

TEMA 5: FÍSICA DE LOS DISPOSITIVOS UNIPOLARES O DE EFECTO DE CAMPO.

Introducción.

5.1. El transistor JFET.

Regiones de funcionamiento:

- Zona lineal
- Desviación del comportamiento lineal.
- Estrangulamiento del canal
- Zona de saturación

5.1.1. Características corriente-tensión.

Conductancia del canal g_D .
Transconductancia del canal g_m .
Tensión de ruptura.

5.2. El transistor MOSFET: Características básicas.

5.2.1. Regiones lineal y de saturación.

5.2.2. La región subumbral.

5.2.3. Tipos de transistores MOSFET.

5.2.4. Integración de dispositivos y efectos de canal corto.

5.2.4.1 Efectos de canal corto.

5.2.4.2. Miniaturización de dispositivos.

TEMA 5: FÍSICA DE LOS DISPOSITIVOS UNIPOLARES O DE EFECTO DE CAMPO.

Introducción.

Se denominan dispositivos unipolares a aquellos dispositivos semiconductores que basan su mecanismo de conducción de forma casi exclusiva en un solo tipo de portador de carga. Podemos considerar en principio cinco tipos de dispositivos unipolares:

- a) Contacto metal-semiconductor. Es la unión de un metal con un semiconductor (barrera Schottky). Equivalente a una unión p-n abrupta a un lado. Además, para el caso de semiconductores fuertemente dopados, este tipo de unión constituye la forma más importante de realizar un contacto óhmico.
- b) El transistor de efecto de campo (JFET). Se trata de una resistencia controlada por tensión. Basa su funcionamiento en una unión p-n inversamente polarizada que se encarga de controlar la resistencia existente entre dos contactos óhmicos, o lo que es lo mismo, el flujo de corriente a su través.
- c) El transistor MESFET (Metal Semiconductor Field Effect Transistor). Idéntico al JFET pero emplea un contacto metal-semiconductor de tipo rectificador en lugar de un unión p-n.
- d) El diodo MOS (Metal-Oxide Semiconductor Diode). Componente muy empleado en el estudio de las superficies semiconductoras, de gran aplicación en los dispositivos CCD.
- e) El transistor MOSFET (Metal-Oxide Semiconductor Field-Effect Transistor). Constituye un diodo MOS con dos uniones p-n adyacentes al diodo MOS. De análogas propiedades eléctricas que los transistores JFET y MESFET es de gran utilidad en la fabricación de microprocesadores, memorias de semiconductor, etc.

Con respecto a los dispositivos bipolares, los unipolares presentan características bien concretas y que en muchos casos son ventajas respecto a los bipolares:

- De fabricación más sencilla. Ocupan menos espacio y, por lo tanto, presentan una mayor capacidad de integración.
- Menor consumo que el bipolar (del orden de nanovatios para el MOS frente a los milivatios de los BJT).
- Más económico.
- Elevada impedancia de entrada (del orden de $10^{10}\Omega$).
- Es el elemento óptimo para la fabricación de memorias.
- Presenta menos ruido que el bipolar.

Sin embargo, por cuestiones de tiempo se tratarán únicamente los dispositivos unipolares transistor JFET y transistor MOSFET. La unión metal-semiconductor ya fue tratada en el tema correspondiente a la unión p-n.

5.1. El transistor JFET.

El transistor JFET fue analizado por primera vez en 1952. Su funcionamiento se basa en la modulación de una resistencia a través de la polarización inversa de una unión p-n. También se trata de un dispositivo unipolar.

Fig. 1a muestra una vista en perspectiva de un transistor JFET. Está formado por un canal conductor que tiene emplazados en sus extremos dos contactos óhmicos llamados fuente (source) y drenador (drain). Un contacto óhmico es un contacto metal-semiconductor que presenta una resistencia prácticamente nula e independiente de la tensión aplicada. Al aplicar una tensión positiva entre drenador y fuente aparece un flujo de electrones desde la fuente hacia el drenador.

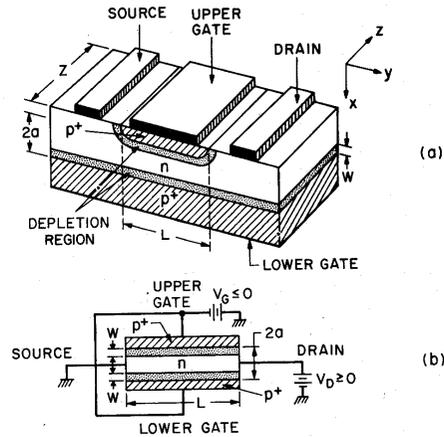


Fig. 1: (a) Vista en perspectiva de un JFET. (b) Sección transversal de la zona central del JFET. La fuente es llevada a tierra y las tensiones de puerta y drenador son las adecuadas para condiciones normales de funcionamiento.

De esta forma la fuente actúa como origen de los portadores y el drenador como el sumidero. Queda el tercer electrodo o contacto llamado puerta, el cual forma una unión rectificadora con el canal.

Para simplificar el estudio se considerará la estructura simétrica de Fig. 1b. Los contactos de la puerta están unidos entre sí, la fuente está puesta a tierra, la puerta a una tensión V_G y el drenador a una tensión V_D. Si el canal es de tipo N, el transistor se denomina JFET de canal N. Este es preferible a un JFET de canal P (en que el canal es de tipo P) debido a la mayor movilidad de sus portadores.

Fig. 2 muestra el símbolo electrónico del JFET:

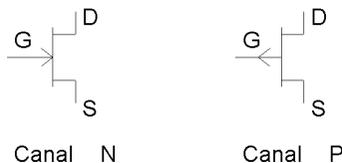


Fig. 2: Símbolo electrónico del JFET.

La resistencia del canal viene dada por:

$$R = \rho \cdot \frac{L}{A} = \frac{L}{q \cdot \mu_n \cdot N_D \cdot A} = \frac{L}{2 \cdot q \cdot \mu_n \cdot N_D \cdot Z \cdot (a - W)}$$

donde:

N_D ≡ Concentración de impurezas dadoras.

