

Profesor: Dr.-Ing. Paola Vega Castillo

Capítulo 4. El Transistor MOSFET

4.1 El concepto de transistor y transistor MOSFET

Un transistor se define como un dispositivo de al menos tres terminales en el cual una terminal controla el flujo de corriente entre las otras dos. La palabra transistor proviene de la combinación de los términos resistor (resistencia) y transfer (transferencia), siendo una resistencia controlada. El término Transfer indica que hay una variable de entrada y una variable de salida; una variable de control que se aplica a una terminal controlando así la corriente que fluye entre las otras 2.

El transistor MOSFET es un dispositivo electrónico semiconductor de cuatro terminales. El acrónimo MOSFET se deriva del nombre "*metal oxide semiconductor field effect transistor*". Este acrónimo se explica al descomponerlo en dos partes: MOS, por las siglas en inglés de metal-óxido-semiconductor y se refiere al sistema de materiales que al estar en contacto hace posible el efecto de campo, y FET es el acrónimo en inglés de field effect transistor. Como el nombre lo indica, el MOSFET es un dispositivo basado en el efecto de campo, es decir, se utiliza un campo eléctrico para controlar la cantidad de corriente que fluye entre dos terminales. En el caso del MOSFET, el transistor es una resistencia controlada por voltaje.

El MOSFET es un dispositivo unipolar; esto significa que el flujo de corriente se debe solamente a un tipo de portadores de carga, o más exactamente, está dominado por un tipo de portadores de carga. Contrariamente, en el caso bipolar, como en el transistor bipolar de junta, se necesita una corriente de electrones y una corriente de huecos para su operación.

El MOSFET es el dispositivo semiconductor más importante debido a que los dispositivos de efecto de campo son los que dominan la industria electrónica, representando un 80% del mercado, por ello cualquier nueva tecnología será comparada con el MOSFET.

El principio de funcionamiento del MOSFET, llamado efecto de campo, fue patentado en 1925 por Julius Lillienfeld. El transistor MOSFET fue concebida en 1928 también por Lillienfeld; la patente le fue otorgada en 1930. Sin embargo, no fue sino hasta muchos años después que la tecnología de materiales logró producir un material aislante (SiO_2) de la calidad necesaria para implementar el transistor MOSFET. Se logró fabricar una capa de dióxido de silicio suficientemente delgada para lograr el efecto de campo pero que a pesar de su reducido espesor se comportara como un buen aislante eléctrico y

con poca contaminación de iones metálicos del proceso de fabricación, pues esta contaminación altera la tensión de activación del transistor (V_{TH}).

La implementación del MOSFET ocasionó una revolución en la industria microelectrónica, hasta entonces dominada por el BJT (bipolar junction transistor). Esta revolución tomó especial fuerza con la introducción de la tecnología CMOS (complementary metal oxide semiconductor) en 1960, así como con la introducción de la teoría de escalamiento de Robert H. Dennard en los años 60's, haciendo posible la integración a gran escala que no hubiera sido posible utilizando transistores BJT.

La tecnología CMOS se originó cuando se logró fabricar una tina N en un sustrato P. Esto permitió fabricar transistores NMOS y PMOS en una sola oblea de silicio, y es la tecnología que utilizamos actualmente.

Debido al continuo escalamiento del MOSFET, para indicar el grado de avance de una tecnología se habla de generaciones tecnológicas, las cuales se diferencian por el concepto de dimensión crítica F (feature size). La dimensión crítica es la dimensión más pequeña que puede fabricarse con el silicio policristalino de compuerta en el plano de la oblea. Así, por ejemplo, en la generación de 14nm el valor de F es de 14nm.

4.2 Estructura del MOSFET

El MOSFET es un dispositivo de cuatro terminales: compuerta, drenador, surtidor (o fuente) y sustrato. Estas terminales se representan con las letras G (gate), D (drain), S (Source) y B (bulk), respectivamente. La compuerta permite controlar, por medio del efecto de campo, la cantidad de corriente que fluye entre drenador y surtidor. La estructura del MOSFET se presenta en la figura 1.

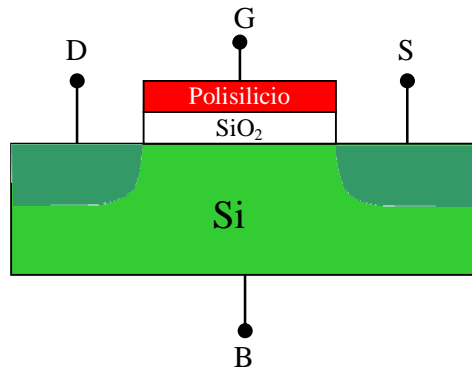


Figura 1. Sección transversal de un transistor MOSFET

El sustrato en el cual se fabrica el MOSFET es una oblea de silicio con espesor de 100 a 800 μm . Normalmente se utilizan obleas tipo P debido a que los transistores NMOS, que son los más rápidos, requieren un sustrato P. En la figura 1 el sustrato se representa con el color verde claro.

El MOSFET consiste en un sustrato de silicio dopado, con dos regiones de dopado complementario a del sustrato a cada lado. Estas regiones constituyen la fuente y el drenado del transistor y se muestran con color verde oscuro en la figura 1. El color verde significa que las regiones de difusión, al igual del sustrato, son de silicio. El tono más oscuro indica que el dopado es mayor. Como la figura 1 sugiere, la fuente y el drenador son idénticos desde el punto de vista de su construcción. La definición de cuál de las regiones constituye la fuente y cuál el drenador se obtiene al aplicar una diferencia de potencial entre ambas terminales.

Sobre el sustrato entre las regiones de difusión se coloca una capa delgada de SiO_2 , representado en la figura 1 con la capa blanca sobre el sustrato. El espesor de la capa de óxido es menor que 10nm y disminuye con cada generación tecnológica. El espesor mínimo de dióxido de silicio es de 1.2nm. Desde hace unos 7 años se reemplazó el SiO_2 por HfO_2 en los transistores de última generación.

Sobre la capa de SiO_2 se encuentra una capa de silicio policristalino, formando la compuerta del transistor; esta se representa con la capa de color rojo en la figura 1. La compuerta tiene un espesor típico de 300nm. Aunque originalmente se concibió que la compuerta del MOSFET fuera de metal, en la práctica se utiliza silicio policristalino altamente dopado, de manera que sus propiedades se acerquen a las de un metal.

La capa de óxido (SiO_2) aísla eléctricamente la superficie del silicio de la capa de silicio policristalino que constituye la compuerta, de forma que el sustrato bajo la compuerta experimente el efecto del campo eléctrico sin que haya conducción de corriente entre compuerta y sustrato. Esta capa de óxido se refiere comúnmente como “óxido de compuerta” (gate oxide), mientras que el término “polisilicio” (polysilicon), o simplemente “*poli*” (poly), es una abreviación comúnmente adoptada en la literatura para el término silicio policristalino.

La figura 2 muestra las dos posibles variaciones en el dopado complementario en la implementación del MOSFET.

