et la zone n au pôle négatif.



Les charges négatives et positives de chaque zone n et p sont réparties uniformément dans celles-ci. Fermons l'interrupteur.

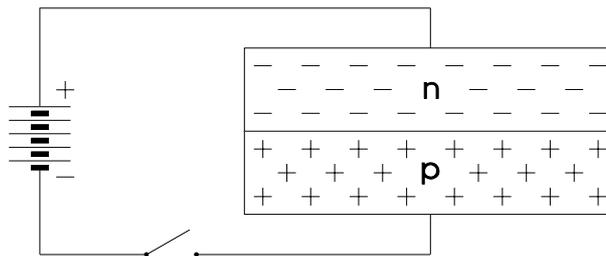
Les trous de la zone p sont repoussés par la borne + du générateur, les électrons de la zone n étant repoussés par sa borne -.

- Les zones semi-conductrices sont désertées par leurs charges libres.
- Par contre, la jonction est le siège de phénomènes spécifiques. Les trous p et les électrons n se rejoignent et s'annulent donc au niveau de la jonction, ce qui "tue" les trous p. Un électron provenant de la borne - du générateur va donc normalement traverser la zone n semi-conductrice, rejoindre la jonction où aucun trou ne viendra "tuer" cette charge -, puis traverser la zone p semi-conductrice. Le courant s'établit donc.



Sens bloqué

Invertissons maintenant les bornes du générateur de tension.

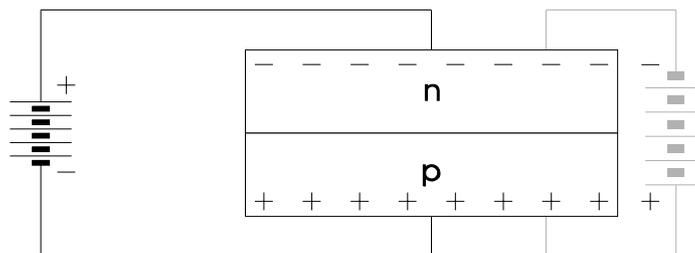


Les charges négatives et positives de chaque zone n et p sont réparties uniformément dans celles-ci. Fermons l'interrupteur.

Les trous de la zone p sont attirés par la borne - du générateur, les électrons de la zone n étant attirés par sa borne +. Un électron provenant de la borne - du générateur va donc entrer dans les trous de la zone p, et ne pourra pas en ressortir, à la fois piégé par l'attraction des trous de la zone p et repoussé par les électrons de la zone n. Le courant ne s'établit donc pas.

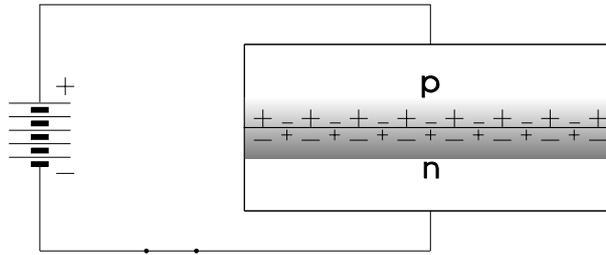
En fait, par le positionnement des électrons et trous de ses semi-conducteurs, la diode se comporte comme un générateur de tension, disposé en sens inverse de celui du circuit, et s'opposant donc au courant dans la diode créé normalement par lui (en grisé).

La diode est bloquée.

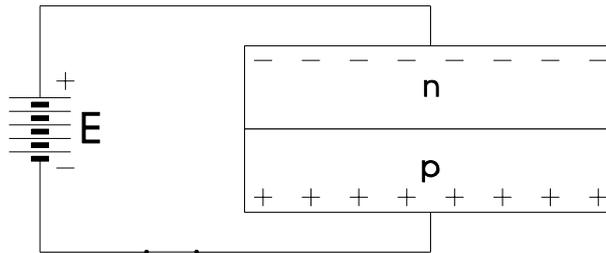


LE TRANSISTOR A JONCTION

Si l'explication précédente simpliste suffit à expliquer le fonctionnement de la diode dans le sens passant, il est nécessaire pour la suite de préciser qu'il existe une légère diffusion des charges d'une zone à l'autre au niveau de la jonction : les électrons arrivent à pénétrer légèrement dans la zone p, tandis que les trous arrivent à pénétrer légèrement dans la zone n. Ces charges, quand même en "pays ennemi", resteront néanmoins minoritaires. Cette zone d'interpénétration n'a donc pas une épaisseur nulle, mais réelle, bien que faible. C'est là la clé de l'effet transistor.