A ≡ Sección transversal que atraviesa la corriente = 2·Z·(a-W)

W ≡ Anchura de la región espacial de carga de la uniones p⁺-n superior e inferior.

L ≡ Longitud del canal.

Z ≡ Anchura del canal.

2·a ≡ Profundidad del canal.

Fig. 3 representa los diferentes casos que pueden considerarse desde el punto de vista de la conducción. Cuando no hay tensión aplicada en la puerta y la tensión de drenador es pequeña aunque ligeramente positiva (Fig. 3a) sólo fluye una pequeña cantidad de corriente I_D a través del canal. El valor de esta corriente viene dado por el cociente V_D/R donde R se obtiene de la expresión anterior. Para estas pequeñas tensiones de polarización V_D, R es prácticamente constante y, en consecuencia, la corriente varía linealmente con la tensión de drenador (**zona lineal**).

Por otra parte se observa un gradual aumento de la anchura W de la región espacial de carga al moverse desde la fuente hacia el drenador. Ello es consecuencia del aumento progresivo de la polarización inversa de las uniones p⁺-n superior e inferior a medida que nos desplazamos desde la fuente (que está a cero voltios) hasta el drenador (con V_D ligeramente positiva). A medida que la polarización V_D aumenta, la anchura W de la región espacial de carga también aumenta y, en consecuencia, la resistencia del canal (pues la sección efectiva de éste disminuye). Como resultado final, la corriente presenta una pendiente menor (**desviación del comportamiento lineal**).

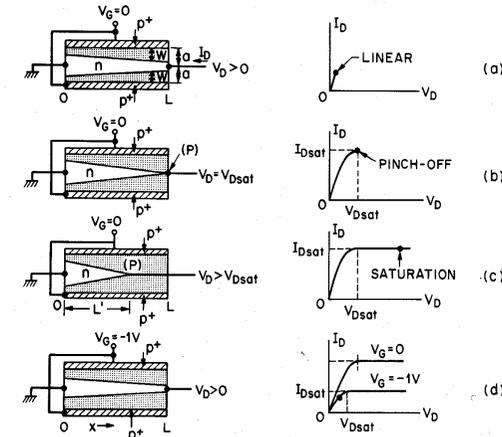


Fig. 3: Variación de la anchura de la región espacial de carga y características de salida de un JFET bajo condiciones de polarización. (a) V_G = 0 y V_D pequeño. (b) V_G = 0 y punto de estrangulamiento del canal. (c) V_G = 0 y V_D > V_{Dsat}. (d) V_G = -1V y V_D pequeño.

Puede darse el caso de que las regiones espaciales de carga superior e inferior se toquen (punto P), Fig. 3b. Ello ocurre cuando $W = a$ en la zona de drenador. Para una unión p⁺-n abrupta el valor de la tensión de drenador para el cual ocurre el **estrangulamiento del canal** se denomina tensión de saturación $V_{D,sat}$ y viene dada por la expresión:

$$V_{D,sat} = \frac{q \cdot N_D \cdot a^2}{2 \cdot \epsilon_s} - V_{bi} \text{ para } V_G = 0$$

donde V_{bi} es el potencial de contacto de la unión con la puerta. Para este valor de la tensión V_D fluye un valor alto de corriente llamada corriente de saturación I_{Dsat} desde el drenador hacia la fuente a través de la región espacial de carga. Esta situación es análoga a la inyección de portadores en una región espacial de carga inversamente polarizada, tal y como ocurría en la unión colector-base de un transistor bipolar.

Si la tensión de drenador sigue aumentando, el punto P de contacto se desplaza hacia la zona de la fuente (Fig. 3c), siendo la tensión en este punto V_{Dsat} . En consecuencia, el número de electrones que fluyendo de la fuente atraviesan el canal y llegan al punto P queda inalterado con respecto a la situación b) pues el punto P no ha variado su potencial.

En consecuencia, para tensiones V_D mayores que V_{Dsat} , la corriente neta no sufre variación apreciable con respecto al valor I_{Dsat} y es independiente de V_D (**zona de saturación**).

Fig. 3d representa el caso en que la tensión aplicada a la puerta sea negativa. Frente a esta situación, aumenta la polarización inversa de la unión p⁺-n y, por tanto, la anchura de la región espacial de carga, este efecto trae consigo a su vez que la resistencia efectiva del canal aumente disminuyendo el flujo de corriente. Para polarizaciones V_D iguales al caso $V_G = 0$ ahora I_D será menor.

Para la situación $V_G = (-)1V$, el flujo de corriente es menor que para $V_G = 0V$ y también se consigue un cierto valor de V_D para el cual las dos regiones espaciales de carga se unen, alcanzándose de nuevo la saturación pero ahora en un nivel inferior de corriente. El valor de V_{Dsat} con polarización negativa en V_G viene dado por:

$$V_{Dsat} = \frac{q \cdot N_D \cdot a^2}{2 \cdot \epsilon_s} - V_{bi} - V_G \text{ para } V_G < 0$$

donde V_G está tomado en valor absoluto para un JFET de canal N.

5.1.1. Características corriente-tensión.

Fig. 4a representa el caso del dispositivo JFET cuando todavía no se ha logrado el estrangulamiento del canal y en Fig. 4b aparece la evolución gradual de la tensión a través del canal.

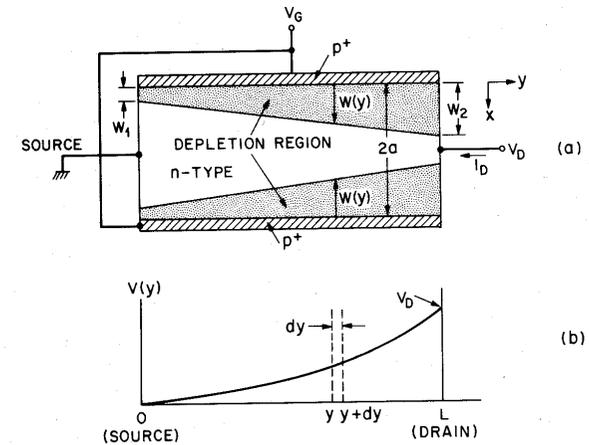


Fig. 4: (a) Vista de la zona del canal. (b) Caída de V_D a lo largo del canal.