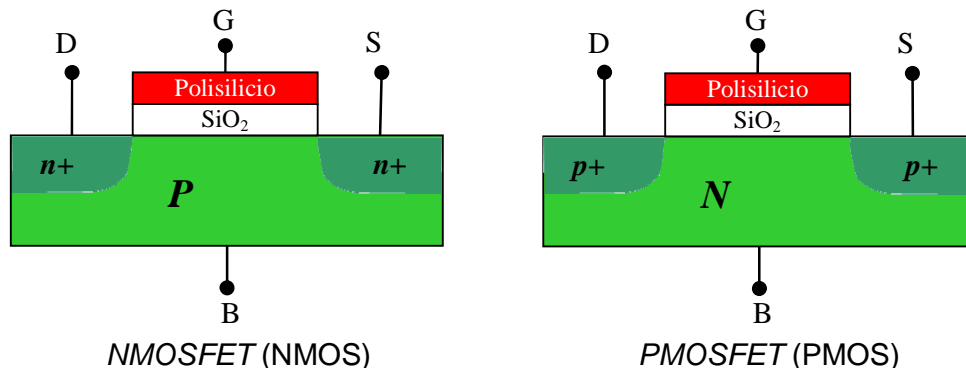


Figura 2. Sección transversal de transistores NMOS y PMOS

Como muestra la figura 2, el sustrato puede tener dopado p o dopado n , mientras que las regiones de drenador y fuente están fuertemente dopadas (pero sin llegar a ser degeneradas) complementariamente al dopado del

substrato. Estas dos variaciones dan origen a los dos tipos posibles de transistores MOSFET: NMOS y PMOS, donde el substrato está dopado respectivamente, con aceptores y donadores, mientras que al activar los transistores, el canal formado entre drenador y fuente es de tipo n en el NMOS y de tipo p en el PMOS, como se verá más adelante.

4.3 Principio de funcionamiento del MOSFET

El principio de funcionamiento del MOSFET se analizará para el caso del transistor NMOS. Este análisis podrá ser luego fácilmente extensivo al transistor PMOS.

4.3.1 Análisis cualitativo

Primeramente se examinará el transistor sin aplicar ninguna tensión en sus terminales. Como se muestra en la figura 3, existen uniones PN entre drenador y substrato y entre fuente y substrato. Estos diodos están conectados en serie por medio del substrato, pero dado que presentan direcciones de conducción contrarias, impiden el flujo de corriente entre drenador y surtidor.

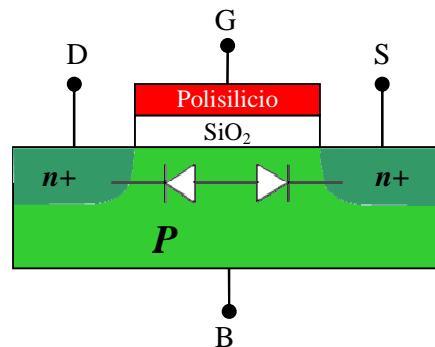


Figura 3.

Al aplicar una tensión de compuerta positiva, y con $V_{GS} > 0$, la influencia del campo eléctrico de la compuerta repele a los huecos cercanos a la superficie del semiconductor, como se muestra en la figura 4.

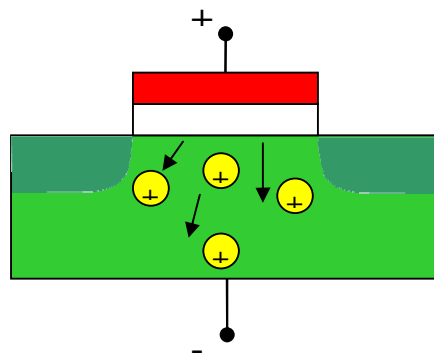


Figura 4. Principio de activación del MOSFET

De esta manera, se agotan los huecos en la superficie del sustrato en la región entre el drenador y la fuente. Este estado se conoce como agotamiento, debido al agotamiento de portadores de carga mayoritarios en la superficie del silicio entre drenador y fuente.

Conforme se continúa aumentando el voltaje de compuerta, electrones libres del sustrato se ven atraídos a la superficie. Cuando el voltaje de compuerta ha alcanzado un valor conocido como el voltaje de umbral del transistor, la concentración de portadores minoritarios en la superficie del sustrato iguala a la concentración de portadores mayoritarios con la que fue dopado el sustrato.

Cuando esto ocurre, la superficie del sustrato entre las regiones de difusión, que fue dopada de tipo p , se comporta como un material de tipo n , debido a la concentración de electrones en la superficie. Este fenómeno se conoce como inversión, puesto que ahora la superficie se comporta como un material de dopado opuesto al del sustrato.

En estas condiciones, los electrones acumulados en la superficie del sustrato entre drenador y fuente forman un canal entre drenador y fuente, permitiendo el flujo de corriente al aplicar un voltaje entre drenador y surtidor. El canal, es decir, la superficie en estado de inversión, se comporta entonces como una superficie del mismo tipo que las regiones de drenador y surtidor, por lo que ya no existe impedimento para el flujo de corriente entre ambas terminales, como se muestra en la figura 5. Una vez formado el canal, sólo se necesita una diferencia de tensión entre drenador y surtidor para que fluya corriente.

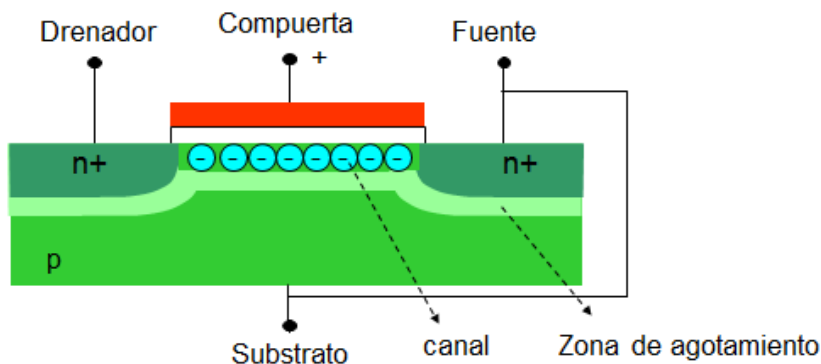


Figura 5. NMOS con el canal formado

Por otra parte, si el voltaje de compuerta aplicado es negativo, los huecos del sustrato (portadores de carga mayoritarios) son atraídos a la superficie del sustrato, aumentando la concentración de huecos en la superficie. Este fenómeno se conoce como acumulación. Puesto que no existe un canal entre drenador y fuente, la corriente entre fuente y drenador es prácticamente cero.

4.3.2 Análisis cuantitativo

Para analizar el comportamiento del MOSFET desde el punto de vista cuantitativo, debe considerarse primero el diagrama de bandas de energía del sistema metal-óxido-semiconductor (MOS), como se muestra en la figura 5.

