

1. FUNDAMENTOS DE OPERACIÓN DEL TRANSISTOR MOS.

En este capítulo se estudiará la estructura MOS (Metal Oxide Semiconductor); se analizará su comportamiento, el cuál permitirá comprender el funcionamiento del transistor MOS. Se abordarán los fenómenos electrostáticos que ocurren dentro de la estructura MOS bajo polarización. Posteriormente se describirá la operación del MOSFET en régimen lineal y en saturación, que permitirá definir algunos parámetros y variables fundamentales para su tratamiento. También, se analizará de manera detallada la parte fundamental del MOSFET: la conductividad del canal. Por último, se describirán los efectos indeseados que ocurren en el MOSFET al reducir la longitud del canal.

1.1. Estructura MOS.

La estructura MOS consiste en un metal, referido como compuerta, un óxido (típicamente óxido de silicio SiO_2) y un semiconductor (tipo N o tipo P) llamado sustrato. El material semiconductor más usado es el silicio (Si). La estructura MOS se muestra en la Figura 1.1.

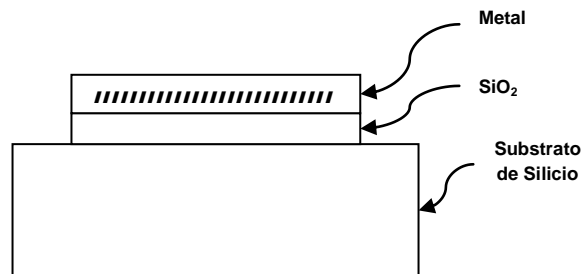


Figura 1.1. Estructura MOS [9] .

En ausencia de polarización, en el semiconductor hay una distribución homogénea de portadores, por lo que no se presenta carga en ninguna de las interfaces: metal-óxido y óxido-semiconductor.

Al aplicar cierto voltaje al metal (usualmente llamado voltaje de compuerta V_G) y poniendo a tierra el sustrato (Figura 1.2), se pueden obtener tres condiciones de operación en la estructura MOS: condición de acumulación, condición de empobrecimiento y condición de inversión [9].

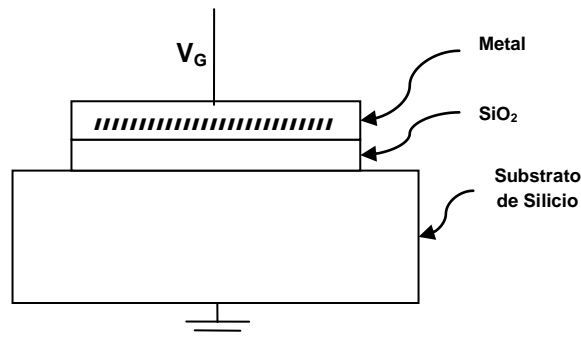


Figura 1.2. Voltaje V_G aplicado a la estructura MOS.

Con el voltaje aplicado, la estructura MOS se comporta como un capacitor de placas planas y paralelas ya que aparece carga tanto en el metal como en el sustrato, formando el capacitor.

La carga por unidad de área del capacitor [10] MOS está dada por:

$$Q = CV_G \quad (1-1)$$

donde C es la capacitancia y V_G es el voltaje aplicado entre la compuerta (metal) y el sustrato. Es importante notar que la concentración de portadores en el sustrato varía con el voltaje aplicado, específicamente en la interfaz óxido-semiconductor, de la que se hará referencia de ahora en adelante como interfaz del óxido simplemente.

La capacitancia por unidad de área [10] presente en el óxido se define como:

$$C_{ox} = \frac{K_{ox}\epsilon_0}{x_{ox}} \quad (1-2)$$

donde K_{ox} es la constante dieléctrica del óxido, ϵ_0 es la permitividad del vacío y x_{ox} es el espesor del óxido. Se puede ver que esta capacitancia es constante.

A continuación se mostrarán las condiciones de operación de la estructura MOS al aplicar cierto voltaje. Se utilizará un sustrato tipo P durante todo el análisis.

Al aplicar un voltaje negativo a la estructura MOS, se presenta la condición de acumulación [9], como se muestra en la Figura 1.3.

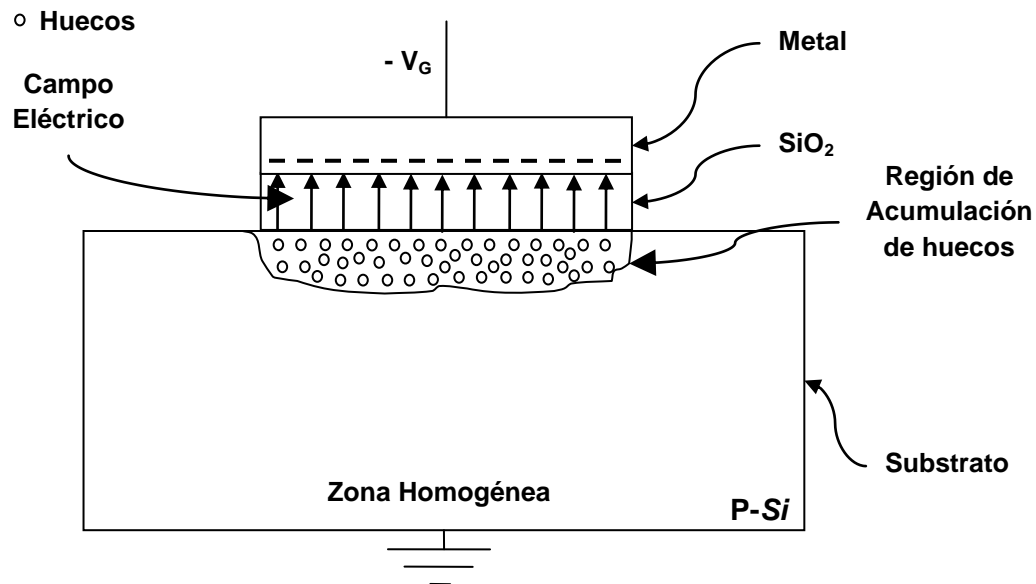


Figura 1.3. Condición de acumulación en la estructura MOS con un semiconductor tipo P.

En esta condición los portadores mayoritarios (huecos), del semiconductor tipo P, comienzan a acumularse en la interfaz del óxido. Esto se debe al campo eléctrico, producto del voltaje negativo aplicado. La región de acumulación que se muestra en la Figura 1.3 se comporta como un semiconductor tipo P más dopado (P^+). También, se muestra una zona homogénea, en la cuál la distribución de carga no varía.

Sí ahora se aplica un voltaje positivo a la estructura MOS, se presenta la condición de empobrecimiento [9] como se muestra en la Figura 1.4.

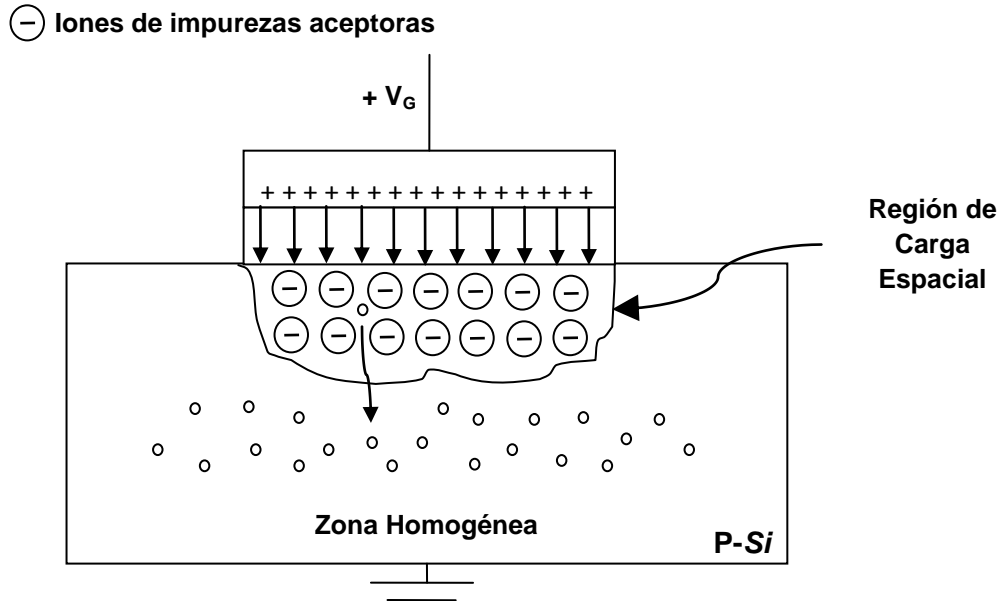


Figura 1.4. Condición de empobrecimiento en la estructura MOS con un semiconductor tipo P.

En esta condición, los portadores mayoritarios son repelidos de la zona cercana al óxido, debido al campo eléctrico. Esto produce una región de carga espacial (RCE) producto de los iones de impurezas [9]. También se tiene una capacitancia por unidad de área, asociada a la RCE en el silicio dada por:

$$C_{si} = \frac{K_{si} \epsilon_0}{x_d} \quad (1-3)$$

donde K_{si} es la constante dieléctrica del silicio y x_d es el ancho de la RCE. Como se puede observar, esta capacitancia cambia con x_d . Al incrementar V_G , el ancho de la RCE aumenta, por lo que C_{si} muestra una dependencia con V_G .