La caída de tensión dV en un elemento dy del canal vendrá dada por la expresión:

$$dV = I_D \cdot dR = \frac{I_D \cdot dy}{2 \cdot q \cdot \mu_n \cdot N_D \cdot Z \cdot [a - W(y)]}$$

donde se ha empleado la expresión anterior para R .

La anchura de la región espacial de carga correspondiente a la unión p⁺-n inversamente polarizada $W(y)$ varía en función de la distancia y del canal de la forma:

$$W(y) = \sqrt{\frac{2 \cdot \epsilon_s \cdot [V(y) + V_G + V_{bi}]}{q \cdot N_D}}$$

Por otro lado, de la expresión anterior:

$$I_D \cdot dy = 2 \cdot q \cdot \mu_n \cdot N_D \cdot Z \cdot [a - W(y)] \cdot dV$$

El valor de dV puede ser obtenido a partir de la expresión anterior para $W(y)$, elevando al cuadrado y diferenciando:

$$W^2(y) = \frac{2 \cdot \epsilon_s \cdot [V(y) + V_G + V_{bi}]}{q \cdot N_D} \Rightarrow 2 \cdot W \cdot dW = \frac{2 \cdot \epsilon_s}{q \cdot N_D} \cdot dV \Rightarrow dV = \frac{q \cdot N_D}{\epsilon_s} \cdot W \cdot dW$$

De esta manera, sustituyendo dV en la igualdad para $I_D \cdot dy$:

$$I_D \cdot dy = 2 \cdot q \cdot \mu_n \cdot N_D \cdot Z \cdot [a - W(y)] \cdot \frac{q \cdot N_D}{\epsilon_s} \cdot W \cdot dW$$

Integrando desde $y = 0$ hasta $y = L$:

$$I_D \cdot \int_0^L dy = I_D \cdot L = \int_{W_1}^{W_2} 2 \cdot q \cdot \mu_n \cdot N_D \cdot Z \cdot [a - W] \cdot \frac{q \cdot N_D}{\epsilon_s} \cdot W \cdot dW$$

con:

$W_1 \equiv$ Anchura de la región espacial de carga en $y = 0$

$W_2 \equiv$ Anchura de la región espacial de carga en $y = L$

Luego:

$$\begin{aligned} I_D &= \frac{1}{L} \cdot \int_{W_1}^{W_2} 2 \cdot q \cdot \mu_n \cdot N_D \cdot Z \cdot [a - W] \cdot \frac{q \cdot N_D}{\epsilon_s} \cdot W \cdot dW = \\ &= \frac{Z \cdot \mu_n \cdot q^2 \cdot N_D^2}{\epsilon_s \cdot L} \cdot \left[a \cdot (W_2^2 - W_1^2) - \frac{2}{3} \cdot (W_2^3 - W_1^3) \right] = \\ &= I_p \cdot \left[\frac{V_D}{V_p} - \frac{2}{3} \cdot \left(\frac{V_D + V_G + V_{bi}}{V_p} \right)^{3/2} + \frac{2}{3} \cdot \left(\frac{V_G + V_{bi}}{V_p} \right)^{3/2} \right] \end{aligned}$$

donde:

$$I_p \equiv \frac{Z \cdot \mu_n \cdot q^2 \cdot N_D^2 \cdot a^3}{\epsilon_s \cdot L}$$

y

$$V_p \equiv \frac{q \cdot N_D \cdot a^2}{2 \cdot \epsilon_s}$$

La tensión V_p es la tensión a la cual se produce el estrangulamiento del canal, es decir, la tensión $V_D + V_G + V_{bi}$ a la cual $W_2 = a$.

En Fig. 5 se representa la característica corriente-tensión para un JFET con tensión de estrangulamiento de 3.2V. Se observan dos regiones. La primera de ellas comprende los valores de tensión de drenador $0 \leq V_D \leq V_{Dsat}$ en la cual la resistencia que ofrece el canal es independiente de la tensión de drenador y la característica corriente-tensión es de tipo óhmico o lineal, es la llamada región lineal. El cálculo de la corriente I en términos de la tensión V viene dado en términos de la expresión de I_D anterior.

A continuación viene una región en la cual para valores de $V_D > V_{Dsat}$ la corriente de drenador es prácticamente independiente de la tensión de drenador aplicada, es la región de saturación para la cual la corriente de drenador es igual a I_{Dsat} . A continuación se tratarán ambas regiones.

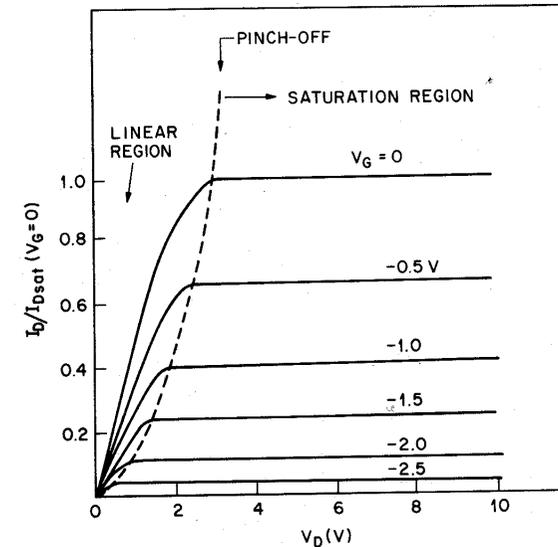


Fig. 5: Característica corriente-tensión ideal normalizada con $V_p=3.2V$.