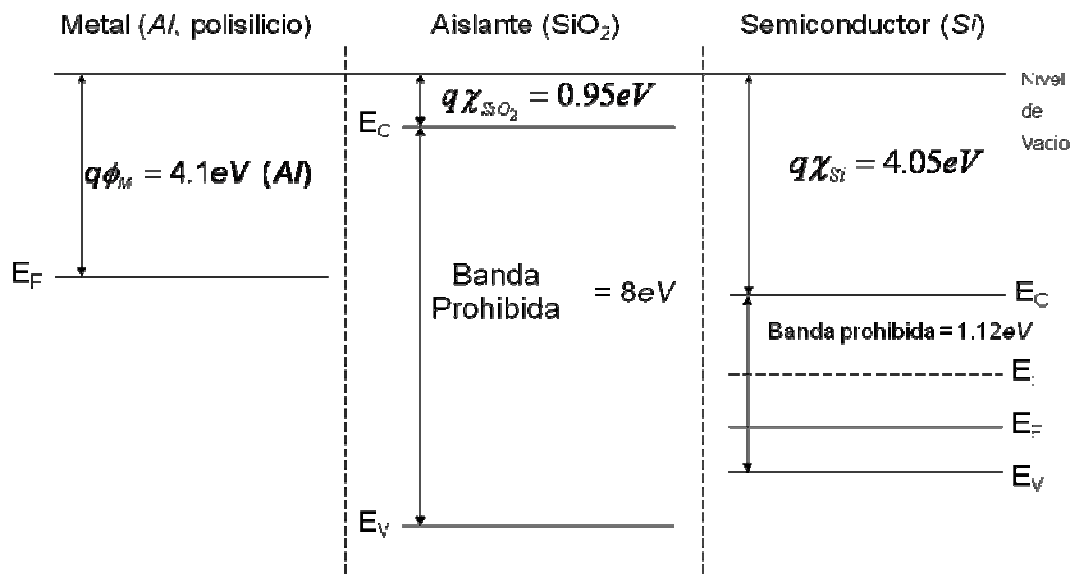


Figura 5. Diagrama de bandas de energía del sistema de materiales MOS antes del contacto

Al poner en contacto el metal, el óxido y el semiconductor, los niveles de Fermi no se alinean debido a que el óxido impide el flujo de portadores de carga entre el metal y el sustrato. En los MOSFET modernos, el metal de la compuerta ha sido reemplazado por silicio policristalino degenerado. En otras palabras, el silicio policristalino ha sido dopado tan fuertemente, que su comportamiento se aproxima al de un metal. Recuérdese aquí, que un semiconductor degenerado es aquel en el cual el nivel de dopado es tan alto que el nivel de Fermi a menos de $3kT$ de la banda de conducción (para dopado n) o de la banda de valencia (para dopado p).

Debido al potencial de contacto de la banda MOS, es decir, la diferencia de funciones de trabajo entre el material de la compuerta (metal o silicio policristalino) y el sustrato, el campo eléctrico resultante conlleva a una caída de tensión en el óxido y en la superficie del semiconductor, la cual se manifiesta en el diagrama de bandas como una deformación de bandas. El potencial intrínseco de Fermi en la cercanía de la interfaz sustrato-óxido se conoce como potencial de superficie ψ , y es la magnitud con la cual se describe el comportamiento del MOSFET.

Desde el punto de vista de V_{GS} y más exactamente del potencial de superficie ψ , se distinguen 4 regiones de operación:

- Acumulación
- Banda plana
- Agotamiento
- Inversión, en la cual se distinguen
 - Inversión débil
 - Inversión fuerte

4.3.2.1. Región de acumulación

Si se aplica una tensión V_{GS} en un transistor, tal que se atraen los portadores mayoritarios del sustrato a la región donde se formaría el canal, se obtiene la condición conocida como acumulación. El canal no puede formarse y con ello el transistor está apagado. Esto se muestra en la figura 6.

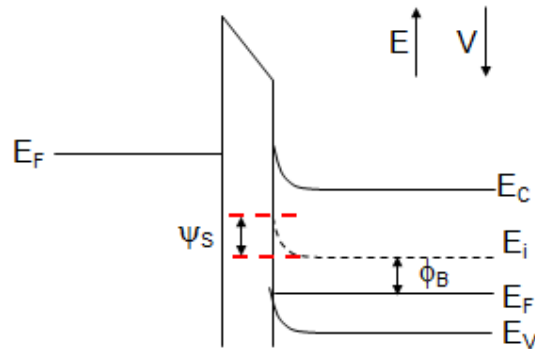


Figura 6. Diagrama de bandas de energía de un NMOS en acumulación

4.3.2.2. Región de banda plana

Para compensar la caída de tensión debido a la diferencia de funciones de trabajo, debe aplicarse una cierta tensión de compuerta, conocida como tensión de banda plana V_{FB} (flat band voltage). Al aplicar este voltaje de compuerta, la deformación de bandas desaparece. Este estado se conoce como condición de banda plana (flat band condition) y se ilustra en la Figura 7.

La caída de tensión también se debe, en parte, a cargas parásitas en el óxido. El voltaje de banda plana se calcula como:

$$V_{FB} = \phi_M - \phi_S - \frac{Q_{ox}}{C_{ox}'}$$

donde ϕ_m es la función de trabajo del material de la compuerta, ϕ_s es la función de trabajo del sustrato, Q_{ox} es la densidad de carga superficial de cargas parásitas en el óxido (en general, en el aislante) y C_{ox}' es la capacitancia de compuerta por unidad de área (F/m^2).

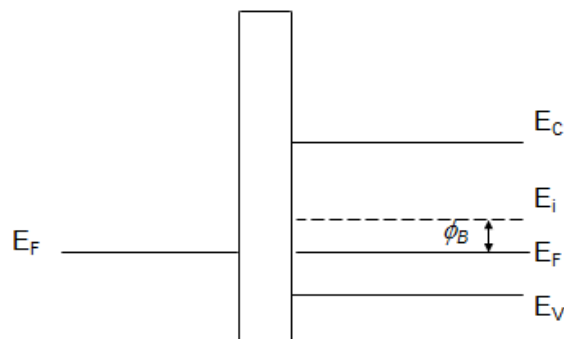


Figura 7. Diagrama de bandas de energía de un NMOS en banda plana

4.3.2.3. Región de agotamiento

Al aumentar el voltaje de compuerta, se forma una zona de agotamiento en la superficie del sustrato cercana a la interfaz $Si-SiO_2$. Al aumentar el voltaje de compuerta, los huecos son repelidos de la superficie, lo cual se manifiesta en el diagrama de bandas como un doblamiento de bandas en el que el nivel intrínseco de Fermi se acerca al nivel de Fermi, como se muestra en la Figura 8. Para $|\psi_s| < |\phi_B|$, el MOSFET se encuentra en la condición de acumulación.

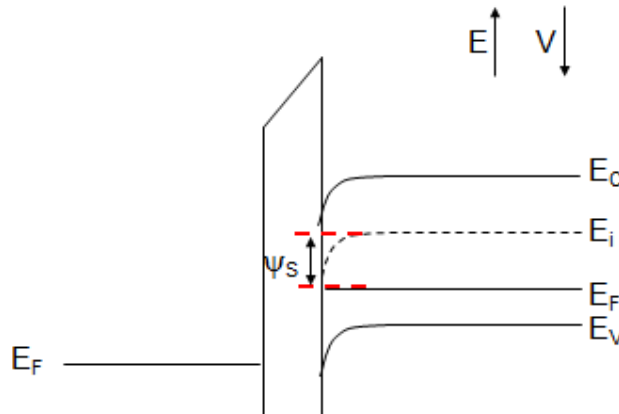


Figura 8. Diagrama de bandas de energía de un NMOS en agotamiento

4.3.2.4. Región de inversión

Si el voltaje V_{GS} continúa aumentando, el transistor entra en la condición de inversión débil, en la que $|\phi_B| \leq |\psi_s| < |2\phi_B|$.