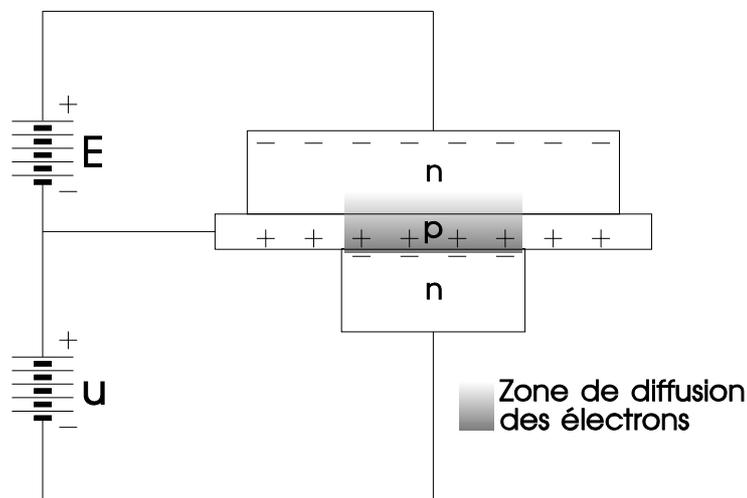


Reprenons le cas de la diode bloquée, alimentée par une tension E .



Nous injectons des électrons au niveau de la liaison entre le fil et la zone p. Ces électrons sont retenus par les trous. Mais, supposons que, comme dans une diode passante, il y ait une certaine profondeur de diffusion de ces électrons. Supposons maintenant que la zone p soit plus mince que la profondeur de diffusion. Des électrons vont réussir à rejoindre la zone n, où ils seront attirés alors par la borne + du générateur de tension, créant alors un courant. Ceci commanderait donc l'état passant de la diode dans un sens où elle est normalement bloquée.

Question : comment faire diffuser ces électrons à travers la zone p ? Le schéma représentant l'état passant de la diode nous apporte la réponse : on établit une zone de diffusion en créant une jonction dans le sens passant. Il suffit d'ajouter une zone n entre le fil et la zone p, l'ensemble des zones n, p puis n constituant un transistor :

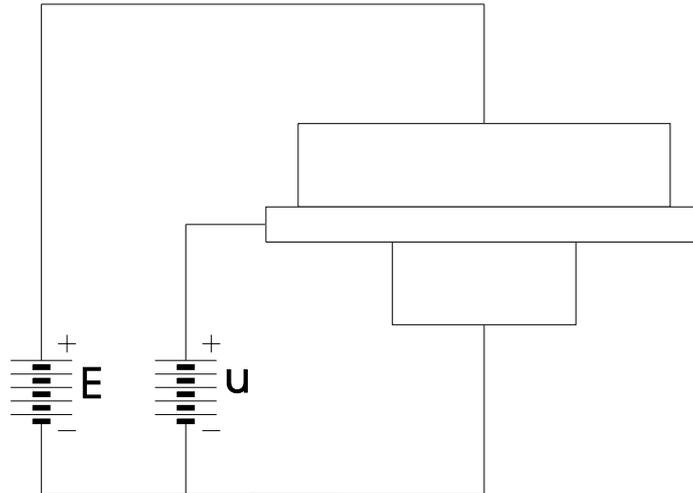


La zone p est considérablement amincie, de manière à être plus fine que la profondeur de diffusion. Par ailleurs, la zone n supplémentaire est petite, de manière à s'assurer qu'un maximum de ses électrons seront capturés par l'autre zone n.

Si u est nul, la présence de la petite zone n ne change rien : nous avons un transistor bloqué. Par contre, en présence de u , un courant s'instaure de la petite zone n à la grande zone n, à travers la zone p : nous avons un transistor passant.

La terminologie définit la petite zone n comme l'*émetteur*, la zone p la *base*, et la grande zone n le *collecteur*. Le transistor que nous venons de décrire est un transistor dit *npn*, par opposition aux transistors *pn*, dont l'émetteur et le collecteur sont dopés p, et la base dopée n (ce qui inverse son fonctionnement).

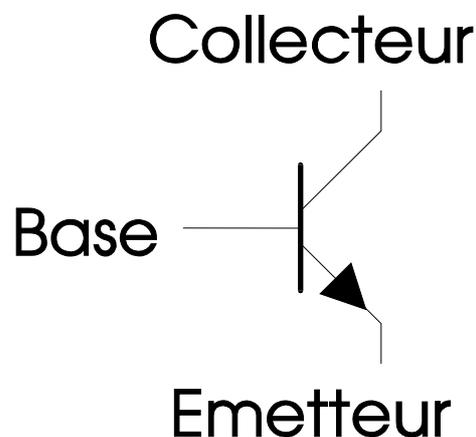
Le montage peut être également réalisé de la manière suivante (dite *émetteur commun*) :



C'est celui que nous utiliserons dans les circuits intégrés, où le transistor est utilisé comme simple interrupteur entre l'émetteur et le collecteur, avec la base comme commande. En appelant V_{cc} (V pour "tension", cc pour Courant Continu, V_{cc} est généralement de l'ordre de quelques volts) la tension d'alimentation du circuit (qui n'est autre que E), nous réalisons l'approximation suivante :

- Si $u = V_{cc}$, l'interrupteur entre le collecteur et l'émetteur est considéré fermé : le courant passe, le transistor est dit *passant*.
- Si $u = 0\text{ V}$, l'interrupteur entre le collecteur et l'émetteur est considéré ouvert : le courant ne passe pas, le transistor est dit *bloqué*.