La región lineal: En esta región se satisface la desigualdad $V_D \ll V_G + V_{bi}$ y se puede expandir el valor de I_D para aproximarlo mediante la expresión:

$$I_D \equiv \frac{I_p}{V_p} \cdot \left[1 - \sqrt{\frac{V_G + V_{bi}}{V_p}} \right] \cdot V_D$$

Se define el parámetro **conductancia del canal g_D** (también llamada conductancia de drenador) como:

$$g_D \equiv \frac{\partial I_D}{\partial V_D} \Big|_{V_G=cte} = \frac{I_p}{V_p} \cdot \left[1 - \sqrt{\frac{V_G + V_{bi}}{V_p}} \right]$$

En esta región puede suponerse que el canal es una resistencia controlada por V_G y de valor, r_d con $r_d \equiv \frac{1}{g_D}$.

Fig. 6 muestra la conductancia normalizada del canal $g_D \cdot \frac{V_p}{I_p}$ en la región lineal en el caso teórico (línea continua) y en la práctica, en un dispositivo que presenta resistencias serie (línea discontinua). Se observa cómo a medida que aumenta la polarización de puerta (V_G) la conductancia disminuye hasta alcanzarse el punto de estrangulamiento en el cual se anula.

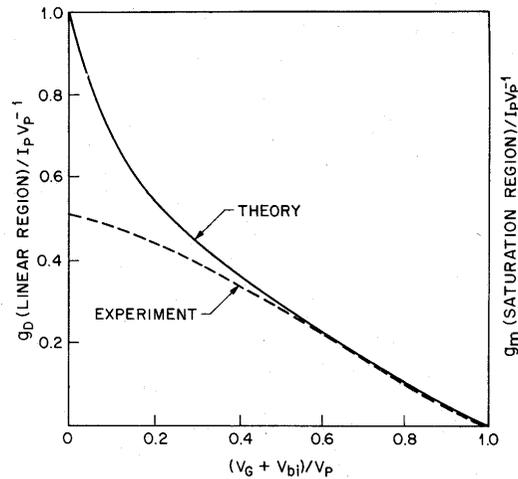


Fig. 6: Variación de los valores normalizados de la conductancia y la transconductancia del canal frente a la tensión de puerta normalizada. La línea continua corresponde al caso ideal, la discontinua a lo que sucede en la práctica debido a la presencia de las resistencias serie.

Otro parámetro importante del dispositivo JFET operando en la zona lineal es la transconductancia g_m , la cual se define como:

$$g_m \equiv \left. \frac{\partial I_D}{\partial V_G} \right|_{V_D = \text{cte}}$$

En la región lineal de operación, g_m se obtiene a partir de la aproximación de I_D anterior:

$$g_m = \frac{I_p}{2 \cdot V_p} \cdot \sqrt{\frac{V_p}{V_G + V_{bi}}} \cdot V_D$$

La región de saturación: En la expresión total de la corriente I_D se puede obtener la corriente de saturación seleccionando la tensión en el valor $V_p = V_{Dsat} + V_G + V_{bi}$:

$$I_{Dsat} = I_p \cdot \left[\frac{1}{3} - \left(\frac{V_G + V_{bi}}{V_p} \right) + \frac{2}{3} \cdot \left(\frac{V_G + V_{bi}}{V_p} \right)^{3/2} \right]$$

y la tensión de saturación viene dada por:

$$V_{Dsat} = V_p - V_G - V_{bi}$$

Respecto a la conductancia del canal ésta vale cero en la región de saturación pues una característica ideal da lugar a la no dependencia de I_{Dsat} respecto a V_D . Por otro lado, la transconductancia g_m se obtiene de la expresión anterior:

$$g_m = \frac{I_p}{V_p} \cdot \left[1 - \sqrt{\frac{V_G + V_{bi}}{V_p}} \right] = \frac{2 \cdot Z \cdot \mu_n \cdot q \cdot N_D \cdot a}{L} \cdot \left[1 - \sqrt{\frac{V_G + V_{bi}}{V_p}} \right]$$

Este resultado nos dice que la transconductancia en la región de saturación coincide con la conductancia en la región lineal por lo que Fig. 6 también es válida para la saturación.

Hasta ahora sólo se ha considerado la resistencia del canal, la cual puede ser modulada en función de la tensión de puerta. En la práctica, existen unas resistencias serie cercanas a los terminales de fuente y drenador como puede observarse en Fig. 7.

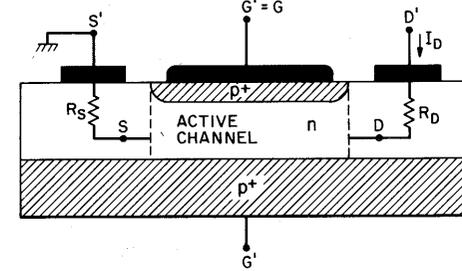


Fig. 7: JFET con las resistencias de fuente y drenador debidas a la conductividad finita en las regiones semiconductoras entre el final del canal activo y los terminales de fuente y drenador.

Estas resistencias provocan caídas de tensión I·R entre los contactos de fuente o drenador y el canal, teniendo efecto sobre los valores de la conductancia y la transconductancia del canal. Las caídas de tensión entre terminales de un JFET considerando las resistencias serie vienen dadas por:

$$V_{D'S'} = V_{DS} + I_D \cdot (R_S + R_D)$$

$$V_{G'S'} = V_{GS} + I_D \cdot R_S$$

La presencia de dichas resistencias serie provoca una reducción con respecto a los valores teóricos tanto de la conductancia como de la transconductancia del canal (Fig. 6, línea discontinua).

Tensión de ruptura: Fig. 8 es una extensión de la característica corriente-tensión de Fig. 5. A medida que la tensión de drenador aumenta hay un momento en el cual el diodo formado entre la puerta y el canal entra en su región de avalancha y, en consecuencia, la corriente de drenador aumenta muy rápidamente. La ruptura se origina en la parte del canal más próxima al drenador ya que es aquí donde la tensión inversa aplicada es mayor:

$$V_B(\text{breakdown voltage}) = V_D + |V_G|$$

En Fig. 8, $V_B = 12V$ para $V_G = 0V$. Si $|V_G| = 1$, V_B permanece todavía en 12V aunque la tensión de drenador en condiciones de ruptura es $V_B - |V_G| = 11V$.