Una vez que el potencial de superficie ha alcanzado la condición $|\psi_s| = |2\phi_B|$ se dice que el transistor está en la condición de inicio de la inversión fuerte.

Para $|\psi_s| > |2\phi_B|$, el transistor está en la condición de inversión fuerte. En el diagrama de bandas puede verse que en esta condición el nivel intrínseco de Fermi en la cercanía de la superficie está ubicado debajo del nivel de Fermi. Esto corresponde al diagrama de bandas de un material n , por lo que el sustrato, en la cercanía de la interfaz, se comporta como un material n . Esto se ilustra en la Figura 9.

Por definición, la condición de inversión es aquel estado en el cual la concentración de portadores minoritarios en la superficie del sustrato iguala la concentración de portadores mayoritarios del sustrato.

Con creciente tensión de compuerta la concentración de electrones en la superficie puede aumentar, pero la profundidad de la zona de agotamiento se mantiene aproximadamente constante.

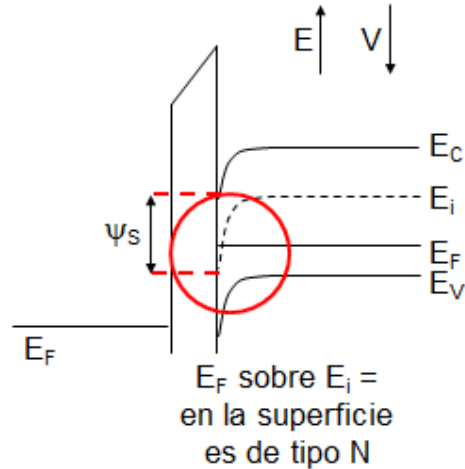


Figura 9. Diagrama de bandas de energía de un NMOS en inversión

4.3.3 Tensión de umbral V_{TH}

Por definición todos los voltajes del MOSFET toman como referencia la tensión del surtidor. En términos generales, se define la tensión de umbral V_{TH} como la tensión entre compuerta y surtidor V_{GS} necesaria para causar inversión de la superficie del sustrato, es decir, la formación del canal.

La tensión de umbral está determinada por las siguientes componentes:

1. El voltaje necesario para invertir la superficie (cambiando el potencial de superficie), $2\phi_B$
2. El voltaje necesario para compensar la carga de la zona de carga espacial, $\frac{Q_{BO}}{C_{OX}}$,
3. La diferencia de función de trabajo entre la compuerta y el sustrato, $\phi_M - \phi_S$
4. El voltaje necesario para compensar cargas parásitas en el óxido, así como en la interfaz Si-SiO₂, $\frac{Q_{OX}}{C_{OX}}$,

Los componentes 3 y 4 fueron previamente contemplados al definir el voltaje de banda plana V_{FB} .

Matemáticamente, el voltaje de umbral se expresa como:

$$V_{TH} = V_{FB} - \frac{Q_{BO}}{C_{OX}} - \frac{Q_{OX}}{C_{OX}} = \phi_M - \phi_S - 2\phi_B - \frac{Q_{BO}}{C_{OX}} - \frac{Q_{OX}}{C_{OX}}$$

cuando el sustrato y la fuente están conectados al mismo potencial, es decir $V_{BS}=0$.

4.4 Clasificación y simbología del MOSFET

El transistor MOSFET puede clasificarse de varias maneras, de acuerdo con:

- El tipo de canal, en PMOS (canal p , sustrato n) y NMOS (canal n , sustrato p).
- El modo de operación, en transistor de enriquecimiento (normalmente inactivo) y de empobrecimiento (normalmente activo).

El MOSFET de enriquecimiento es aquel en el cual el canal debe ser formado aplicando una tensión de compuerta. Esto implica para un NMOS, una tensión de umbral positiva, y para el PMOS, una tensión de umbral negativa.

Por el contrario, el MOSFET de empobrecimiento es aquel en el cual el canal ya está formado, y se aplica una tensión de compuerta para remover el canal. Esto implica, para NMOS una tensión de umbral negativa y para PMOS una tensión de umbral positiva. Esto se ilustra en la Figura 10. La tensión de umbral de un transistor puede ajustarse con un perfil de dopado especial en la región del canal aplicando implantación iónica.

El MOSFET puede representarse con símbolos de tres o cuatro terminales. En el símbolo de tres terminales, el sustrato y el surtidor están conectados al mismo potencial, razón por la cual no se muestra explícitamente la conexión el sustrato. En este símbolo, la flecha indica dirección de corriente técnica entre drenador y surtidor. Si no se usan flechas, un círculo sin rellenar se añade a la compuerta de los transistores PMOS. Estos símbolos se encuentran en la Figura 11. El transistor de empobrecimiento se distingue por tener una línea más gruesa entre drenador y surtidor.

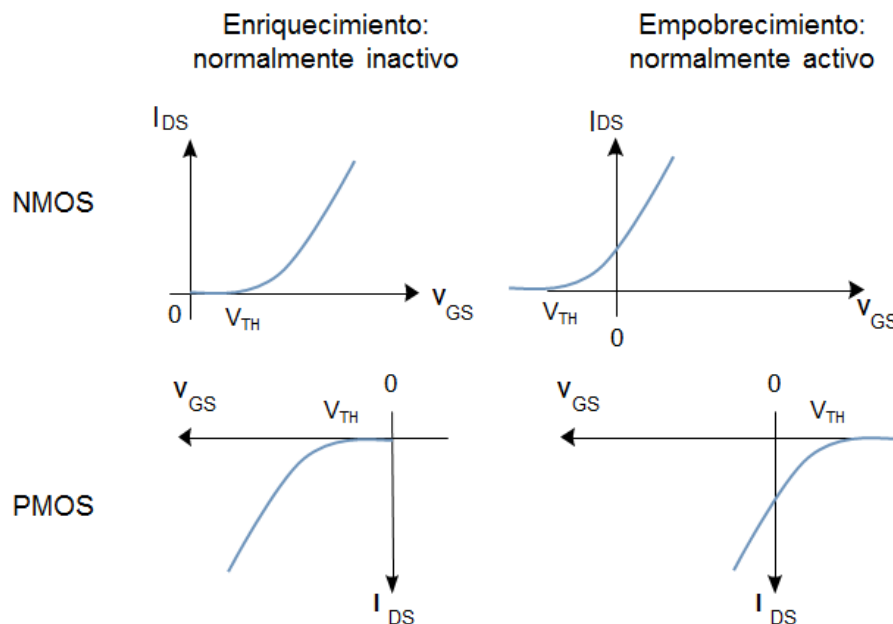


Figura 10. Curvas características de entrada de los transistores NMOS y PMOS de enriquecimiento y empobrecimiento