La symbolique classique utilisée dans les schémas électroniques de type téléviseur ou chaîne hi-fi, où le transistor à jonction est un composant distinct, est la suivante (pour un transistor npn) :



LE TRANSISTOR MOS

Les impératifs informatiques actuels imposent d'implanter plusieurs millions de transistors sur quelques millimètres carrés de silicium. La "loi de Moore", qui n'est pas véritablement une loi mais plutôt une prédiction visionnaire datant de 1965 et toujours vérifiée depuis, dit que le nombre de transistors dans un microprocesseur double tous les 18 mois ! Cela implique d'avoir aujourd'hui des transistors de taille inférieure au micron (1 micron, noté $1\ \mu\text{m}$, égale à $1/1000^{\text{e}}$ de millimètre), interdisant donc l'utilisation de composants distincts, présentant un coût unitaire très faible ainsi qu'une faible consommation. Ceci a conduit à utiliser un type de transistor particulièrement adapté à ces spécifications : le transistor MOS (Metal Oxyde Semiconductor).

Constitution

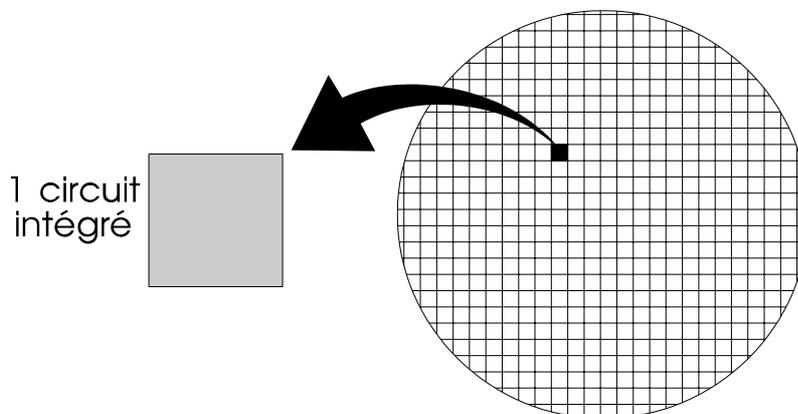
Le transistor MOS présente la spécificité d'avoir une couche isolante très mince entre le fil normalement relié à la base d'une part, et la base d'autre part. Les terminologies diffèrent :

- la base est appelée *canal*
- les émetteur et collecteur deviennent le *drain* et la *source*
- le fil connecté à la base devient la grille. Ce mot est particulièrement mal adapté : la terminologie historique de ce mot est *gate*, signifiant *porte* en américain. Or, remplacer le *porte* français par le *gate* américain ne permettait pas de conserver la même initiale (problème qui ne se posait pas avec le drain, la source et le canal). Il a donc été choisi le mot *grille*.

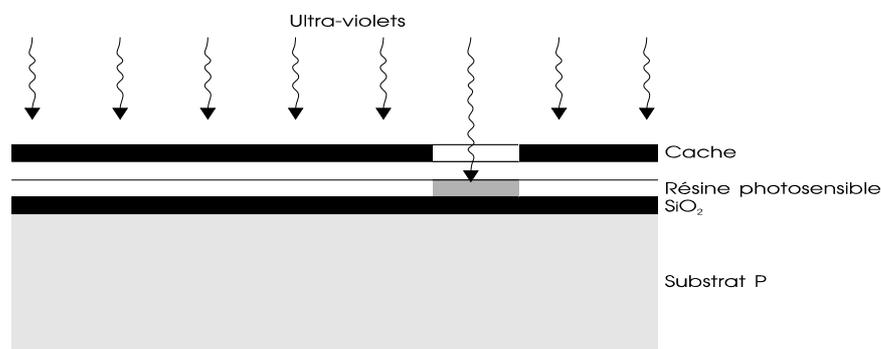
Le meilleur moyen de savoir comment un tel transistor est constitué est de suivre sa fabrication.