Sí se continua incrementando el voltaje positivo aplicado se llega a la condición de inversión [9], como se muestra en la Figura 1.5.

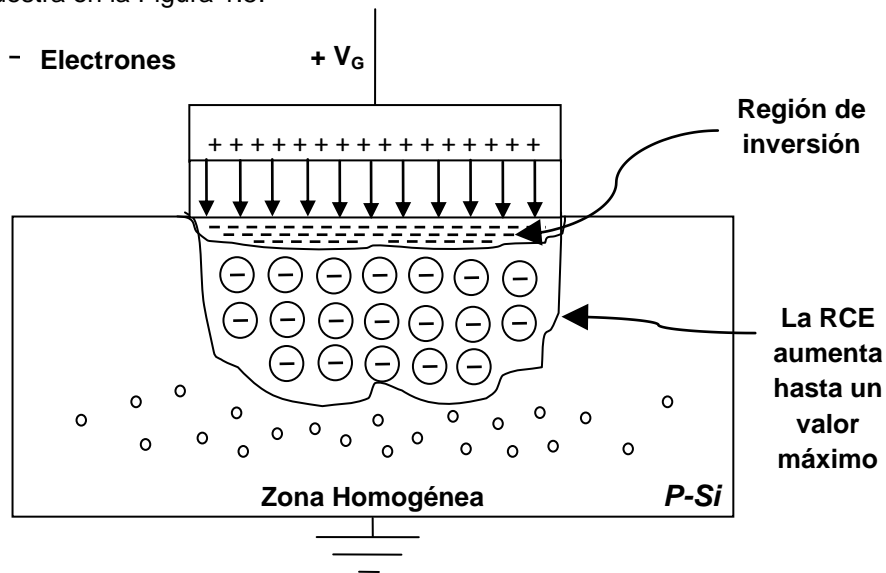


Figura 1.5. Condición de inversión en la estructura MOS con un semiconductor tipo P.

Con este incremento de voltaje la RCE sigue aumentando. En la interfaz del óxido, el campo eléctrico confina a los electrones libres generados en la superficie del silicio, esto produce que la región entre el sustrato y el óxido, se comporte como un semiconductor tipo N. Por lo tanto, se ha invertido el tipo de semiconductor, es decir, el tipo de conductividad. El comienzo de este proceso se conoce como inversión débil [9].

La RCE se incrementa hasta un límite, en el cuál un incremento en el voltaje aplicado, produce un incremento en la concentración de carga móvil en la región de inversión y por lo tanto la RCE ya no se incrementa. Cuando ocurre esto se tiene una inversión fuerte [9].

En el caso de tener una estructura MOS con un semiconductor tipo N, se logra la condición de acumulación al aplicar un voltaje positivo al metal y poniendo a tierra el sustrato. Al aplicar un voltaje negativo se obtiene la condición de empobrecimiento. Con un voltaje negativo más intenso se obtiene la condición de inversión, produciendo una región entre el óxido y el sustrato tipo P.

Para observar mejor los fenómenos que ocurren en la estructura MOS, se analizará el diagrama de bandas de energía de la estructura MOS, en el que se verán de manera clara las tres condiciones de operación mostradas anteriormente. Para tal efecto, se considerará una estructura MOS ideal.

En la Figura 1.6 se presenta el diagrama de bandas de energía de la estructura MOS sin polarizar. Nuevamente se emplea un semiconductor tipo P.

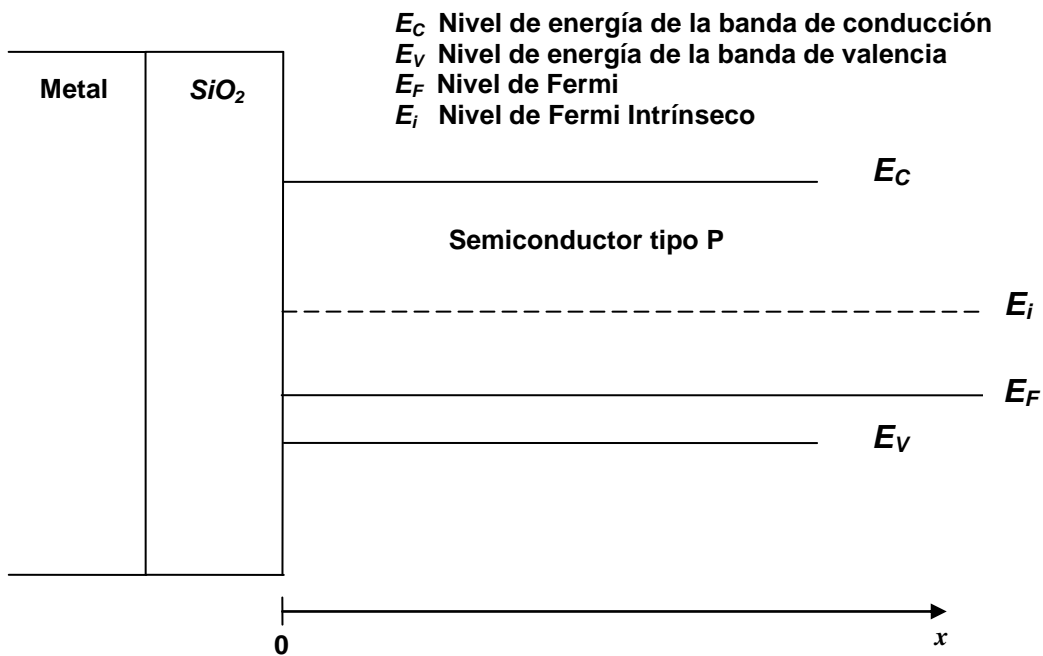


Figura 1.6. Diagrama de Bandas de la estructura MOS ideal sin voltaje aplicado.

Cuando no hay voltaje aplicado, la estructura MOS está en condición de bandas planas (Flat Banda-FB). Además, independientemente del voltaje aplicado, el nivel de Fermi permanece constante. La distribución de carga contenida en el semiconductor permanece homogénea [9].

En la Figura 1.7 se muestra el diagrama de bandas de la estructura MOS cuando está en condición de acumulación.

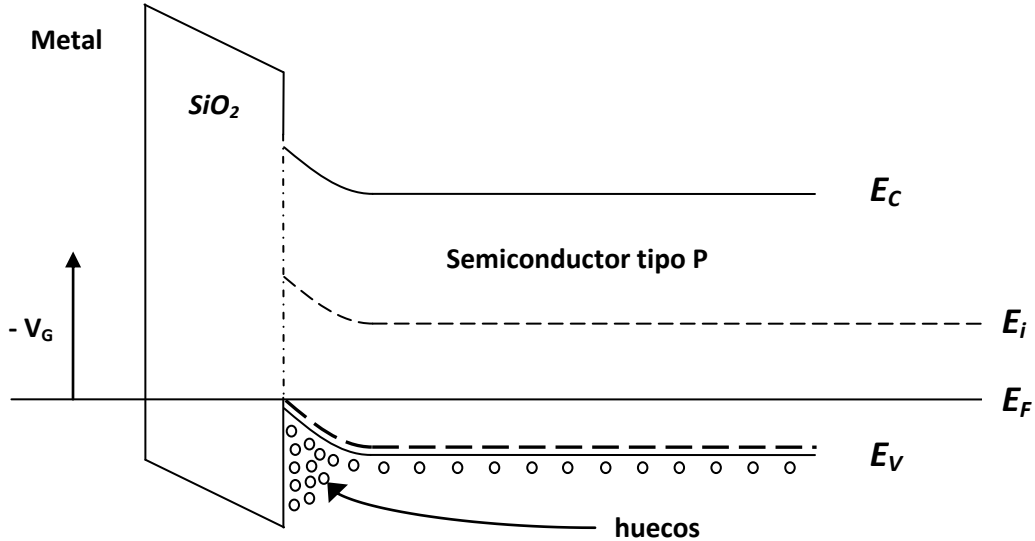


Figura 1.7. Diagrama de bandas de la estructura MOS ideal en condición de acumulación ($V_G < 0$) [11].

Como se puede observar en la Figura 1.7, al aplicar un voltaje negativo hay un doblamiento de bandas. Se ve el efecto de acumulación de portadores mayoritarios (huecos), ya que en la interfaz del óxido el nivel de Fermi (E_F) está más próximo al nivel de energía de la banda de valencia (E_V), por lo que tenemos una región de semiconductor tipo P mas dopada (P^+) [9].

En la Figura 1.8 se muestra el diagrama de bandas de la estructura MOS al aplicar un voltaje positivo, en este caso las bandas se doblan hacia abajo.