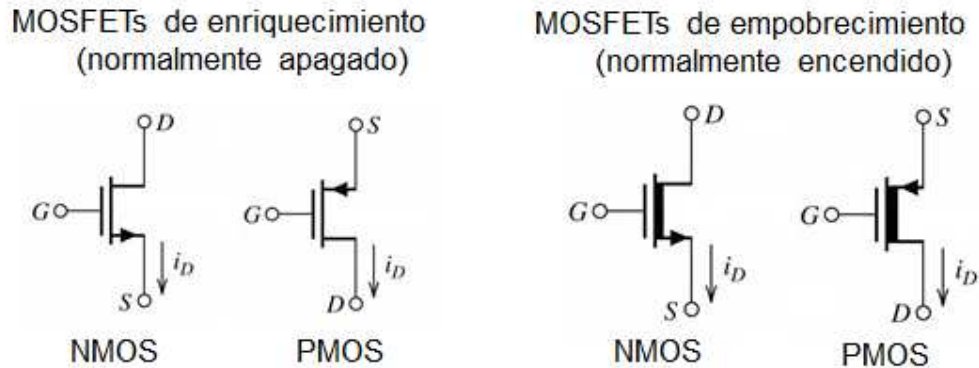


Figura 11. Símbolos de tres terminales del transistor MOSFET

Por otra parte, el símbolo de cuatro terminales muestra explícitamente la conexión del sustrato. En este caso, la flecha en la terminal de sustrato apunta de la región P a la región N. Estos símbolos se encuentran en la Figura 12. El transistor de empobrecimiento se distingue por tener una línea más continua entre drenador y surtidor.

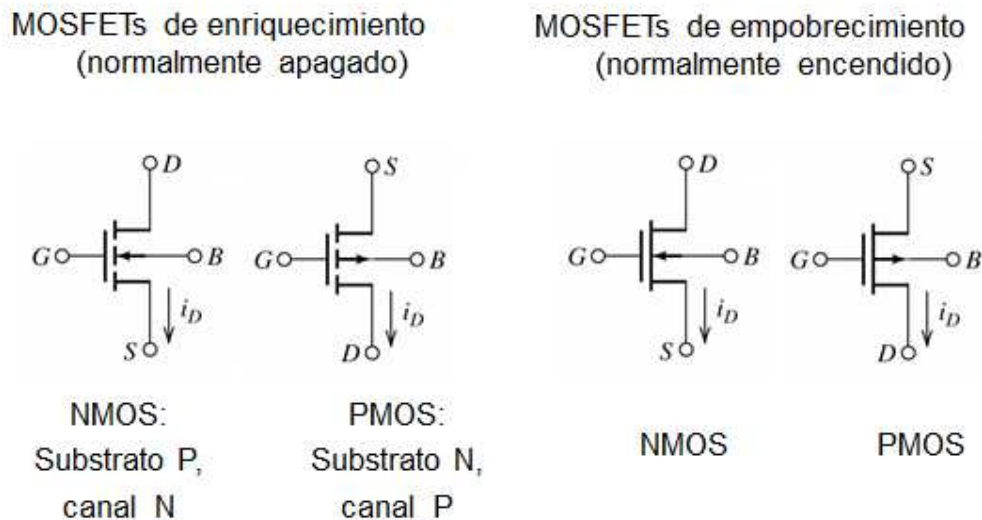


Figura 12. Símbolos de cuatro terminales del transistor MOSFET

4.5 Curvas características del MOSFET

Para estudiar las curvas características del MOSFET es necesario definir primero las dimensiones involucradas en las ecuaciones. Estas dimensiones se muestran en la figura 13: el largo del canal L , el ancho del canal W y el espesor del dieléctrico de compuerta (óxido) t_{ox} . Cuando se habla de generaciones tecnológicas del MOSFET, la dimensión crítica F es, como una primera aproximación, el valor del largo del canal L .

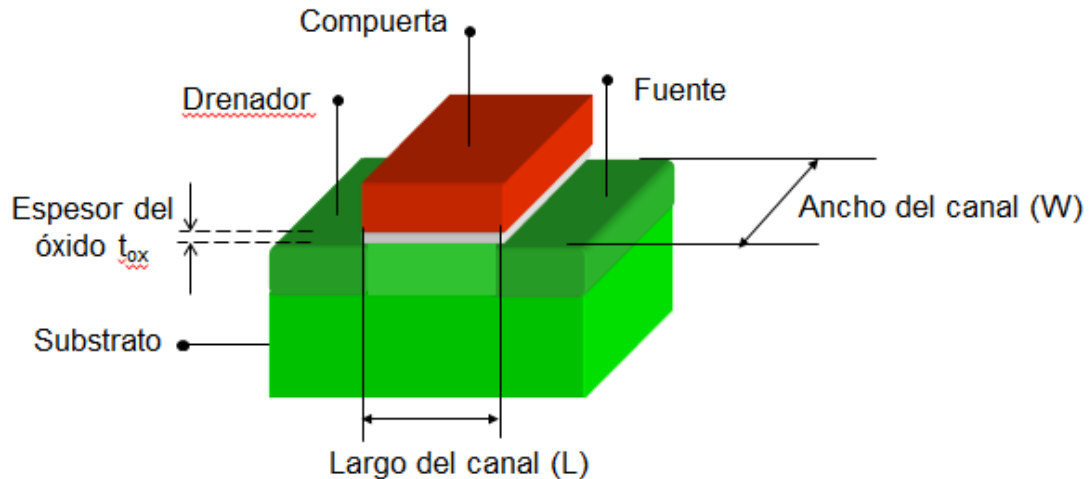


Figura 13. Dimensiones importantes en un transistor MOSFET

Con base en la curva IV de salida del MOSFET, pueden distinguirse tres importantes regiones de operación para aplicaciones analógicas, como se muestra en la Figura 14:

- La región de corte
- La región lineal
- La región de saturación

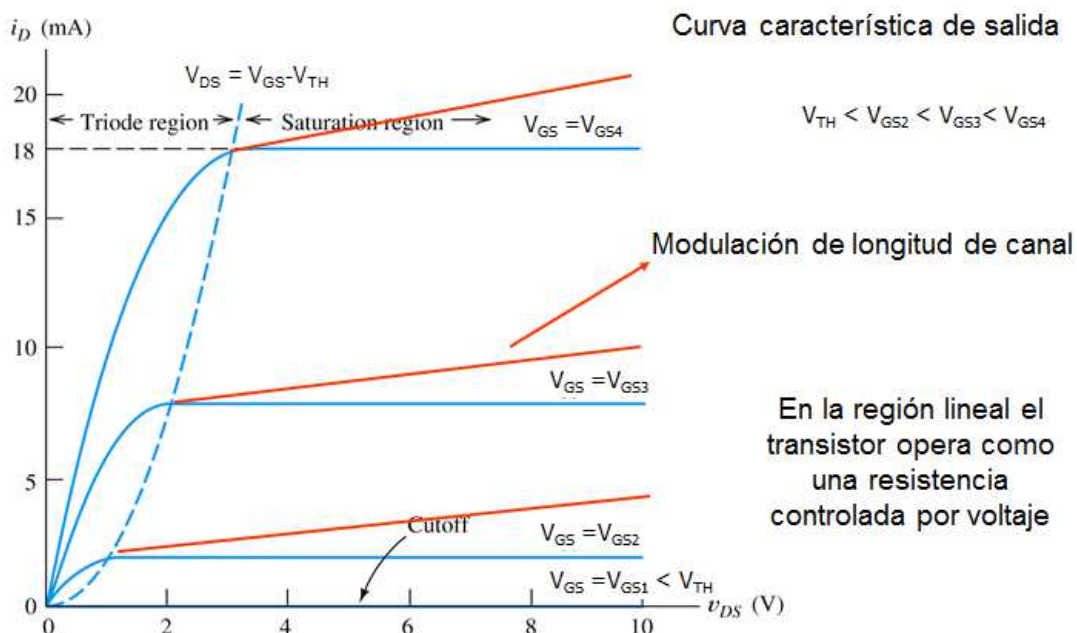


Figura 14. Curva característica de salida del MOSFET

4.5.1 Región de corte

Cuando $V_{GS} < V_{TH}$, el MOSFET se encuentra operando en la región de corte o subumbral. Aunque idealmente la corriente I_{DS} es cero cuando el transistor está en la región de corte, en realidad existe un flujo de corriente de fuga entre