Fabrication

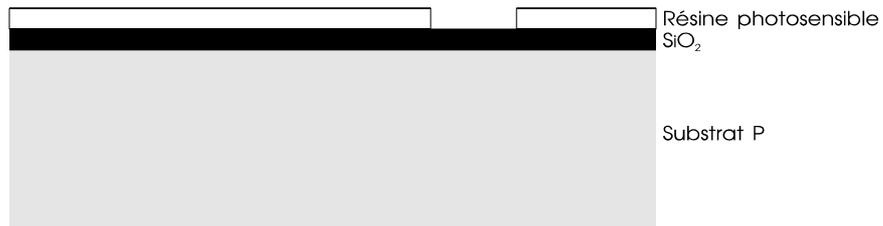
- Du silicium purifié est fondu en un cylindre d'environ 10cm de diamètre, puis découpé en rondelles d'1/2mm d'épaisseur. C'est sur chacune de ces rondelles, constituant le *substrat*, que nous allons graver, à plusieurs centaines d'exemplaires, des circuits intégrés de quelques mm². Nous nous attacherons à la réalisation d'un transistor à canal p, puis par similitude d'un transistor à canal n, sur un substrat dopé p.



- Le substrat est soumis à une atmosphère oxydante (atmosphère humide), ce qui génère une couche d'oxyde de silicium (SiO_2) d'une épaisseur de l'ordre du μm à sa surface.
- Puis, par une méthode que nous utiliserons par la suite systématiquement sans la détailler (le *masquage*), un évidement, laissant le substrat à nu, est réalisé dans cette couche d'oxyde. La méthode est la suivante :
 - Dépôt d'une couche de résine photosensible, ayant la propriété d'être dissoute à l'acide si elle est exposée à des rayons ultra-violettes (notés UV).
 - Un masque, présentant des trous devant les endroits où on veut réaliser l'évidement dans le SiO_2 , est amené au-dessus de la résine.
 - Des UV sont projetés sur le cache et polymérisent la résine aux endroits illuminés, de la même manière qu'on peint au pochoir.



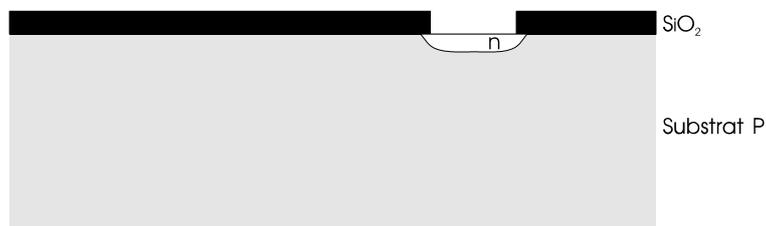
- Le masque est retiré, et la pastille plongée dans un bain d'acide, qui retire la résine là où elle est polymérisée.



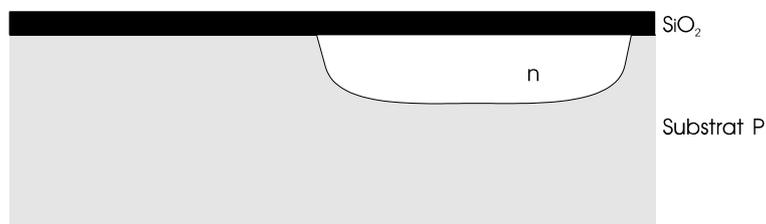
- La surface de la pastille reçoit un bain d'un deuxième acide, qui retire le SiO₂ là où la résine a disparu.
- Un solvant retire la couche de résine restante. Il nous reste donc une couche de SiO₂, laissant le substrat à nu à l'endroit désiré.



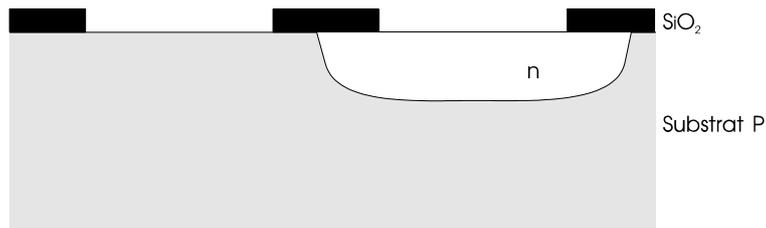
- Puis les impuretés dopantes sont déposées par diffusion ionique, formant une mince couche n (en général, du phosphore).



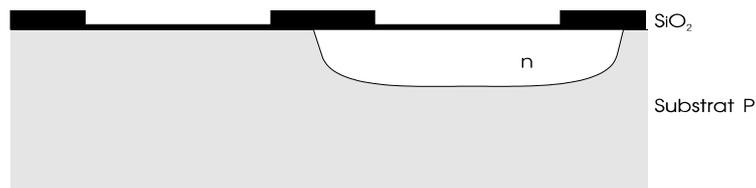
- L'ensemble est porté à haute température (1400 °C), de manière à augmenter par diffusion la taille de la couche n, qu'on appelle alors *caisson n*, le tout sous atmosphère humide, de manière à régénérer la couche de SiO₂. Ce caisson aura pour fonction de réaliser un "pseudo-substrat" n pour le transistor canal p.