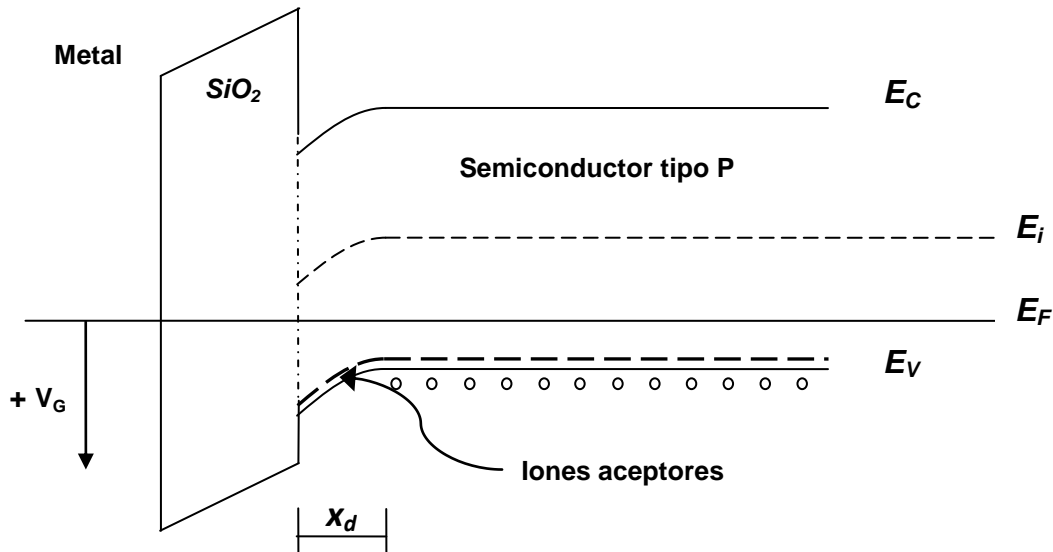


Figura 1.8. Diagrama de bandas de la estructura MOS ideal en condición de empobrecimiento ($V_G > 0$) [11].

Con el voltaje positivo aplicado, se genera una región de carga espacial debido a los iones de impurezas. Se ve el efecto de empobrecimiento, ya que en la interfaz del óxido, el nivel de Fermi (E_F) se encuentra más alejado del nivel de energía de la banda de valencia (E_V), por lo que hay una pérdida de portadores mayoritarios [9]. Así, la carga por unidad de área Q_s , contenida en el semiconductor, estará dada por la carga contenida dentro de la región de carga espacial [11]:

$$Q_s = -qN_B x_d \quad (1-4)$$

donde q es la carga del electrón, N_B es la concentración de impurezas del sustrato y x_d es el ancho de la región de carga espacial.

En la Figura 1.9 se muestra el diagrama de bandas de la estructura MOS en condición de inversión, es decir, cuando se aplica un voltaje positivo intenso.

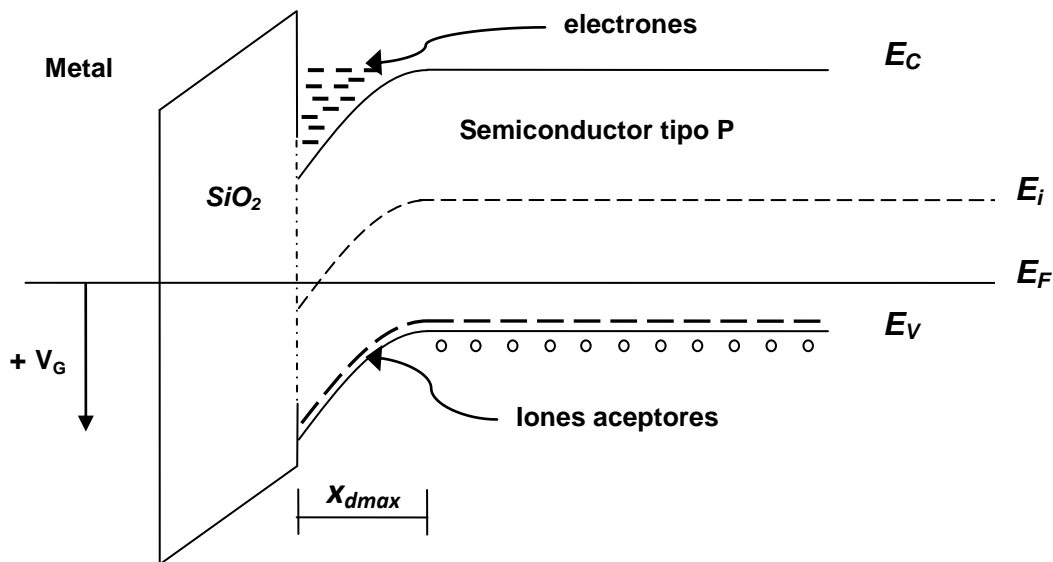


Figura 1.9. Diagrama de bandas de la estructura MOS ideal en condición inversión ($V_G > 0$) [11].

Si se comienza a incrementar el voltaje positivo aplicado, el ancho de la región de carga espacial aumentará. Además, la variación total de potencial electrostático en el silicio también aumentará. Sin embargo, como las bandas comienzan a doblarse cada vez más, el nivel de energía de la banda de conducción (E_C) eventualmente se irá acercando al nivel de Fermi (E_F), adicionalmente el nivel intrínseco (E_i) se ubicará por debajo del nivel de Fermi. Debido a lo anterior, la concentración de electrones cerca de la interfaz del óxido pronto aumentará considerablemente. Después de esto, la mayor cantidad de carga inducida en el semiconductor consistirá de carga Q_n debida a electrones confinados en una estrecha capa, es decir, la región de semiconductor cercana al óxido estará invertida, ahora se comportará como una región de semiconductor tipo N [11].

Una vez que se ha formado la capa de inversión, el ancho de la región de carga espacial alcanzará un máximo, debido a que bajo inversión fuerte, un pequeño incremento en el voltaje aplicado a la compuerta, producirá un mayor doblamiento de las bandas, y como resultado, un incremento de la carga móvil dentro de la capa de inversión. Así, en condición de inversión fuerte, la carga por unidad de área inducida en el semiconductor estará dada por:

$$Q_s = Q_n - qN_B x_{dmax} \quad (1-5)$$

donde x_{dmax} denota el ancho máximo de la región de carga espacial [11].

Para describir de una forma mejor la variación de carga en el sustrato al aplicar un voltaje, se puede emplear la ecuación de Poisson [10], que relaciona la densidad de carga en el semiconductor con el potencial aplicado ϕ .

$$\nabla^2 \phi = -\frac{\rho}{K_{Si} \epsilon_0} \quad (1-6)$$

donde ϕ es el potencial, ρ es la densidad de carga en el semiconductor y K_{Si} es la constante dieléctrica del semiconductor.

Para simplificar los cálculos se considerará que la variación de carga en el semiconductor sólo se da en el eje x (Figura 1.10).

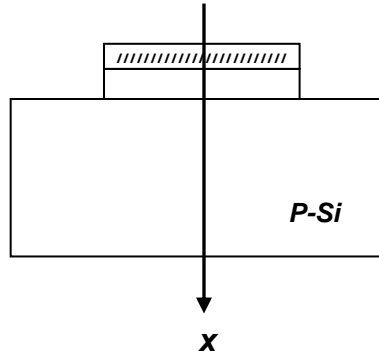


Figura 1.10. Variación de carga en el eje x de la estructura MOS.

Además se considerará que solo se presentan fenómenos a lo largo del eje x [9]. Por lo que:

$$\frac{d^2 \phi}{dx^2} = -\frac{\rho(x)}{K_{Si} \epsilon_0} \quad (1-7)$$

$$\rho(x) = q[-N_A + N_D - n(x) + p(x)] \quad (1-8)$$

$$p(x) = p_{p0} e^{-\frac{q\phi(x)}{kT}} \quad (1-9)$$

$$n(x) = n_{p0} e^{\frac{q\phi(x)}{kT}} \quad (1-10)$$

N_A es la concentración de impurezas aceptoras, N_D es la concentración de impurezas donadoras, $p(x)$ es la variación de la concentración de huecos en dirección del eje x , $n(x)$ es la variación de la concentración de electrones en dirección del eje x , k es la constante de Boltzmann, T es la temperatura, n_{p0} es la concentración de electrones en el material tipo P en equilibrio y p_{p0} es la concentración de huecos en el material tipo P en equilibrio.