drenador y surtidor. Esta corriente se conoce como corriente de subumbral y se calcula como sigue:

$$I_{DS} = I_{D0} e^{\frac{(V_{GS}-V_{TH})}{mV_t}} \left[1 - e^{\frac{-V_{DS}}{V_t}} \right]$$

donde $I_{DS,0} = I_{DS}(V_{GS} = V_{TH}) \frac{W}{L} \approx 0.1 \mu A \frac{W}{L}$ y $m = 1 + \frac{C_{dep}}{C_{ox}}$, donde C_{dep} es la capacitancia de agotamiento.

Para $V_{DS} > 4V_t$, la corriente de subumbral puede aproximarse como sigue:

$$I_{DS} \approx I_{D0} e^{\frac{(V_{GS}-V_{TH})}{mV_t}}$$

La curva característica de entrada se muestra en la Figura 15. Esta misma curva, pero ahora mostrando la región de subumbral, se muestra en la Figura 16. Un parámetro de gran importancia relacionado con la corriente de subumbral es la llamada pendiente de subumbral S , que se define como:

$$S = \left[\frac{d(\log I_{DS})}{dV_{GS}} \right]^{-1} = \ln 10 \cdot V_t \cdot m$$

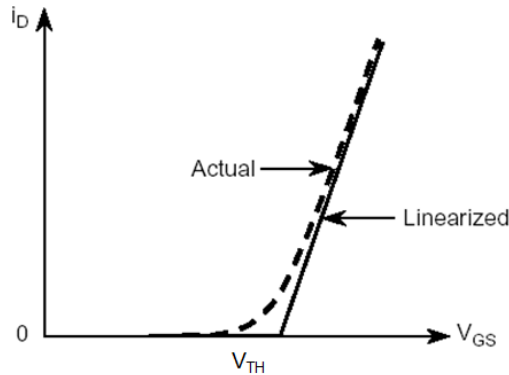


Figura 15. Curva característica de entrada del MOSFET

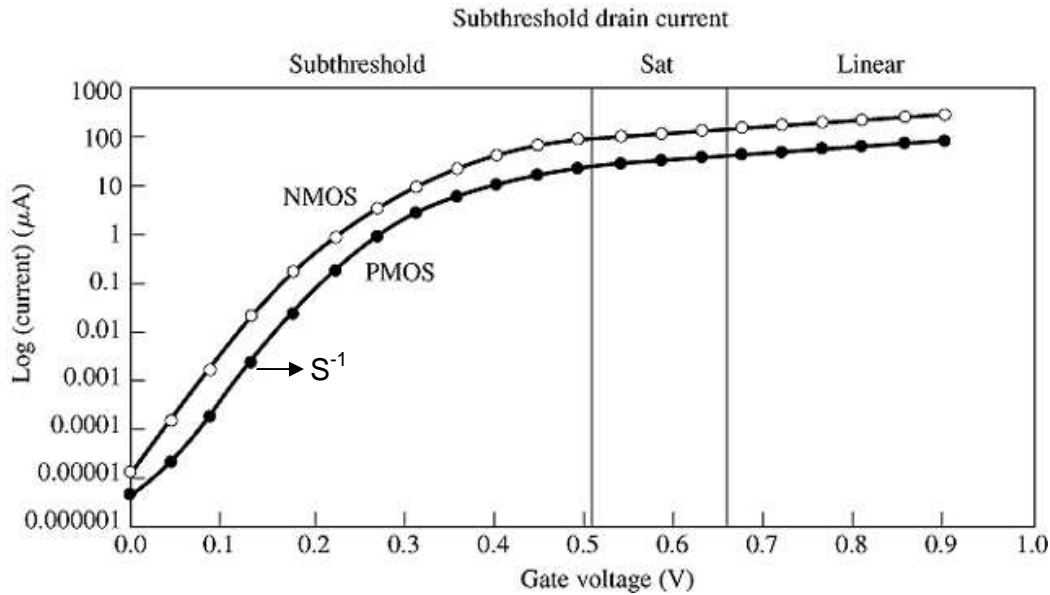


Figura 16. Curva característica de entrada del MOSFET mostrando la región de subumbral

La pendiente de subumbral S representa el cambio en V_{GS} necesario para obtener una variación de una década (un factor de 10) en I_{DS} . Las unidades de S son mV/dec . S varía con la temperatura. Valores típicos de S a temperatura ambiente ($\approx 300K$) son de 80 a 85 mV/dec . Por otra parte, S es aproximadamente 100 mV/dec a 100°C.

El valor ideal de S se obtiene al reducir la capacitancia de agotamiento a cero. En este caso,

$$S = \ln 10 \cdot V_t \cdot m = \ln 10 \cdot V_t \cdot 1 = 60 \frac{mV}{dec}$$

con lo cual el valor ideal de S es de 60 mV/dec .

4.5.2 Región lineal

En la región lineal el transistor se comporta como una resistencia (controlada), como lo muestra la curva característica de salida (I_{DS} vs V_{DS}) de la figura 14.

En MOSFET opera en la región lineal en la condición

$$V_{GS} - V_{TH} > V_{DS} \quad , \quad V_{GS} \geq V_{TH}$$

La ecuación que describe I_{DS} en la región lineal es

$$I_{DS} = K(V_{GS} - V_{TH} - \frac{V_{DS}}{2})V_{DS}$$

El valor de la resistencia del MOSFET en la región lineal se obtiene con la definición de resistencia:

$$R = \frac{\partial V}{\partial I} \quad \text{en este caso} \quad R_{ch} = \left(\frac{\partial I_{DS}}{\partial V_{DS}} \right)^{-1}$$

donde R_{ch} es la resistencia del MOSFET en la zona lineal y es llamada resistencia de canal. Aplicando la definición anterior,

$$\frac{\partial I_{DS}}{\partial V_{DS}} = K(V_{GS} - V_{TH} - V_{DS})$$

Por lo tanto:

$$R_{ch} = \frac{1}{K(V_{GS} - V_{TH} - V_{DS})}$$

lo cual para $V_{DS} \ll V_{DS\ sat}$ puede aproximarse a:

$$R_{ch} \approx \frac{1}{K(V_{GS} - V_{TH})}$$

4.5.3 Región de saturación

Por otra parte, en la región de saturación la corriente es, en el caso ideal, independiente del voltaje, manteniendo un valor constante. Este comportamiento corresponde al de una fuente de corriente. El paso de la región lineal a la región de saturación está determinado por el voltaje de saturación $V_{DS\ sat}$.

$$V_{DS\ sat} = V_{GS} - V_{TH} \quad , \quad V_{GS} \geq V_{TH}$$

El MOSFET se encuentra en la región de saturación cuando $V_{DS} \geq V_{DS\ sat}$.