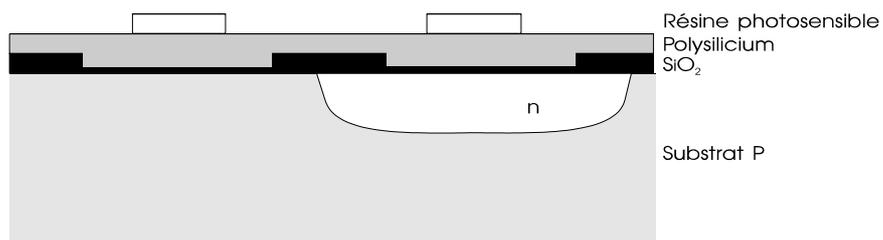
- Par la méthode du masquage, nous réalisons alors deux évidements dans le SiO₂.



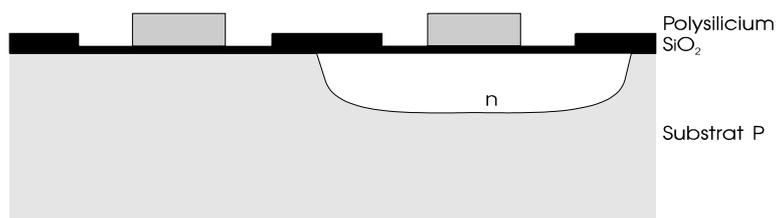
- Puis on réoxyde légèrement le substrat, de manière à régénérer une mince couche de SiO₂ (environ 50 Å), zone appelée *oxyde mince*.



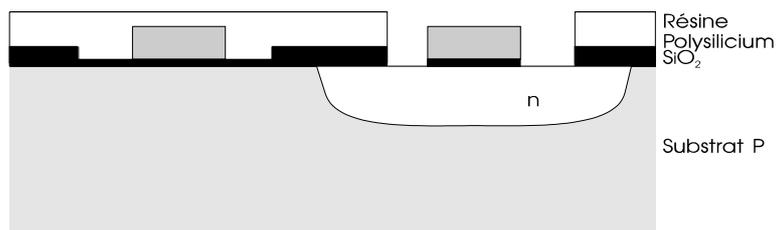
- On réalise alors un masquage quelque peu différent des précédents : on dépose une couche de polysilicium, puis de résine. Après mise en place du masque, insolation aux UV et polymérisation de la résine, on attaque celle-ci à l'acide. Mais celui-ci attaque la résine qui n'a pas polymérisé.



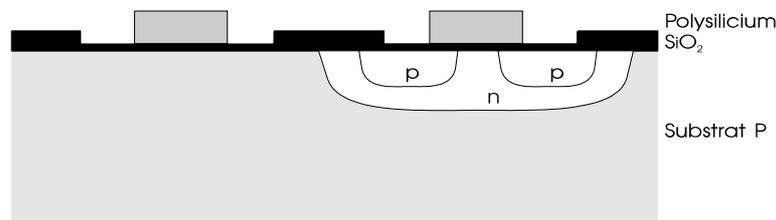
- Par un deuxième acide, le polysilicium est retiré là où il n'y a pas de résine. Enfin, la résine est retirée par un solvant.



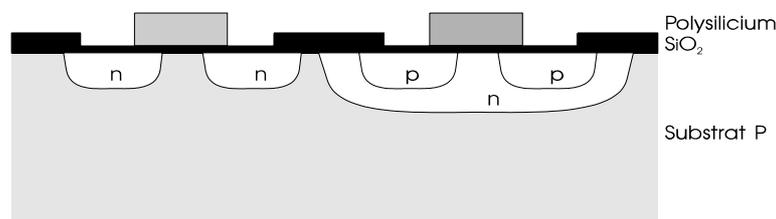
- Par un nouveau masquage, et grâce à un acide sélectif, on enlève l'oxyde mince apparent sur le transistor n (on n'enlève pas encore la résine) :



- On diffuse alors dans le caisson n des ions dopés p (en général, du bore), qui pénètrent sur une faible épaisseur, puis on retire la résine par du solvant. De la même manière que pour le caisson n, on chauffe de manière à augmenter la diffusion du bore. Enfin, la structure est refermée par une oxydation en surface du silicium.



- Avec des méthodes similaires, nous fabriquons un transistor canal n (sans caisson, car le substrat présente le bon dopage) :



Par les mêmes méthodes, il est possible de graver des résistances, des diodes et des capacités. Bien entendu, nous avons décrit les opérations pour un transistor, mais chaque opération (masquage, bain d'acide, insolation UV, etc...) traite l'ensemble des composants de la rondelle de silicium. De plus, en perçant des petits trous au niveau de l'oxyde mince, afin d'accéder aux sources et drains, et en gravant par les mêmes méthodes des pistes conductrices (généralement en aluminium), on peut connecter les grilles, sources et drains des différents transistors (ainsi que les résistances, diodes et capacités), et ainsi constituer les circuits les plus complexes. Les composants n'étant pas séparables de leur support, comme dans le cas de cartes électroniques où les composants sont distincts (on dit *discrets*), mais logés de manière continue dans celui-ci, on parle de *circuit intégré*.