También se considerará que hay neutralidad de carga en el semiconductor, por lo que la densidad de carga en todo el volumen del semiconductor será cero [11], es decir:

$$\rho(x) = q[-N_A + N_D - n(x) + p(x)] = 0$$

Esto implica que:

$$N_A - N_D = p_{p0} - n_{p0} \quad (1-11)$$

Sustituyendo la ecuación (1-9), (1-10) y (1-11) en la ecuación (1-8):

$$\rho(x) = q \left[-p_{p0} + n_{p0} + p_{p0} e^{-\frac{q\phi(x)}{kT}} - n_{p0} e^{\frac{q\phi(x)}{kT}} \right] \quad (1-12)$$

Entonces la ecuación de Poisson que se debe resolver es:

$$\frac{d^2 \varphi}{dx^2} = -\frac{q}{K_{si} \varepsilon_0} \left[-p_{p0} + n_{p0} - n_{p0} e^{\frac{q\varphi(x)}{kT}} + p_{p0} e^{-\frac{q\varphi(x)}{kT}} \right] \quad (1-13)$$

El campo eléctrico en función del potencial eléctrico [10] está dado por:

$$E = -\frac{d\varphi}{dx} \quad (1-14)$$

Así, sustituyendo el campo eléctrico E, en la ecuación (1-13) se tiene:

$$E dE = -\frac{q}{K_{si} \varepsilon_0} \left[-p_{p0} + n_{p0} - n_{p0} e^{\frac{q\varphi(x)}{kT}} + p_{p0} e^{-\frac{q\varphi(x)}{kT}} \right] dV \quad (1-15)$$

Al integrar y ordenar los términos se obtiene la siguiente ecuación:

$$E^2 = -\frac{2p_{p0}q}{K_{si} \varepsilon_0} \left\{ -\frac{kT}{q} e^{-\frac{q\varphi(x)}{kT}} - \varphi(x) - \frac{n_{p0}}{p_{p0}} \left[\frac{kT}{q} e^{\frac{q\varphi(x)}{kT}} + \varphi(x) \right] \right\} + C \quad (1-16)$$

La constante de integración C, se obtiene considerando la condición de frontera: cuando $x \rightarrow \infty$ el potencial eléctrico $\varphi(x) = 0$ por lo que el campo eléctrico también es cero ($E = 0$).

$$C = -\frac{2p_{p0}q}{K_{si} \varepsilon_0} \frac{kT}{q} \left[1 + \frac{n_{p0}}{p_{p0}} \right] \quad (1-17)$$

Sustituyendo C en la ecuación (1-16) y reorganizando los términos:

$$E^2 = -\frac{2p_{p0}q}{K_{si} \varepsilon_0} \left(-\frac{kT}{q} \right) \left\{ e^{-\frac{q\varphi(x)}{kT}} + \frac{q\varphi(x)}{kT} - 1 + \frac{n_{p0}}{p_{p0}} \left[e^{\frac{q\varphi(x)}{kT}} - \frac{q\varphi(x)}{kT} - 1 \right] \right\} \quad (1-18)$$

Definiendo F [9] como:

$$F = \left\{ e^{-\frac{q\varphi(x)}{kT}} + \frac{q\varphi(x)}{kT} - 1 + \frac{n_{p0}}{p_{p0}} \left[e^{\frac{q\varphi(x)}{kT}} - \frac{q\varphi(x)}{kT} - 1 \right] \right\}^{\frac{1}{2}} \quad (1-19)$$

Sustituyendo el valor de F en la ecuación (1-18) y reorganizando los términos:

$$E^2 = \frac{2q^2 p_{p0}}{K_{si} \varepsilon_0 kT} \left(\frac{kT}{q} \right)^2 F^2 \quad (1-20)$$

Se puede sustituir la longitud de Debye L_D^1 en la ecuación (1-20):

$$E^2 = \frac{1}{L_D} \left(\frac{kT}{q} \right)^2 F^2 \quad (1-21)$$

donde:

$$L_D = \frac{kT K_{si} \varepsilon_0}{2q^2 N_B} \quad (1-22)$$

Sí ahora se aplica la Ley de Gauss [10] a la ecuación (1-21) y se toma el valor del campo eléctrico en la interfaz óxido-semiconductor, es decir, aplicando:

¹ La longitud de Debye es la longitud del campo localizado dentro del semiconductor.

$$E(0) = \frac{Q_s}{K_{si} \epsilon_0} \quad (1-23)$$

Se obtiene la variación de carga en la interfaz óxido-semiconductor [9], dada por:

$$Q_s = \frac{K_{si} \epsilon_0 kT}{qL_D} F \quad (1-24)$$

En la Figura 1.11 se presenta una grafica representativa de la variación de carga en la interfaz del óxido en función del potencial eléctrico [9].

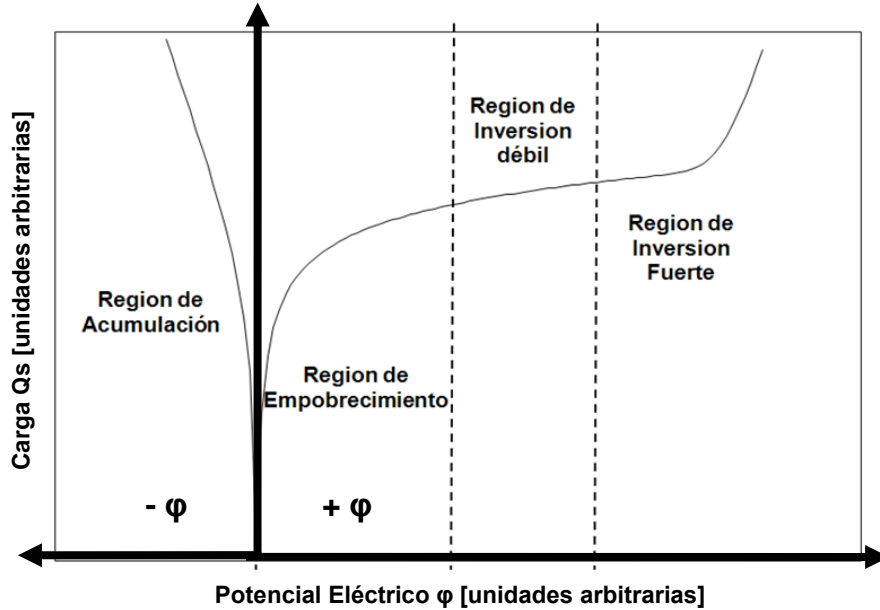


Figura 1.11. Variación de carga en la interfaz óxido-semiconductor en función del potencial eléctrico.

En la gráfica se observan las regiones de acumulación, empobrecimiento, inversión débil e inversión fuerte. Para un potencial $\phi > 2\phi_F$ (donde ϕ_F se define como el potencial de Fermi²) se presenta la región de inversión fuerte, este potencial corresponde al doblamiento total de las bandas [11].

A fin de determinar la magnitud del voltaje necesario para alcanzar la inversión fuerte, es necesario considerar que el voltaje aplicado a la compuerta se repartirá entre el óxido y el semiconductor de la estructura MOS [11]. Así:

$$V_G = V_{ox} + V_s \quad (1-25)$$

V_{ox} es el potencial en el óxido y V_s el potencial eléctrico en el semiconductor dado por:

$$V_s = \frac{qN_B x_d^2}{2K_{si} \epsilon_0} \quad (1-26)$$

En ausencia de carga localizada en la interfaz entre el óxido y el semiconductor, la ley de Gauss requiere que el desplazamiento eléctrico sea continuo en esta interfaz, es decir:

$$K_{si} \epsilon_0 E_s = K_{ox} \epsilon_0 E_{ox} \quad (1-27)$$

² Potencial de Fermi $\phi_F = (E_i - E_F)/q$

Si no hay carga presente dentro del óxido, el campo eléctrico es uniforme [11] y esta dado por:

$$E_{ox} = \frac{V_{ox}}{x_{ox}} \quad (1-28)$$

donde x_{ox} es el espesor del óxido. El campo eléctrico en el silicio, según la ley de Gauss [11], es:

$$E_s = -\frac{Q_s}{K_{si}\epsilon_0} \quad (1-29)$$

Con una combinación de estas últimas tres ecuaciones se obtiene el voltaje presente en el óxido [11], de la siguiente forma:

$$V_{ox} = -\frac{x_{ox}}{K_{ox}\epsilon_0} Q_s = -\frac{Q_s}{C_{ox}} \quad (1-30)$$

Donde $C_{ox} \equiv K_{ox}\epsilon_0/x_{ox}$ es la capacitancia por unidad de área del óxido [11].

Entonces el voltaje en la compuerta es:

$$V_G = -\frac{Q_s}{C_{ox}} + V_s \quad (1-31)$$

Y para inversión fuerte [11] es necesario que:

$$V_s = 2\phi_F \quad (1-32)$$

Así, el voltaje para tener una inversión fuerte en la estructura MOS, también llamado voltaje de umbral V_T [11], estará dado por:

$$V_G = V_T = -\frac{Q_s}{C_{ox}} + 2\phi_F \quad (1-33)$$

Ahora, si se toma en cuenta la aproximación de empobrecimiento [9], en donde se considera que en la región de carga espacial no hay portadores móviles:

$$Q_s \equiv Q_B \quad (1-34)$$

$$Q_B = -qN_B x_d \quad (1-35)$$

donde Q_B es la carga de impurezas ionizadas. Entonces el voltaje para inversión fuerte quedará como:

$$V_T = 2\phi_F - \frac{Q_B}{C_{ox}} \quad (1-36)$$

Estructura MOS real.

Hasta ahora se ha tratado la estructura MOS como ideal. Ahora se analizarán algunos efectos que permitirán modelar la estructura MOS de forma más real.

En el modelado de la estructura MOS real se toman en cuenta básicamente dos elementos importantes: la diferencia en las funciones de trabajo (del metal y semiconductor) y la presencia de carga en el óxido [9].