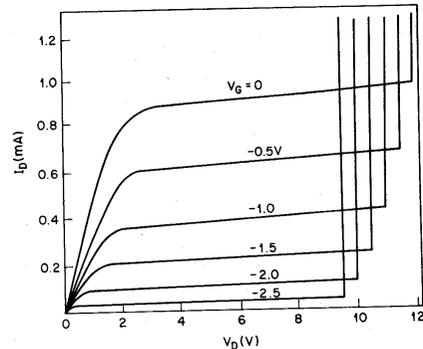


Fig. 8: Característica corriente-tensión de un JFET real.

5.2. El transistor MOSFET: Características básicas.

El transistor MOSFET (acrónimo de Metal-Oxide-Semiconductor Field-Effect Transistor) es el dispositivo semiconductor más importante desde el punto de vista de la fabricación a gran escala de dispositivos y circuitos integrados (memorias, microprocesadores). El primer transistor MOSFET fue fabricado en 1960 a título experimental. Se trata de un dispositivo unipolar de características corriente-tensión análogas a las del JFET.

5.2.1. Regiones lineal y de saturación.

Fig. 9 representa una perspectiva del transistor MOSFET. Se trata de un dispositivo de cuatro terminales. En primer lugar aparece un sustrato semiconductor de tipo P, en el cual hay dispuestas dos regiones de tipo n⁺. Este dispositivo MOSFET se dice que es de canal N. El de canal P está formado por un sustrato de tipo N e invirtiendo las polarizaciones existentes.

Las regiones de tipo n⁺ son las correspondientes a la fuente y el drenador. El contacto metálico con el óxido es la puerta. El óxido actúa de aislante entre el metal y el sustrato semiconductor.

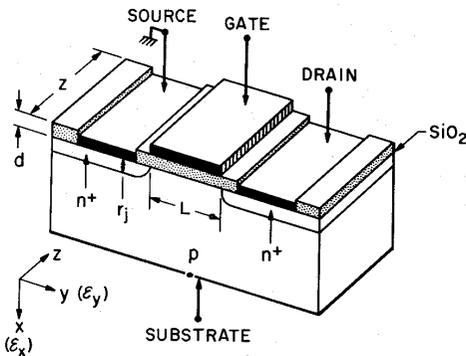


Fig. 9: Vista en perspectiva de un MOSFET.

Desde el punto de vista geométrico, los parámetros que caracterizan al dispositivo son la longitud del canal L (distancia de las dos uniones n⁺-p); la anchura del canal Z, espesor del óxido dieléctrico, d, el espesor r_j de las regiones n⁺ y el dopado del sustrato N_A.

Las propiedades conductoras del dispositivo MOSFET están basadas fundamentalmente en el comportamiento del llamado canal óhmico o región situada entre las dos zonas de tipo n⁺. La disposición formada por la puerta, el aislante SiO₂ y el sustrato semiconductor de tipo P, conforma un condensador de placas paralelas con dieléctrico SiO₂ el cual bajo determinadas condiciones de polarización es susceptible de crear cargas inducidas.

Supóngase una tensión en la puerta positiva con respecto a la fuente, al estar el contacto metálico a un potencial positivo dicha tensión producirá un campo eléctrico en el dieléctrico el cual inducirá por efecto condensador cargas de sentido opuesto (cargas negativas) al otro lado del dieléctrico, es decir, en el sustrato de tipo P. Sin embargo, al ser muy deficitario en electrones, las cargas negativas van a ser aportadas o van a proceder de las regiones n⁺ laterales correspondientes a la fuente y el drenador, los cuales son excedentarios en portadores de carga negativa (Fig. 10).

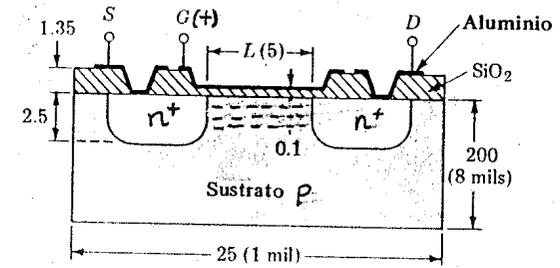


Fig. 10: Proceso de formación del canal.

La región situada entre las zonas n⁺ está poblada ahora de cargas, en este caso negativas y confiere propiedades de conductividad a esta región. Dicha zona se suele llamar canal, el cual aumenta su conductividad eléctrica a medida que crece la polarización positiva puerta-fuente. Se ha de decir que la aparición de las cargas en el canal se consigue a partir de haber superado una tensión umbral de puerta V_{GS(th)}. La conductividad del canal puede modularse variando la polarización de puerta.

Para comprender el modo de conducción en un transistor MOSFET, considérese que se aplica una tensión positiva a la puerta respecto a la fuente, tal que se crea una región de cargas negativas (canal), Fig. 11a. Si ahora se aplica a la región del drenador una pequeña tensión positiva se producirá un flujo de electrones desde la fuente hacia el drenador (o bien una corriente en sentido contrario) que tiene paso libre a través del canal creado. En esta situación, el canal actúa como una resistencia, y la corriente de drenador I_D es proporcional a la tensión de drenador V_D aplicada (Fig. 11a). Véase la línea recta en la gráfica I_D-V_D.

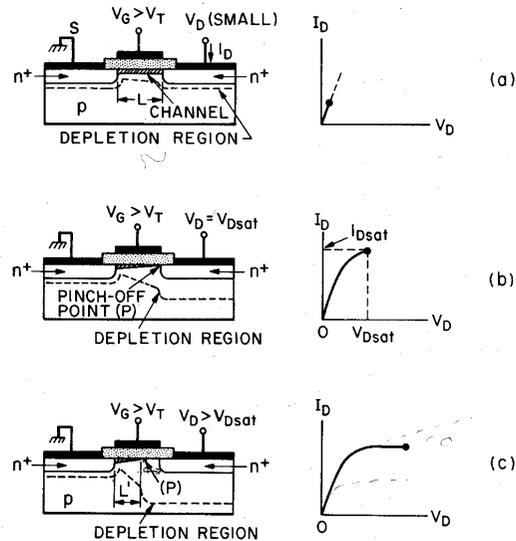


Fig. 11: Funcionamiento del MOSFET y característica I-V. (a) Baja tensión de drenador. (b) Inicio de la saturación. El punto P indica el punto de estrangulamiento del canal. (c) Más allá de la saturación.