Nous avons ainsi gravé quelques centaines d'exemplaires de notre circuit intégré sur la rondelle de silicium. Celle-ci est alors découpée en autant de petits carrés, qui seront logés dans un boîtier en plastique, appelé *puce* (français) ou *chip* (homologue anglais), et reliés aux connexions de celui-ci par de minces fils d'or (élément particulièrement conducteur et inoxyidable).

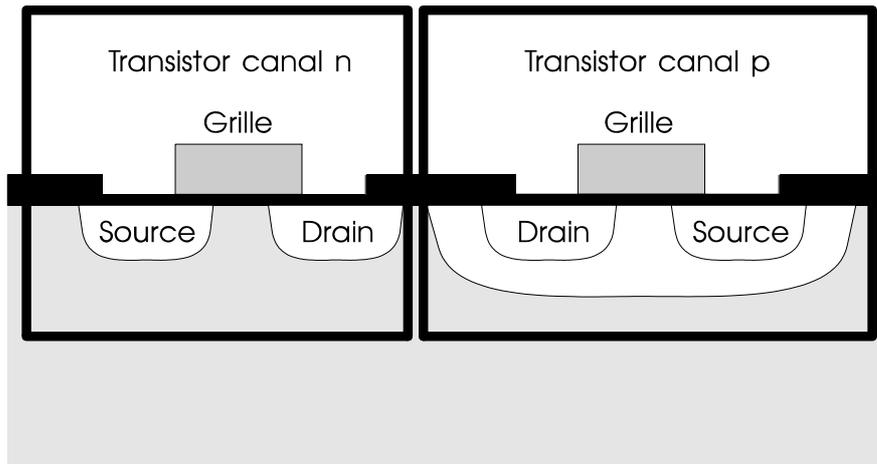
La fiabilité requise pour l'élaboration d'un transistor doit être faramineuse : si elle est d'un composant défectueux sur 1000, ce qui semble *a priori* être une bonne fiabilité pour un process de fabrication, cela est catastrophique pour la fiabilité du microprocesseur, qui comporte plusieurs millions de transistors, et qui comporterait donc plusieurs milliers de transistors défectueux. Il y a deux parades à cela :

- taux de fiabilité du procédé de fabrication d'un transistor très élevé
- circuits de correction d'erreur intégrés (par exemple un circuit triplé, le résultat majoritaire étant retenu).

La taille du canal est déterminante pour le temps de commutation du transistor (temps pendant lequel celui-ci passe de l'état bloqué à passant, ou inversement), et donc sur le nombre d'opérations auxquelles celui-ci pourra participer en une seconde. A mi-1999, la taille des canaux était de l'ordre de 0,25 μm (soit 2500 \AA), à comparer au pas de la structure cristalline du silicium (distance entre deux atomes consécutifs), de l'ordre de 3 \AA : cela signifie qu'il y a moins d'un millier d'atomes de silicium entre le drain et la source ! Le nombre de charges participant au courant dans le canal est donc très réduit. Or, cela constituera une limite technologique lorsque le canal aura une longueur de l'ordre de 100 à 500 \AA , soit entre 30 et 170 atomes de silicium entre la source et le drain, où les lois qui régissent le fonctionnement des transistors ne s'appliqueront plus. Au rythme actuel, nous devrions y arriver dans les prochaines années.

Fonctionnement

Le transistor est ainsi constitué :

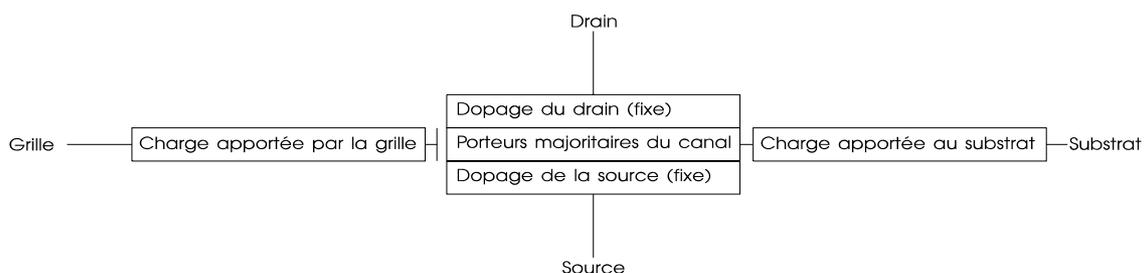


La symbolique standard utilisée est la suivante (la flèche indique le sens passant de la diode) :