En la práctica, la modulación de ancho de canal se manifiesta como un incremento de la corriente de drenador con valores crecientes de V_{DS} . Este comportamiento puede modelarse por medio de una fuente de corriente real. La modulación de largo de canal consiste en el acortamiento del canal al aumentar V_{DS} , y dicho acortamiento disminuye la resistencia del canal, aumentando la corriente del transistor.

Comparando el modelo de una fuente de corriente real y la ecuación de la corriente de drenador en la región de saturación, se tiene:

$$I_{DS} = \underbrace{\frac{1}{2} K(V_{GS} - V_{TH})^2}_{\text{constante}} + \underbrace{\frac{1}{2} K(V_{GS} - V_{TH})^2 (\lambda V_{DS})}_{I = f(V_{DS}) \rightarrow \text{resistencia}}$$

El término independiente de V_{DS} representa una fuente de corriente ideal, mientras que el término dependiente de V_{DS} representa una resistencia y

permite modelar el efecto de largo de canal. En la ecuación, esto se manifiesta en el término que contiene V_{DS} y el coeficiente de modulación de largo de canal λ . Como puede observarse en la ecuación, las unidades de λ son V^{-1} .

En el circuito equivalente de la figura 13, el efecto de modulación de largo de canal está representado por la resistencia de salida $r_o = \frac{1}{\lambda I_{DSsat}}$, con

$$I_{DSsat} = \frac{1}{2} K (V_{GS} - V_{TH})^2.$$

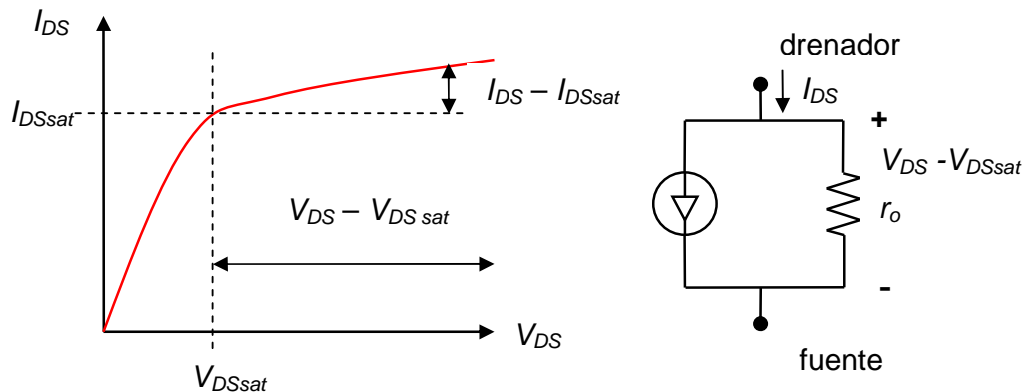


Figura 13. Modelo analógico en la región de saturación

4.6 Polarización del sustrato

Puesto que el MOSFET es un dispositivo de cuatro terminales, debe considerarse también la polarización del sustrato. Como puede apreciarse en la Figura 14, la presencia de las regiones de difusión en el sustrato da origen a diodos parásitos, los cuales deben mantenerse inactivos para operar el transistor. Para ello, debe conectarse el sustrato P al potencial más bajo disponible en el circuito (por ejemplo, GND), y el sustrato N al potencial más alto disponible en el circuito (por ejemplo, V_{DD}).

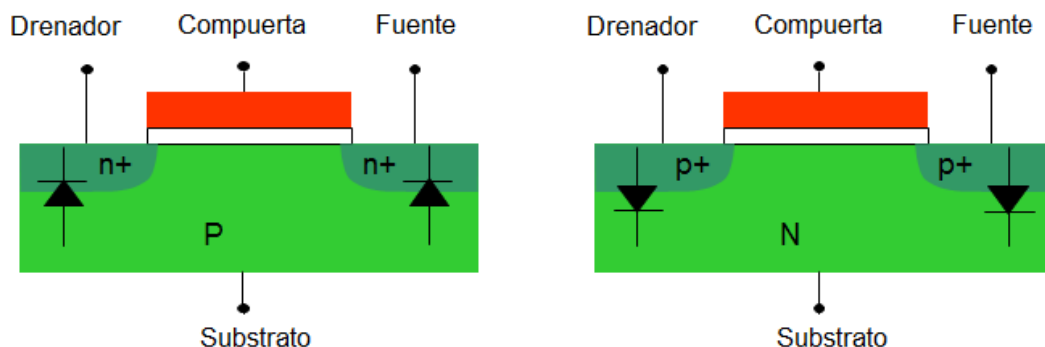


Figura 14. Diodos parásitos difusión-sustrato

La Figura 15 muestra el diagrama de bandas de energía drenador-sustrato-surtidor de un transistor NMOS sin tensión aplicada y con el sustrato

conectado al mismo potencial que el surtidor. Como se observa en la figura, existen barreras de energía (potencial de contacto) que impide el flujo de portadores de carga.

Al aplicar una tensión V_{GS} para lograr la inversión del sustrato, el diagrama de bandas de energía drenador-sustrato-surtidor de un transistor NMOS se modifica como se ilustra en la Figura 16. Como se observa en la figura, ahora al aplicar una tensión V_{DS} diferente de cero se posibilita el flujo de portadores de carga. En este diagrama el sustrato está conectado al mismo potencial que el surtidor.

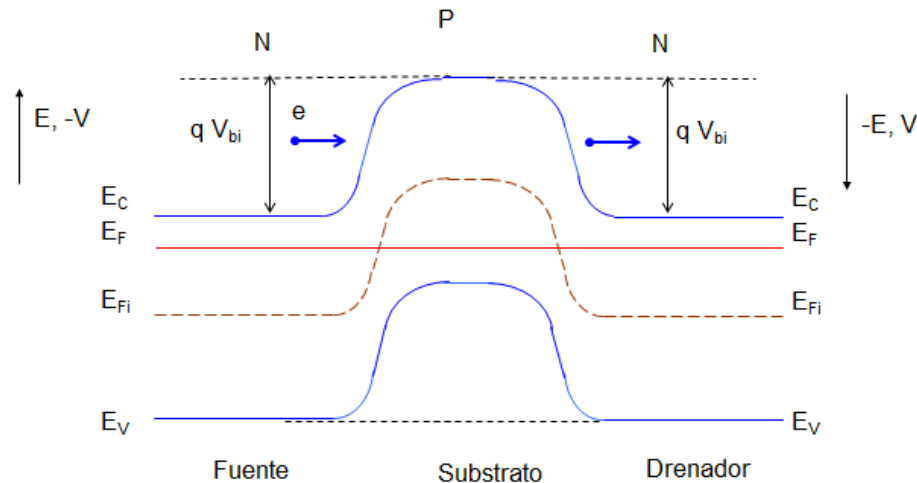


Figura 15. Diagrama de bandas drenador-sustrato-surtidor de un transistor NMOS sin tensión aplicada