La energía de los electrones en el nivel de Fermi, en el metal y el semiconductor de una estructura MOS, son diferentes. Tal diferencia de energía es usualmente expresada como una diferencia en

las funciones de trabajo, que es la energía necesaria para extraer un electrón desde el nivel de Fermi [11].
 En la Figura 1.12 se muestran las diferencias en los niveles de Fermi y las funciones de trabajo de los materiales de la estructura MOS (los materiales no están en contacto).

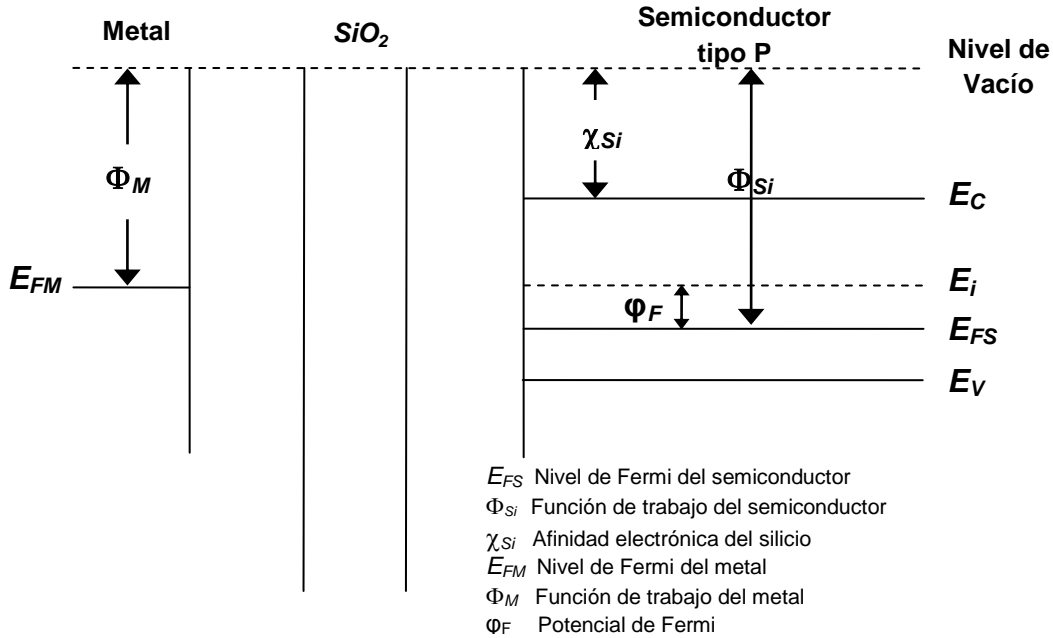


Figura 1.12. Funciones de trabajo del metal y el semiconductor (Φ_M y Φ_{Si}).

La función de trabajo del metal (Φ_M) es la diferencia de energía entre el nivel de vacío y el nivel de Fermi del metal [9].

La afinidad electrónica del semiconductor (χ_{Si}) es la diferencia de energía entre el nivel de vacío y el nivel de energía de la banda de conducción del semiconductor [9].

La función de trabajo del silicio (Φ_{Si}) es la diferencia de energía entre el nivel de vacío y el nivel de Fermi del silicio [9]. Es una cantidad que varía con la concentración de impurezas, de la siguiente forma:

$$\Phi_{Si} = \chi_{Si} + \frac{1}{2} E_g + \varphi_F \quad (1-37)$$

donde $\varphi_F = \frac{kT}{q} \ln \left(\frac{N_B}{n_i} \right)$; n_i es la densidad intrínseca de portadores.

La banda prohibida en el óxido, es mucho más grande que la del semiconductor ($E_g = E_C - E_V$), dadas sus características dieléctricas.

Las diferencias mostradas anteriormente, hacen que la condición de bandas planas no se mantenga en la estructura MOS sin polarización. Por lo tanto, las diferencias de las funciones de trabajo entre el metal y el semiconductor provocan la presencia de un potencial en el óxido de silicio [11], como se ve en la Figura 1.13.

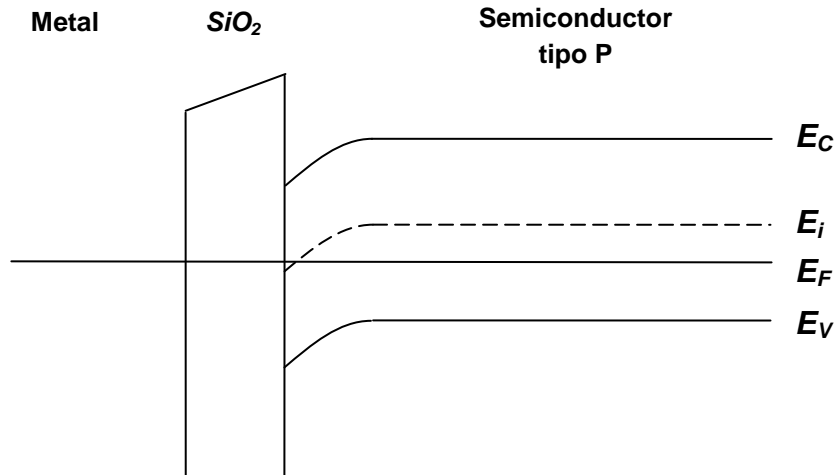


Figura 1.13. Efecto de la diferencia en las funciones de trabajo del metal y el semiconductor [11].

Para eliminar este efecto, se requiere aplicar un voltaje a la compuerta que compense las diferencias para llegar de nuevo a la condición de bandas planas, como se muestra en la Figura 1.14. Este voltaje requerido para establecer la condición de bandas planas es llamado voltaje de Flat-Band V_{FB} [11], claramente se ve que es la diferencia de las funciones de trabajo que se ha indicado:

$$V_G = V_{FB} = \phi_M - \phi_{Si} \equiv \phi_{MS} \quad (1-38)$$

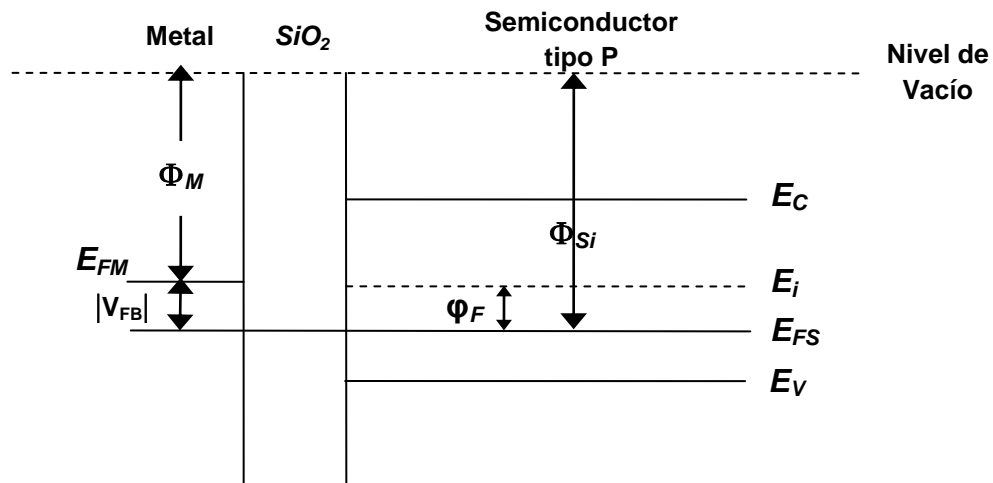


Figura 1.14. Voltaje de Flat Band (V_{FB}) para restablecer la condición de bandas planas [11].

Otro aspecto que se debe tomar en cuenta es la presencia de carga dentro del óxido. Esto ocurre durante el proceso de fabricación, ya que diversas cargas pueden quedar atrapadas en el óxido que forma la estructura MOS. Esta carga es independiente del voltaje que se aplique a la compuerta, pero influye sobre el comportamiento de la estructura [9].

Los defectos del óxido originan un plano de carga, por lo que se producen estados energéticos dentro de la banda prohibida del silicio para mantener la neutralidad de carga. Así, aparece una carga imagen en el silicio generando dipolos, lo que ocasiona un campo eléctrico que produce discontinuidades. Se debe aplicar un voltaje en la compuerta para compensar esta carga [9].

Los estados no deseados que aparecen, provocan un ligero doblamiento de bandas, formando un pequeño empobrecimiento de la interfaz del óxido, si la carga dentro del óxido es positiva. Para restablecer la condición de bandas planas se necesita aplicar un voltaje [11] dado por:

$$V_G = V_{FB} = -\frac{qN_T}{C_{ox}} \quad (1-39)$$

donde N_T es la densidad total de carga en el óxido.

Con estos fenómenos tratados, en la estructura MOS real, se puede deducir que el voltaje aplicado a la estructura MOS para restablecer la condición de bandas planas estará dado por:

$$V_{FB} = \Phi_{MS} - \frac{qN_T}{C_{ox}} \quad (1-40)$$

Como se ha mostrado hasta ahora, la estructura MOS es de gran importancia para el funcionamiento del transistor MOS, ya que ésta es la base de su fabricación y de varios dispositivos de estado sólido de efecto de campo usados en la electrónica analógica y digital.