Canal n		ou bien	
Canal p		ou bien	

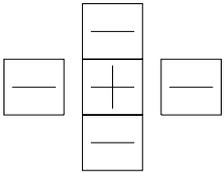
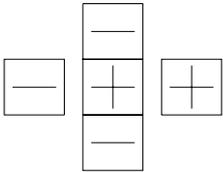
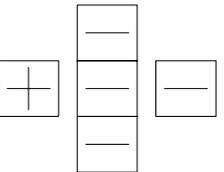
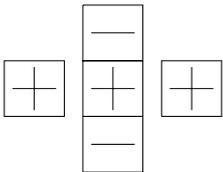
Le principe de fonctionnement de tels transistors est schématiquement le suivant : le but est de faire passer ou non du courant entre le drain et la source. Grâce aux tensions du substrat et de la grille (sachant que celle du substrat sera fixée par la construction du transistor, et que celle de la grille pourra être soit 0V, soit V_{cc}), nous allons attirer (mais non apporter comme dans le cas du transistor à jonction) des charges tantôt positives, tantôt négatives dans le canal, afin de retrouver un effet transistor.

Pour plus de compréhension, nous allons schématiser le transistor à l'aide de cinq cases, représentant la nature (positive ou négative) des porteurs majoritaires de chacune des trois zones le constituant, et les charges (positives ou négatives) influençant la nature des porteurs majoritaires du canal :



Nous allons voir à quelle tension doit être relié le substrat pour obtenir un interrupteur électrique, c'est-à-dire le passage ou non d'un courant entre le drain et la source, selon la tension appliquée à la grille.

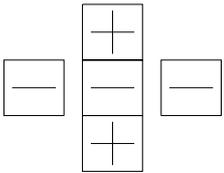
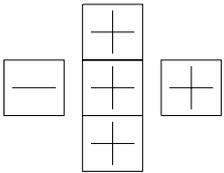
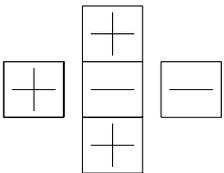
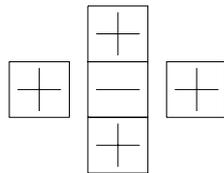
Cas du transistor canal n

	Substrat relié à 0V	Substrat relié à Vcc
Grille à 0V	 <p style="text-align: center;">BLOQUE</p> <p>La grille et le substrat amenant les mêmes charges, ils n'ont aucune influence sur le canal, qui garde ses propres porteurs, majoritairement positifs (dopé p). Ainsi, du drain vers le canal, nous avons une diode bloquée.</p>	 <p style="text-align: center;">BLOQUE</p> <p>La grille attire les charges positives fournies par le substrat, donc le canal est majoritairement positif. Ainsi, du drain vers le canal, nous avons une diode bloquée.</p>
Grille à Vcc	 <p style="text-align: center;">PASSANT</p> <p>La grille attire les charges négatives fournies par le substrat, donc le canal devient majoritairement négatif. Ainsi, il n'y a pas de diode bloquée, et le courant passe entre le drain et la source.</p>	 <p style="text-align: center;">BLOQUE</p> <p>La grille et le substrat amenant les mêmes charges, ils n'ont aucune influence sur le canal, qui garde ses propres porteurs, majoritairement positifs (dopé p). Ainsi, du drain vers le canal, nous avons une diode bloquée.</p>
Interrupteur réalisé ?	OUI (la tension à la grille permet de commander le passage ou non du courant entre le drain et la source)	NON (la tension à la grille n'a aucune influence sur le passage ou non du courant entre le drain et la source)

Nous veillerons donc à relier le substrat des transistors canal n à 0V, ainsi :

- si la grille est à 0V : le transistor bloque le courant entre le drain et la source ;
- si la grille est à Vcc : le transistor laisse passer le courant entre le drain et la source.

Cas du transistor canal p

	Substrat relié à 0V	Substrat relié à Vcc
Grille à 0V	 <p style="text-align: center;">BLOQUE</p> <p>La grille et le substrat amenant les mêmes charges, ils n'ont aucune influence sur le canal, qui garde ses propres porteurs, majoritairement négatifs (dopé n). Ainsi, du canal vers la source, nous avons une diode bloquée.</p>	 <p style="text-align: center;">PASSANT</p> <p>La grille attire les charges positives fournies par le substrat, donc le canal devient majoritairement positif. Ainsi, il n'y a pas de diode bloquée, et le courant passe entre le drain et la source.</p>
Grille à Vcc	 <p style="text-align: center;">BLOQUE</p> <p>La grille attire les charges négatives fournies par le substrat, donc le canal est majoritairement négatif. Ainsi, du canal vers la source, nous avons une diode bloquée.</p>	 <p style="text-align: center;">BLOQUE</p> <p>La grille et le substrat amenant les mêmes charges, ils n'ont aucune influence sur le canal, qui garde ses propres porteurs, majoritairement négatifs (dopé n). Ainsi, du canal vers la source, nous avons une diode bloquée.</p>
Interrupteur réalisé ?	NON (la tension à la grille n'a aucune influence sur le passage ou non du courant entre le drain et la source)	OUI (la tension à la grille permet de commander le passage ou non du courant entre le drain et la source)