Sin embargo, a medida que la tensión de drenador va aumentando, la región del canal más próxima al drenador pierde progresivamente su carácter negativo eléctrico, y en consecuencia disminuye la anchura del canal en esta zona próxima al drenador con respecto a la que está al lado de la fuente. Se pierde la relación lineal constante entre I_D y V_D . Si se sigue aumentando la tensión de drenador, se llega a un valor de ésta V_{Dsat} , en el cual la anchura del canal en la zona próxima al drenador se ha anulado, no hay canal, es el llamado punto de estrangulamiento del canal, Fig. 11b.

Más allá de este valor de tensión ($V_D > V_{Dsat}$), la corriente que fluye por el canal permanece prácticamente invariable debido a que la tensión en el punto P permanece prácticamente en el valor V_{Dsat} . El número de portadores que llega al punto P procedentes de la fuente permanece invariable. El cambio más significativo es la disminución de la longitud del canal a un valor L' inferior a L ($L' < L$), Fig. 11c.

La región de la característica I_D - V_D del transistor MOSFET en la cual la corriente es proporcional a la tensión se denomina región lineal, a partir del valor de tensión en el drenador V_{Dsat} , la región en la característica se denomina región de saturación.

A continuación se van a obtener de forma analítica las características estáticas corriente-tensión en un transistor MOSFET. Para ello se va a partir de un modelo idealizado, el cual ayudará a comprender mejor su funcionamiento. En este sentido deben hacerse una serie de suposiciones encaminadas a simplificar el modelo:

- 1°.- Sólo se considerará la corriente de arrastre.
- 2°.- La movilidad de los portadores en la región óhmica es constante.
- 3°.- El dopado en el canal es constante.

4°.- El campo eléctrico transversal (dirección de arriba a abajo) es mucho mayor que el campo longitudinal (paralelo al flujo de corriente). Esta es la llamada aproximación del canal gradual.

En Fig. 12a se representa un transistor MOSFET operando en la zona lineal. Bajo las hipótesis anteriores, puede decirse que la carga total inducida en el semiconductor por unidad de área, Q_s , a una distancia y de la fuente, viene dada por la expresión:

$$Q_s(y) = -[V_G - \psi_s(y)] \cdot C_o$$

donde $\psi_s(y)$ es el potencial de superficie en el punto y , y $C_o \equiv \epsilon_{ox}/d$ es la capacidad por unidad de área. ϵ_{ox} es la constante dieléctrica del óxido. Asimismo, la carga en el canal viene dada por:

$$Q_n(y) = Q_s(y) - Q_{sc}(y) = -[V_G - \psi_s(y)] \cdot C_o - Q_{sc}(y)$$

siendo Q_{sc} la carga en la superficie de la región de distribución espacial de carga, la cual puede obtenerse a partir de la expresión:

$$Q_{sc}(y) = -q \cdot N_A \cdot W_m \equiv -\sqrt{2 \cdot \epsilon_s \cdot q \cdot N_A} \cdot [V(y) + 2 \cdot \psi_B]$$

donde $V(y)$ aparece en Fig. 12c y es la polarización inversa entre el punto y y el electrodo de fuente (a potencial cero), ψ_B es la diferencia de potencial entre el nivel de Fermi E_F y el nivel de Fermi intrínseco E_i . De esta forma:

$$Q_n(y) \equiv -[V_G - V(y) - 2 \cdot \psi_B] \cdot C_o + \sqrt{2 \cdot \epsilon_s \cdot q \cdot N_A} \cdot [V(y) + 2 \cdot \psi_B]$$

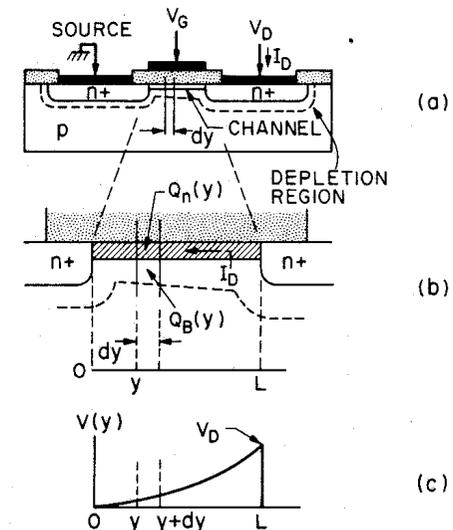


Fig. 12: (a) MOSFET operando en la zona lineal. (b) Ampliación de la zona del canal. (c) Caída de V_D a lo largo del canal.

La conductividad del canal en la posición y puede ser aproximada por:

$$\sigma(x) = q \cdot n(x) \cdot \mu_n(x)$$

bajo la hipótesis de una movilidad constante, la conductancia del canal viene dada por:

$$g = \frac{Z}{L} \cdot \int_0^{x_i} \sigma(x) \cdot dx = \frac{Z \cdot \mu_n}{L} \cdot \int_0^{x_i} q \cdot n(x) \cdot dx$$