1.2. Estructura del Transistor MOS.

Entre los dispositivos más importantes derivados de la estructura MOS se encuentra el transistor de efecto de campo conocido como MOSFET (Metal Oxide Semiconductor Field-Effect Transistor). Un dispositivo cuyo funcionamiento se basa en el efecto de campo eléctrico superficial [12].

El MOSFET (Figura 1.15) consiste en un sustrato de silicio tipo P dentro del cual hay dos regiones difundidas tipo N, llamadas surtidor (Source-S) y drenador (Drain-D). La región entre el surtidor y el drenador está bajo la influencia de un contacto de metal, llamado compuerta (Gate-G). Si se aplica un voltaje positivo intenso a la compuerta, la región de la superficie del sustrato tipo P puede ser invertida, esto inducirá un canal tipo N conectando el surtidor y el drenador. La conductividad de este canal puede ser modulada variando el voltaje de compuerta [11].

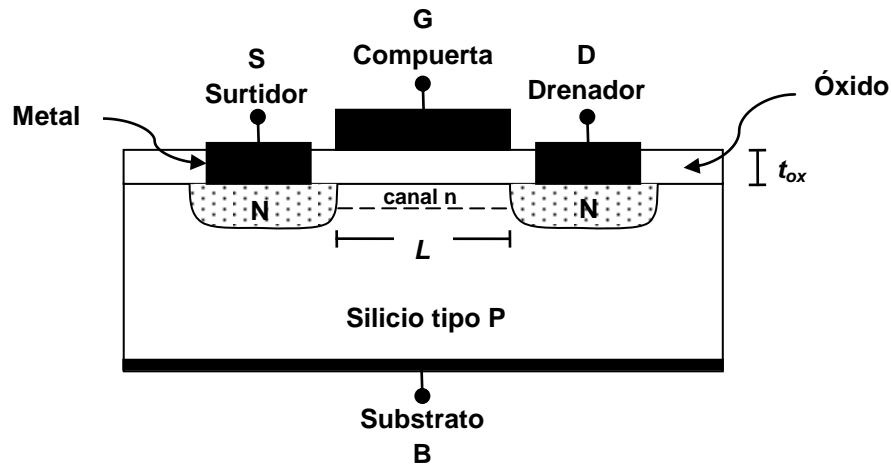


Figura 1.15. Esquema básico del transistor MOSFET [12].

De acuerdo al tipo de conductividad del canal, hay transistores MOS canal N y canal P. Además, existen transistores MOSFET de enriquecimiento o inversión en donde se aplica un voltaje a la compuerta para generar el canal que une el surtidor y drenador. También hay transistores MOSFET de agotamiento o empobrecimiento, en donde el canal se forma mediante una difusión adicional durante el proceso de fabricación. En esta tesis se trabajará con transistores de inversión.

1.3. Operación del Transistor MOS.

Como primer paso para observar el funcionamiento del MOSFET, se considerará un voltaje positivo intenso aplicado a la compuerta V_G para inducir el canal entre el surtidor y el drenador, como se muestra en la Figura 1.16. Con un pequeño voltaje positivo en el drenador V_D , se polarizará en inversa la unión PN que se forma entre la región del drenador y el substrato, por lo que sólo existirá corriente a lo largo del canal, el cuál se comportará como una resistencia [11]. Esta resistencia estará dada por:

$$R_{ch} = \rho \frac{L}{A} = \rho \frac{L}{Wt_i} \quad (1-41)$$

ρ es la resistividad del canal, L es la longitud del canal, $A = Wt_i$ es el área transversal del canal dada por el producto del ancho del canal W y su espesor t_i .

La resistividad está dada por:

$$\rho = \frac{1}{q\mu_n n} \quad (1-42)$$

donde μ_n es la movilidad de los portadores mayoritarios y n es el número de portadores mayoritarios por unidad de área en el canal. Sustituyendo la resistividad en la ecuación (1-41):

$$R_{ch} = \frac{1}{q\mu_n n} \frac{L}{Wt_i} = \frac{1}{\mu_n Q_n} \frac{L}{Wt_i} \quad (1-43)$$

donde Q_n es la carga total por unidad de área en el canal. Esta carga es controlada por el voltaje de compuerta V_G , en otras palabras, se puede observar que se tiene una resistencia controlada por el voltaje de compuerta V_G , que corresponde a la operación básica del MOSFET [9].

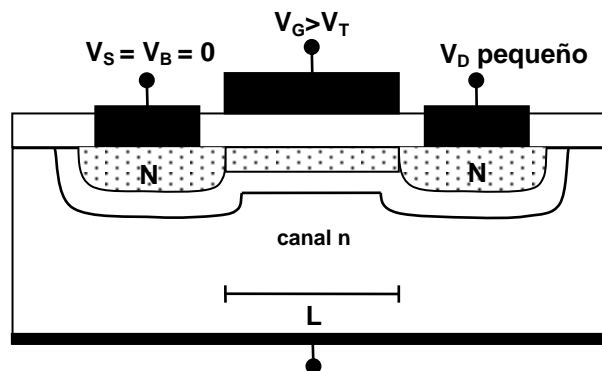


Figura 1.16. Esquema básico del MOSFET con V_D pequeño; la resistencia del canal es constante [12].

1.3.1. Región lineal.

Cuando el voltaje de drenador V_D es pequeño, la resistencia del canal permanecerá constante, por lo que es de esperarse que las características de corriente-voltaje (I_D - V_D), al variar el voltaje de compuerta, sean prácticamente lineales, como se ve en la Figura 1.17.

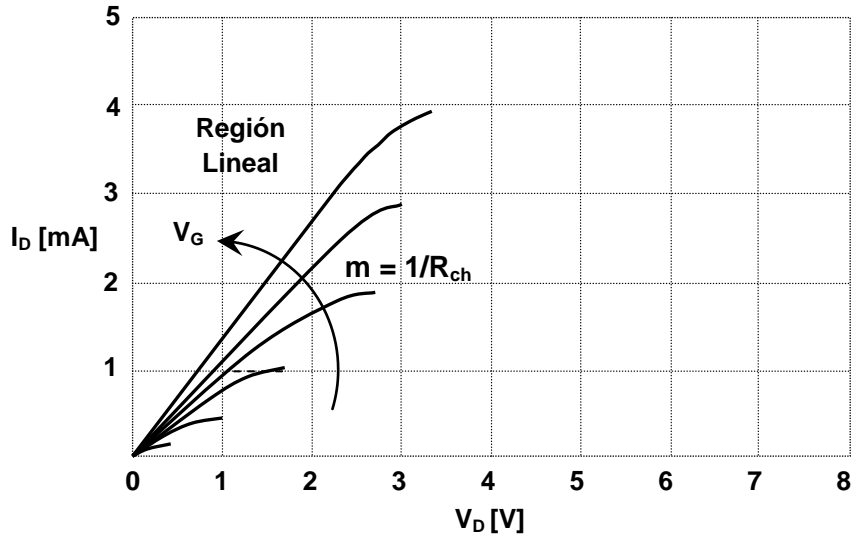


Figura 1.17. Características de corriente-voltaje en la región lineal para algunos valores de V_G .

1.3.2. Región de saturación.

Con el incremento del voltaje en el drenador V_D , el voltaje efectivo para mantener la inversión, junto al drenador, comenzará a aumentar. Como resultado, Q_n disminuirá y la resistencia del canal aumentará. Así, las características de I_D - V_D (Figura 1.17) comenzarán a doblarse hacia abajo, como se muestra en la Figura 1.18 [11].

A medida que V_D sigue aumentando, el voltaje efectivo necesario para mantener la inversión en la región cercana al drenador seguirá incrementándose, hasta que eventualmente será mayor que el voltaje aplicado a la compuerta. En ese momento, el canal se cerrará en el borde del drenador, como se muestra en la Figura 1.19. Esta condición se conoce como oclusión del canal o condición de Pinch-Off [12]. El voltaje de drenador necesario para alcanzar esta condición se conoce como voltaje de saturación V_{Dsat} [11].

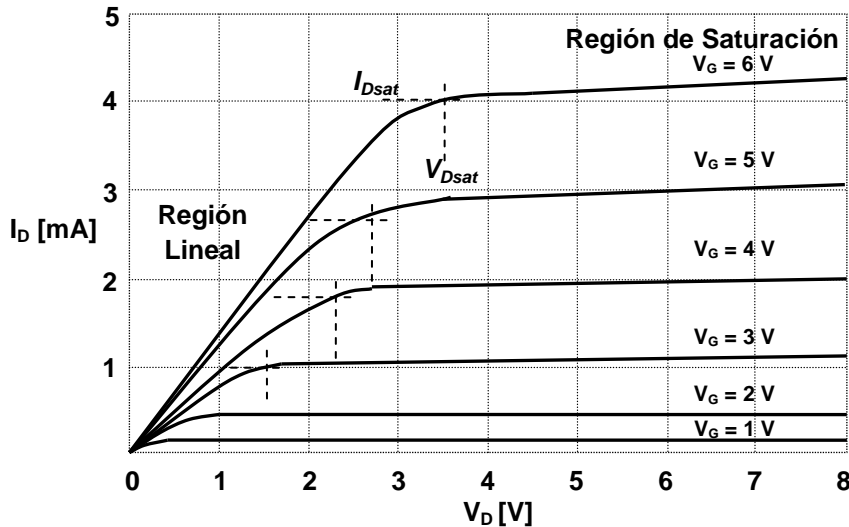


Figura 1.18. Características de corriente-voltaje de un MOSFET de canal n [11].