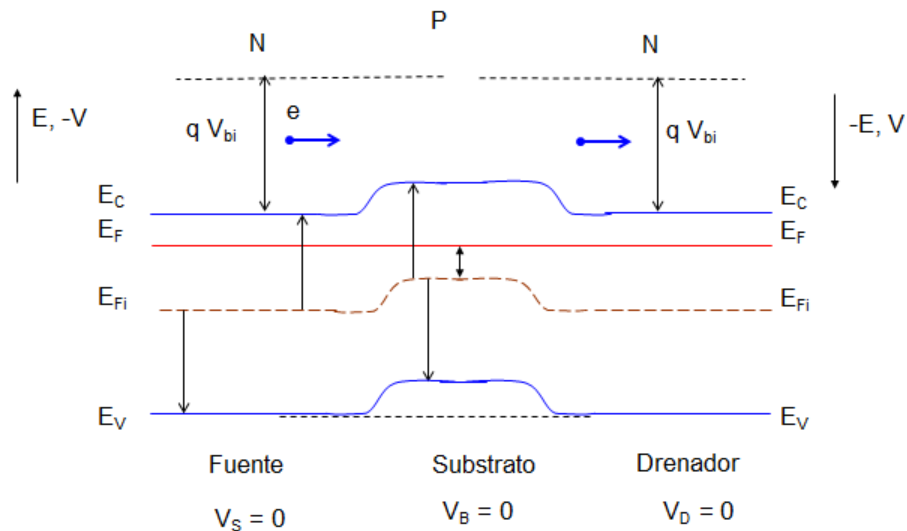


Figura 16. Diagrama de bandas drenador-sustrato-surtidor de un transistor NMOS con tensión V_{GS} positiva aplicada

Hasta ahora se ha estudiado el funcionamiento del MOSFET asumiendo que la tensión del sustrato es igual a la tensión de surtidor. Al aplicar una tensión $V_{BS} \neq 0$ se altera la tensión de umbral del transistor. Si la tensión aplicada al sustrato favorece la inversión del sustrato, la tensión de umbral disminuye. Por el contrario, si la tensión aplicada al sustrato favorece la acumulación, la

tensión de umbral aumenta. Esto se ilustra en las figuras 17 y 18, respectivamente. Esta alteración de la tensión de umbral se conoce como efecto de cuerpo o efecto de sustrato.

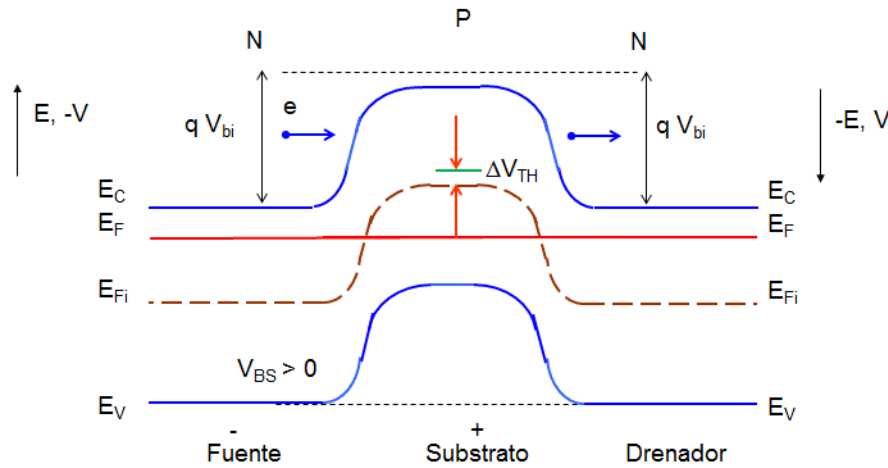


Figura 17. Diagrama de bandas drenador-sustrato-surtidor de un transistor NMOS con efecto de cuerpo disminuyendo la tensión de umbral

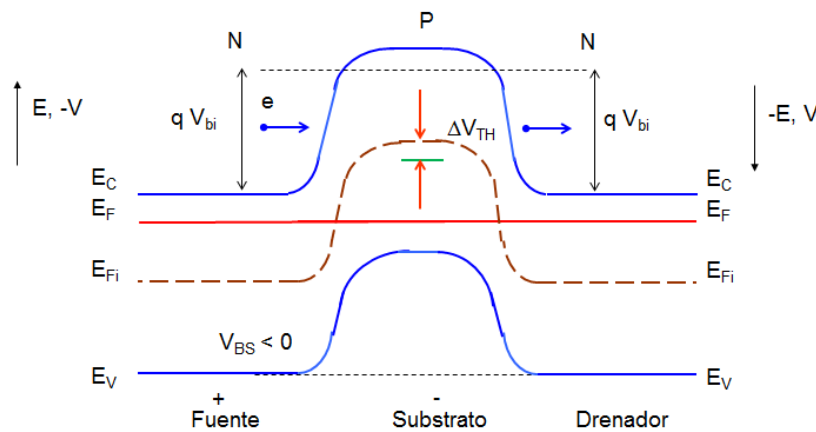


Figura 18 Diagrama de bandas drenador-sustrato-surtidor de un transistor NMOS con efecto de aumentando la tensión de umbral

La curva característica de entrada en la figura 19 muestra el efecto de cuerpo para un transistor NMOS de enriquecimiento. El transistor opera en saturación. En este caso, al tener una tensión V_{BS} negativa, la tensión de umbral aumenta y la curva característica se desplaza hacia la derecha, es decir, a valores más positivos de V_{GS} .

El voltaje de umbral tomando en cuenta el efecto de cuerpo se calcula como:

$$V_{TH} = V_{TH0} + \gamma(\sqrt{2\phi_B - V_{BS}} - \sqrt{2\phi_B})$$

Donde el factor γ es el coeficiente de efecto de cuerpo (coeficiente de efecto de sustrato) y ϕ_B es la diferencia entre el nivel de Fermi intrínseco y el nivel de Fermi del sustrato.

La permitividad del silicio ϵ_{Si} es $11.7\epsilon_0$.

$$\gamma = \frac{\sqrt{2qN_A\epsilon_{Si}}}{C'_{ox}}$$

$$\phi_B = V_t \cdot \ln\left(\frac{N_A}{n_i}\right)$$

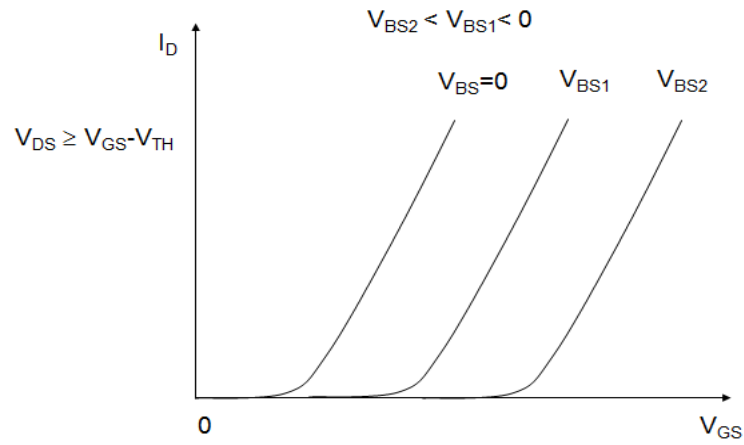


Figura 19. Curva característica de entrada de un NMOS mostrando el efecto de cuerpo