La anterior integral corresponde a la carga total por unidad de área en el canal, es decir, $|Q_n|$, con lo cual:

$$g = \frac{Z \cdot \mu_n}{L} \cdot |Q_n|$$

La resistencia del canal en una sección dy (Fig. 12b) es:

$$dR = \frac{dy}{g \cdot L} = \frac{dy}{Z \cdot \mu_n \cdot |Q_n(y)|}$$

y la caída de tensión en este elemento dy es:

$$dV = I_D \cdot dR = \frac{I_D \cdot dy}{Z \cdot \mu_n \cdot |Q_n(y)|}$$

I_D es la corriente de drenador, la cual es independiente de y. Si ahora se sustituye la expresión para $Q_n(y)$ en esta última y se integra entre $y=0$ (con $V=0$) e $y=L$ (con $V=V_D$) se obtiene:

$$I_D \cong \frac{Z}{L} \cdot \mu_n \cdot C_o \cdot \left[\left(V_G - 2 \cdot \psi_B - \frac{V_D}{2} \right) \cdot V_D - \frac{2}{3} \cdot \frac{\sqrt{2 \cdot \epsilon_s \cdot q \cdot N_A}}{C_o} \cdot \left((V_D + 2 \cdot \psi_B)^{3/2} - (2 \cdot \psi_B)^{3/2} \right) \right]$$

Fig. 13 recoge la representación gráfica de la anterior ecuación para un cierto valor dado V_G . Corresponde a la característica corriente-tensión de un transistor MOSFET ideal. Puede observarse en ella, la región lineal en la cual la corriente I_D aumenta de forma lineal con la tensión de drenador.

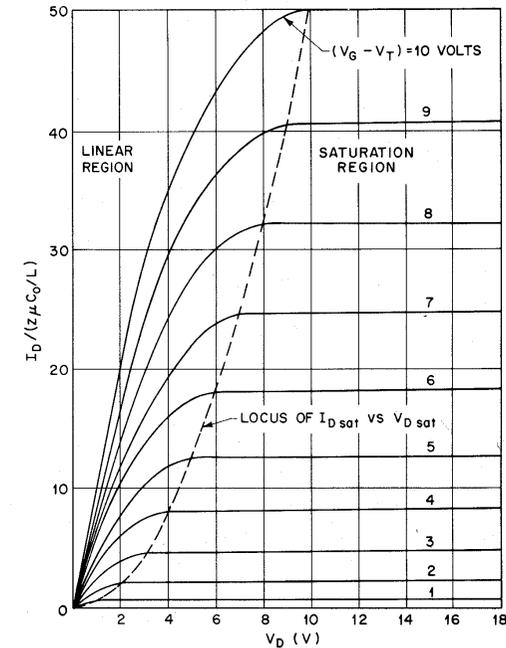


Fig. 13: Característica idealizada de drenador de un MOSFET. Para $V_D \geq V_{Dsat}$, la corriente de drenador permanece constante.

Esto ocurre hasta un cierto valor de la tensión de drenador, llamado V_{Dsat} en el cual se alcanza el valor máximo de la corriente, llamado I_{Dsat} . A partir de este valor de la tensión de drenador, la corriente ya no aumenta linealmente, por el contrario permanece invariable con la tensión e igual a I_{Dsat} , es la llamada región de saturación. La línea de puntos indica el lugar geométrico de los valores (I_{Dsat}, V_{Dsat}).

A continuación se va a pasar a discutir las dos regiones características. En primer lugar cabe considerar la región lineal. La ecuación general para I_D puede aproximarse, para pequeños valores de V_D por la expresión:

$$I_D \cong \frac{Z}{L} \cdot \mu_n \cdot C_o \cdot (V_G - V_T) \cdot V_D \quad \text{para} \quad V_D \ll V_G - V_T$$

siendo V_T la anterior tensión umbral dada por una expresión del tipo:

$$V_T \cong \frac{\sqrt{2 \cdot \epsilon_s \cdot q \cdot N_A \cdot 2 \cdot \psi_B}}{C_o} + 2 \cdot \psi_B$$

En la región lineal, a partir de la aproximación hecha para la corriente I_D se puede obtener la conductancia del canal g_D y la transconductancia g_m :

$$g_D \equiv \left. \frac{\partial I_D}{\partial V_D} \right|_{V_G = \text{cte}} \equiv \frac{Z}{L} \cdot \mu_n \cdot C_o \cdot (V_G - V_T)$$

$$g_m \equiv \left. \frac{\partial I_D}{\partial V_G} \right|_{V_D = \text{cte}} \equiv \frac{Z}{L} \cdot \mu_n \cdot C_o \cdot V_D$$

La región de saturación se alcanza cuando la carga eléctrica $Q_n(y)$ en el canal para el valor $y=L$ se anula. Es bajo esta condición cuando el número de electrones móviles en el drenador se reduce drásticamente. Este es el llamado punto de *pinch-off* o de estrangulamiento del canal, análogo al que se daba en el transistor JFET. La corriente de drenador y la tensión en este punto son V_{Dsat} e I_{Dsat} .

Por encima de la tensión V_{Dsat} , la corriente queda prácticamente constante. Su valor se obtiene a partir de la condición $Q_n(y)$:

$$V_{Dsat} \equiv V_G - 2 \cdot \psi_B + K^2 \cdot \left[1 - \sqrt{1 + 2 \cdot V_G / K^2} \right]$$

donde:

$$K \equiv \frac{\sqrt{\epsilon_s \cdot q \cdot N_A}}{C_o}$$

Con el valor de V_{Dsat} y sustituyendo en la característica inicial corriente-tensión, se obtiene I_{Dsat} :

$$I_{Dsat} \equiv \frac{Z \cdot \mu_n \cdot \epsilon_{ox}}{2 \cdot d \cdot L} \cdot [V_G - V_T]^2 \text{ pues } C_o \equiv \frac{\epsilon_{ox}}{d}$$

Puede observarse como en la región de saturación la conductancia del canal es cero y la transconductancia toma la expresión:

$$g_m \equiv \left. \frac{\partial I_D}{\partial V_G} \right|_{V_D = \text{cte}} \equiv \frac{Z \cdot \mu_n \cdot \epsilon_{ox}}{d \cdot L} \cdot [V_G - V_T]$$

tal y como ocurría en el JFET, g_m en la región de saturación coincide con g_D en la región lineal.

5.2.2. La región subumbral.

Esta región o modo de conducción del transistor MOSFET se corresponde con tensiones de puerta por debajo de la tensión umbral V_T . El estudio de esta región es importante en aplicaciones en que el transistor es empleado como conmutador a fin de tener una buena comprensión de cómo el transistor conmuta a corte y a conducción. Dentro de dichas aplicaciones cabe destacar las aplicaciones en lógica digital y memorias.