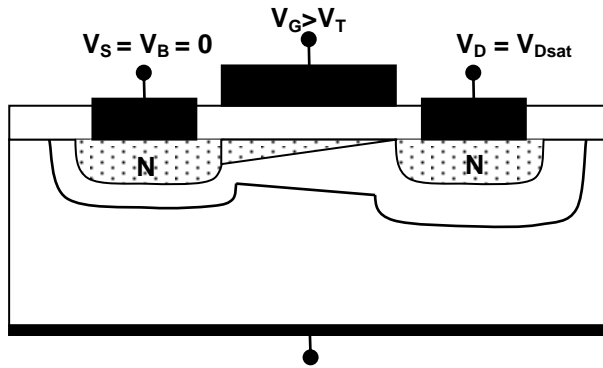


Figura 1.19. Esquema básico del MOSFET para $V_D = V_{Dsat}$; condición de Pinch-Off [12].

Una vez que el voltaje de drenador excede V_{Dsat} , el potencial efectivo de la región final del canal (junto al drenador) permanecerá constante, independientemente de cualquier incremento en el voltaje de drenador V_D , entonces se dice que se llegó a la saturación [11]. Sin embargo la longitud del canal se reducirá un poco [12], como se ve en la Figura 1.20.

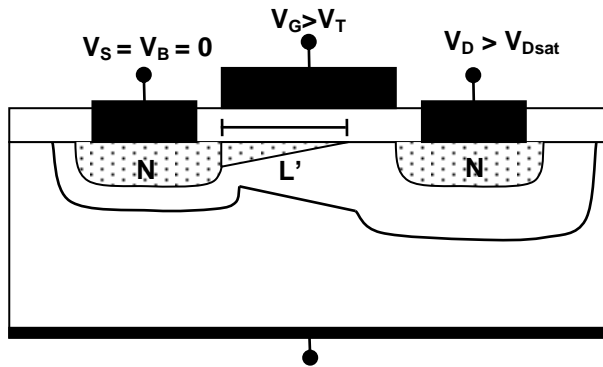


Figura 1.20. Esquema básico del MOSFET para $V_D > V_{Dsat}$; no se incrementa la corriente de drenador [12].

La corriente no cambiará de manera significativa con el incremento del voltaje de drenador, ya que depende del voltaje de compuerta y éste potencial permanece constante. Así, para voltajes de drenador mayores que V_{Dsat} la corriente no cambiará substancialmente y permanecerá en un valor I_{Dsat} [11].

Así, podemos distinguir dos regiones de operación del MOSFET. A voltajes pequeños de drenador, las características de corriente-voltaje son cuasi-lineales (región lineal). Mientras que a valores grandes la corriente se satura con el incremento de V_D (región de saturación) [11].

1.3.3. Modelo básico de corriente.

El canal del MOSFET es el que tiene mayor importancia para un análisis profundo, ya que sus características definen las propiedades de operación del MOSFET. Es por esto que se analizará con detalle la conductividad en el canal.

Se considerará un MOSFET trabajando en la región lineal [11], como se muestra en la Figura 1.21.

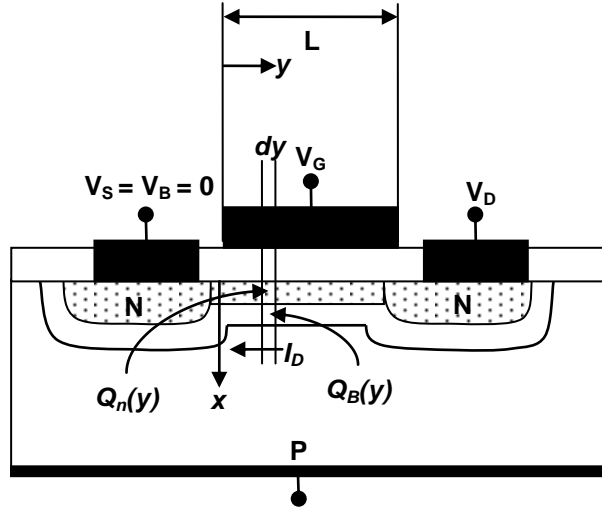


Figura 1.21. Sección elemental de canal de un MOSFET que contiene carga $Q_n(y)$ [11].

En esta Figura se muestra el sistema de coordenadas a utilizar. Se considerará que los parámetros solo varían perpendicularmente a la superficie. Además, se tomará en cuenta la aproximación de canal gradual, en donde la variación del campo eléctrico a lo largo del canal (eje y) es mucho menor que la magnitud del campo normal al canal (eje x) [12], es decir:

$$E_y \ll E_x \text{ y que } \frac{dE_y}{dy} \ll \frac{dE_x}{dx} \quad (1-44)$$

Con esto se considerará que los fenómenos se deben principalmente al campo E_x . La variación de potencial a lo largo del canal será $V(y)$ [12]. Entonces la densidad de corriente en el canal, considerando que es homogéneo y la sección transversal es constante, será:

$$J = \sigma E_y = qn\mu_n E_y = \frac{i}{A} = \frac{i}{t_i W} \quad (1-45)$$

La corriente [10] en el canal está dada por:

$$i = \iint J \cdot dS \quad (1-46)$$

Se deben considerar todos los elementos infinitesimales de corriente a lo largo del eje x para obtener la corriente total en el canal [12], por lo tanto:

$$di = \sigma E_y dS = \sigma \left[-\frac{dV(y)}{dy} \right] W dx \quad (1-47)$$

Integrando a lo largo de x , es decir desde $x=0$ hasta $x=t_i$:

$$I(y) = \int_0^{t_i} di = -W \frac{dV(y)}{dy} \int_0^{t_i} \sigma dx \quad (1-48)$$

La corriente total a lo largo del canal se define por la variación en el eje y [12]:

$$\int_0^L Idy = -W \int_0^L \frac{dV(y)}{dy} \int_0^{t_i} \sigma dx dy$$

Integrando:

$$I = -\frac{W}{L} \iint \sigma dx dV \quad (1-49)$$

Sustituyendo el valor de la conductividad se obtiene una ecuación general para la corriente:

$$I = -\frac{W}{L} \iint q \mu_n n dx dV \quad (1-50)$$

Para resolver esta ecuación se considerará que la movilidad no cambia con los potenciales aplicados al dispositivo [9], por lo que:

$$I = -\frac{W}{L} \mu_n \iint q n dx dV \quad (1-51)$$

La carga en la región del canal [9] es:

$$Q_n = \int_0^{t_i} q n dx \quad (1-52)$$

Sustituyendo Q_n en la ecuación (1-51):

$$I = -\frac{W}{L} \mu_n \iint Q_n dV \quad (1-53)$$

Ahora se debe conocer la carga Q_n en el canal, que como se ha mencionado anteriormente depende del voltaje de compuerta V_G . Para esto se puede tomar en cuenta la ecuación vista en la estructura MOS (1-31):

$$V_G = \frac{Q_s}{C_{ox}} + V_s$$

donde Q_s es la carga en el semiconductor, C_{ox} es la capacitancia asociada al óxido y $V_s = \phi$ es el potencial en el sustrato tipo P:

$$V_G = -\frac{Q_s}{C_{ox}} + \phi \quad (1-54)$$

Despejando Q_s :

$$Q_s = -C_{ox}(V_G - \phi) \quad (1-55)$$

La carga presente en el sustrato Q_s es carga debida a la región de inversión y a la RCE, así:

$$\begin{aligned} Q_s &= Q_n + Q_B \\ Q_n &= Q_s - Q_B \end{aligned} \quad (1-56)$$

Q_B es la carga por unidad de área debida a los iones de impurezas de la RCE:

$$Q_B = -q N_B x_{dmax} \quad (1-57)$$

$$x_{dmax} = \sqrt{\frac{2K_{si} \epsilon_0}{q N_B} 2\phi_F} \quad (1-58)$$

Entonces Q_n queda definida por:

$$Q_n = -C_{ox}(V_G - \varphi) + qN_B \sqrt{\frac{2K_{si}\epsilon_0}{qN_B} 2\varphi_F}$$

$$Q_n = -C_{ox} \left[V_G - \varphi - \frac{\sqrt{q2K_{si}\epsilon_0 N_B}}{C_o} \sqrt{2\varphi_F} \right] \quad (1-59)$$

Nombrando γ como el factor de cuerpo [9]:

$$\gamma = \frac{\sqrt{q2K_{si}\epsilon_0 N_B}}{C_{ox}} \quad (1-60)$$

Sustituyendo γ en la ecuación (1-59):

$$Q_n = -C_{ox}(V_G - \varphi - \gamma\sqrt{2\varphi_F}) \quad (1-61)$$

Para lograr inversión en cualquier punto de la superficie del sustrato [9]:

$$\varphi = 2\varphi_F + V(y) + V_{FB} \quad (1-62)$$

Sustituyendo esta última ecuación en (1-61):

$$Q_n = -C_{ox}(V_G - 2\varphi_F - V(y) - V_{FB} - \gamma\sqrt{2\varphi_F + V(y)}) \quad (1-63)$$

Ahora ya se tiene la carga del canal Q_n y se puede sustituir en la ecuación (1-53):

$$I = -\frac{W}{L} \mu_n \int_0^{V_D} [-C_{ox}(V_G - 2\varphi_F - V(y) - V_{FB} - \gamma\sqrt{2\varphi_F + V(y)})] dV$$

Integrando:

$$I = I_D = \frac{W}{L} \mu_n C_{ox} \left\{ (V_G - 2\varphi_F - V_{FB})V_D - \frac{V_D^2}{2} - \gamma \frac{2}{3} \left[(2\varphi_F + V_D)^{\frac{3}{2}} - (2\varphi_F)^{\frac{3}{2}} \right] \right\} \quad (1-64)$$

Esta ecuación de corriente es válida sólo en la región lineal del MOSFET, es decir, mientras el canal se encuentra en conducción [12].