Nous veillerons donc à relier le substrat des transistors canal p à Vcc, ainsi :

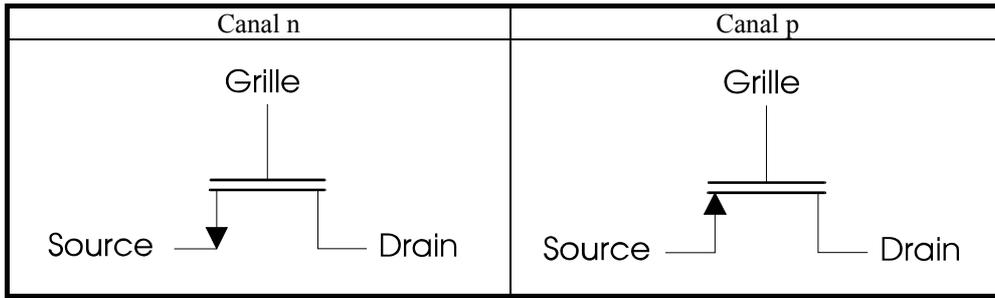
- si la grille est à 0V : le transistor laisse passer le courant entre le drain et la source ;
- si la grille est à Vcc : le transistor bloque le courant entre le drain et la source.

Utilisation

Question : quelle est la différence fondamentale entre un transistor à jonction et le transistor MOS, et finalement, pourquoi préférer ce dernier pour les circuits intégrés ?

Toute la différence provient du fait que dans le transistor à jonction, il y a une connexion électrique directe entre la base et le fil qui l'alimente. Ainsi, les charges présentes dans la base sont réellement celles qui proviennent du fil. Dans le MOS, une mince couche d'oxyde isole la grille de la zone dopée intermédiaire (entre le drain et la source). Il n'y a donc quasiment aucun courant qui circule depuis la grille. L'ensemble "grille-isolant-canal" forme un condensateur qui permet à la grille, par le champ électrique qu'elle produit, d'influencer la nature du canal (majoritairement positive ou négative) sans utiliser de courant (le peu de courant qui passe par la grille est celui nécessaire à charger cet infime condensateur) : les charges présentes dans le canal ne sont pas celles qui proviennent de la grille, mais celles qui ont été attirées (ou non) par elle. Pour cette raison, ce type de transistor est appelé *transistor à effet de champ*. Ceci lui confère une consommation particulièrement basse, et donc un échauffement réduit (souvenez-vous de vos vieux cours d'électricité : $U = RI$, $P = UI$, donc $P = RI^2$: la puissance thermique dégagée est proportionnelle au carré de l'intensité), ce qui est extrêmement important quand il s'agit d'en implanter plusieurs millions sur quelques millimètres carrés.

Dans le souci de simplifier la lecture des schémas dans les pages suivantes, nous utiliserons la symbolisation électrique suivante pour les transistors MOS (le substrat étant relié conformément à ce que nous venons de voir) :

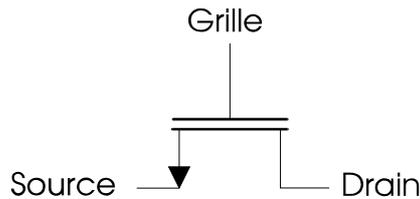


Remarque : les explications fournies ici ont pour objectif d'introduire la suite, qui ne concerne que le fonctionnement des ordinateurs. Ceci justifie certaines approximations, suffisantes pour notre sujet, ainsi que les impasses réalisées sur les possibilités et fonctionnalités des transistors, matières qui remplissent à elles seules de nombreuses bibliothèques.

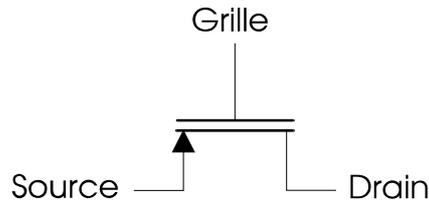
RESUME

Un transistor est un appareil électrique présentant trois connexions : le drain, la source et la grille. Son fonctionnement est celui d'un interrupteur laissant ou pas passer le courant entre la source et le drain selon la tension appliquée à la grille :

- un transistor canal n laisse passer le courant si une tension est appliquée à la grille, ne le laisse pas passer sinon. Symbole :



- un transistor canal p ne laisse passer le courant si une tension est appliquée à la grille, le laisse passer sinon. Symbole :



On peut implanter des millions de transistors, capacités, diodes et résistances sur une plaque de quelques millimètres carrés de silicium, constituant alors un circuit intégré.