Voltaje de umbral

Como anteriormente se vio en la estructura MOS, para llegar a una inversión fuerte, es necesario aplicar un voltaje mayor o igual al voltaje de umbral V_T , en el caso del MOSFET, este voltaje inducirá el canal que conecta al surtidor y el drenador. Así, en el MOSFET, este valor queda definido por el voltaje necesario para comenzar la formación del canal junto al surtidor, donde se puede tomar el potencial en ese punto como referencia, es decir, para $y=0$, $V(0)=0$ [12]. Esta condición es cuando la carga se hace cero, entonces de la ecuación (1-63):

$$Q_n|_{\text{surtidor}} = -C_{ox}(V_G - 2\varphi_F - V_{FB} - \gamma\sqrt{2\varphi_F}) = 0$$

$$V_G - 2\varphi_F - V_{FB} - \gamma\sqrt{2\varphi_F} = 0$$

Así, el voltaje de umbral en el MOSFET queda definido como:

$$V_T = V_G = 2\varphi_F + V_{FB} + \gamma\sqrt{2\varphi_F} \quad (1-65)$$

Voltaje de saturación

Para conocer el punto a partir del cuál comienza la saturación del transistor MOS (V_{Dsat}), es necesario hacer algunas consideraciones.

Con el incremento del voltaje de drenador, la región del canal cercana al drenador va desapareciendo (Figura 1.19). Cuando esto ocurre se puede considerar que la carga en ese punto es nula [12], es decir:

$$Q_n(y = L) = 0$$

Así:

$$Q_n = -C_o(V_G - 2\varphi_F - V_{FB} - V_D - \gamma\sqrt{2\varphi_F + V_D}) = 0$$

$$V_G - 2\varphi_F - V_{FB} - V_D - \gamma\sqrt{2\varphi_F + V_D} = 0$$

Entonces, el voltaje V_{Dsat} para $y=L$, es:

$$V_{Dsat} = V_G - 2\varphi_F - V_{FB} + \frac{\gamma^2}{2} \left[1 - \sqrt{1 + \frac{4}{\gamma^2} (V_G - V_{FB})} \right] \quad (1-66)$$

1.3.4. Modelo de primera aproximación.

Las expresiones presentadas en este Capítulo, para el MOSFET, son muy complejas para su aplicación, es por esto que se utilizarán algunas expresiones aproximadas para una evaluación más rápida. Estas expresiones son llamadas de primera aproximación. Surgen de considerar que la carga en la RCE Q_B es constante e igual a su valor en la superficie junto al surtidor cuando se forma el canal [12]. Las expresiones se presentan a continuación:

$$I_D = \frac{W}{L} \mu_{eff} C_{ox} \left[(V_G - V_T) V_D - \frac{V_D^2}{2} \right] \quad (1-67)$$

$$V_{Dsat} = V_G - V_T \quad (1-68)$$

$$I_{Dsat} = \frac{W}{L} \mu_{eff} C_{ox} \left[\frac{(V_G - V_T)^2}{2} \right] \quad (1-69)$$

La conductancia en la región lineal será:

$$g_d = \frac{W}{L} \mu_{eff} C_{ox} (V_G - V_T - V_D) \quad (1-70)$$

y la transconductancia:

$$g_m = \frac{W}{L} \mu_{eff} C_{ox} V_D \quad (1-71)$$

La transconductancia en la región de saturación será:

$$g_{msat} = \frac{W}{L} \mu_{eff} C_{ox} (V_G - V_T) \quad (1-72)$$

La movilidad μ_{eff} , utilizada en estas formulas de primera aproximación, caracteriza la variación de la movilidad con los potenciales aplicados [12]. Así, la movilidad queda expresada como:

$$\mu_{eff} = \frac{\mu_0}{[1+\theta(V_G-V_T)]\left[1+\frac{\mu_s}{Lv_{max}}V_D\right]} \quad (1-73)$$

donde:

$$\mu_s = \frac{\mu_0}{1+\theta(V_G-V_T)} \quad (1-74)$$

μ_0 es la movilidad superficial máxima, el parámetro θ incluye todos los factores que afectan la disminución de la movilidad con el campo transversal. Para considerar los efectos del campo longitudinal, se incluye la velocidad de saturación de los portadores v_{max} . Las consideraciones anteriores, son importantes al trabajar con canales muy cortos y campos intensos [12].

1.4. Efectos de Canal Corto (Short Channel Effects - SCE).

La reducción de las dimensiones del MOSFET ha sido fundamental a fin de hacerlos competitivos para aplicaciones analógicas de muy alta frecuencia. Sin embargo, con el escalado de los transistores MOS, comienzan a presentarse efectos indeseados en la operación del MOSFET, conocidos como efectos de canal corto. A medida que se reduce la longitud del canal, las regiones de carga espacial de las uniones del surtidor y drenador se acercarán, de forma que comenzarán a interferir con la región del canal quitándole el control de los fenómenos electrostáticos a la compuerta [9]. Son varias las consecuencias de los SCE, las más importantes son: Incremento de la corriente en régimen de saturación, incremento de la pendiente sub-umbral (S) e incremento de la corriente de fuga en condición de corte.

En régimen de saturación, un pequeño incremento del voltaje de drenador, producirá una pequeña reducción de la longitud del canal, es decir, la condición de Pinch-Off se desplazará hacia el interior del canal, tal como se mostró en la Figura 1.20. Para dispositivos de canal muy pequeño, este efecto se vuelve muy importante, de modo que la corriente de drenado se ve incrementada, apareciendo una pendiente en la zona de saturación [9].

Otro efecto producido por los SCE, se da en la región sub-umbral, la cual se tiene cuando el voltaje aplicado a la compuerta es menor que V_T . Bajo esta condición, la estructura MOS se encuentra en inversión débil y la corriente no depende del voltaje de drenador V_D , depende del voltaje de compuerta V_G [9], como se muestra en la Figura 1.22.

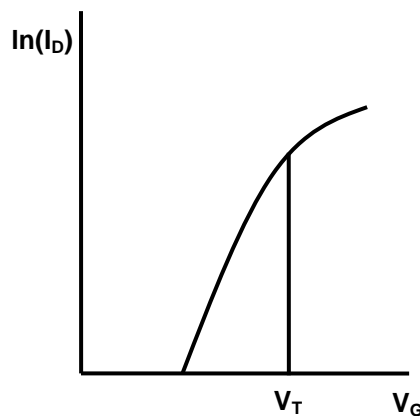


Figura 1.22. Región de la corriente sub-umbral.

Esta corriente se suele evaluar por un parámetro conocido como pendiente sub-umbral, que se refiere al voltaje necesario para que la corriente varíe una década [9], es decir:

$$S = \frac{\ln 10}{\frac{d(\ln I_D)}{dV_G}} \quad (1-75)$$

Sus unidades son [mV/Dec] y de manera ideal presenta un valor de 60 mV/Dec. Como resultado de los SCE, la pendiente de la grafica de la Figura 1.22 disminuye por lo que la corriente de fuga se incrementa [9].

Finalmente, los SCE provocan una reducción del V_T , lo cual produce incrementos de la corriente de fuga del dispositivo bajo condición de corte.

Los conceptos presentados en este Capítulo dan una base para entender el funcionamiento y operación del MOSFET y así poder modelar su comportamiento y utilizarlo en diversas aplicaciones, ya que es uno de los elementos más utilizados en la electrónica analógica y digital, y por lo tanto en las muchas áreas de telecomunicaciones, debido a que prácticamente todos los circuitos integrados de uso comercial se basan en el MOSFET.

El estudio del MOSFET siempre ha tomado gran importancia. Desde su aparición, se ha desarrollado de una manera acelerada, donde cada vez se buscan mejorar sus parámetros tecnológicos a fin de incrementar su frecuencia de operación, lo cual ha permitido el desarrollo de aplicaciones analógicas a frecuencias cada vez mayores, de modo que actualmente se prevé su uso para aplicaciones de microondas y posiblemente ondas milimétricas. Por esto se vuelve fundamental su estudio de cara a las tecnologías actuales en sistemas de telecomunicaciones.