En el modo de conducción subumbral, el transistor está regido más por el fenómeno de la difusión de portadores que por el de arrastre. En esta situación se tiene:

$$I_D = -q \cdot A \cdot D_n \cdot \frac{dn}{dy} = q \cdot A \cdot D_n \cdot \frac{n(0) - n(L)}{L}$$

donde A es la sección transversal del canal y $n(0)$, $n(L)$ la concentración de portadores en la fuente y drenador, respectivamente. Dichas concentraciones vienen dadas por las expresiones:

$$n(0) = n_i \cdot e^{q(\psi_s - \psi_B)/K.T}$$

$$n(L) = n_i \cdot e^{q(\psi_s - \psi_B - V_D)/K.T}$$

donde ψ_s es el potencial de superficie en la fuente. Sustituyendo ambos resultados en la expresión de I_D se obtiene:

$$I_D = \frac{q \cdot A \cdot D_n \cdot n_i \cdot e^{-q\psi_B/K.T}}{L} \cdot (1 - e^{-qV_D/K.T}) \cdot e^{q\psi_s/K.T}$$

el potencial de superficie ψ_s es aproximadamente $V_G - V_T$. En consecuencia la corriente de drenador decrece exponencialmente cuando V_G es menor que V_T , es decir:

$$I_D \propto e^{q(V_G - V_T)/K.T}$$

Fig. 14 recoge una característica típica en la región subumbral. Se observa la dependencia exponencial de I_D cuando $V_G < V_T$. Si se desea reducir la corriente I_D en la región subumbral a un valor despreciable en la práctica, el transistor MOSFET se polariza a la mitad o menos de V_T .

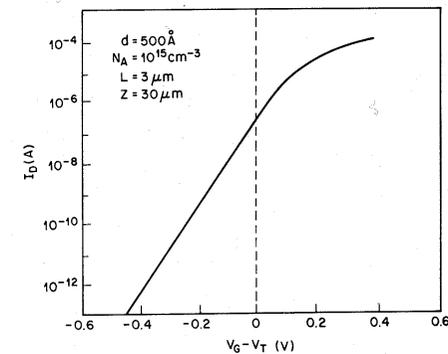


Fig. 14: Característica subumbral del MOSFET.

5.2.3. Tipos de transistores MOSFET.

Básicamente existen cuatro tipos fundamentales de transistores MOSFET, los cuales dependen del tipo de canal creado. Por ejemplo, con polarización de puerta nula, la conductancia del canal es muy baja y si, al aplicar una tensión de puerta positiva al canal de tipo N, este aumenta progresivamente su conductividad se dice que se trata de un transistor de canal N de enriquecimiento. Por el contrario, si con polarización nula en la puerta, la conductividad del canal es apreciable se tiene un transistor MOSFET de canal N de empobrecimiento o vaciamiento. También suelen llamarse respectivamente transistores normalmente OFF o normalmente ON. De igual forma se puede tener con los transistores de canal P, sólo que ahora se cambian las polaridades.

Fig. 15 recoge una representación de los cuatro tipos posibles en cuanto a sección transversal, características de salida, de transferencia y símbolo electrónico.

Obsérvese la diferencia entre el transistor MOSFET normalmente OFF y el normalmente ON. Para este último fluye corriente apreciable con tensión nula entre puerta y fuente. El normalmente OFF necesita de una tensión por encima de la umbral para acusar un flujo de corriente apreciable.

TYPE	CROSS SECTION	OUTPUT CHARACTERISTICS	TRANSFER CHARACTERISTICS	Símbolo electrónico
n-CHANNEL ENHANCEMENT (NORMALLY OFF)				
n-CHANNEL DEPLETION (NORMALLY ON)				
p-CHANNEL ENHANCEMENT (NORMALLY OFF)				
p-CHANNEL DEPLETION (NORMALLY ON)				

Fig. 15: Sección transversal, características de salida y de transferencia y símbolos electrónicos de los cuatro tipos de MOSFETs.

5.2.4. Integración de dispositivos y efectos de canal corto.

Para aumentar el número de componentes integrados en una pastilla es necesario lógicamente reducir las dimensiones de cada uno de los transistores. En la actualidad se han conseguido dispositivos MOS con longitud de la puerta entre 1-2µm y se espera bajar de la región del µm en un futuro muy próximo. Sin embargo, en este afán por la miniaturización aparecen nuevos efectos que deben ser considerados, entre ellos los de más relevancia son los debidos a la longitud extremadamente corta del canal.

Se considerarán en primer lugar los efectos debidos al canal corto y, a continuación, cómo evitarlos de forma que se consiga un aumento en la capacidad de integración.

5.2.4.1 Efectos de canal corto.

A medida que la longitud del canal L disminuye como consecuencia de la miniaturización, las anchuras de las regiones espaciales de carga de drenador y fuente W_D y W_S son comparables con ella. Si se utiliza la aproximación de una unión abrupta, las anchuras de las regiones espaciales de carga son:

$$W_S = \sqrt{\frac{2 \cdot \epsilon_s}{q \cdot N_A} \cdot (V_{bi} + V_{BS})}$$

$$W_D = \sqrt{\frac{2 \cdot \epsilon_s}{q \cdot N_A} \cdot (V_D + V_{bi} + V_{BS})}$$

donde V_{BS} es la magnitud de la polarización del sustrato.

Cuando se cumpla la condición W_S+W_D=L ocurrirá el efecto “punch-through” ambas regiones espaciales de carga se solaparán y la puerta perderá el control de la corriente de drenador.

En una aplicación típica, la tensión de polarización V_D se mantiene en un valor constante (por ejemplo, 5V). A medida que la longitud del canal se reduce, la componente longitudinal del campo eléctrico ε_y se incrementa, la movilidad del canal se hace dependiente del campo eléctrico, llegándose a una eventual saturación de la velocidad de los portadores. Incluso para un pequeño campo eléctrico ε_y, la movilidad de los portadores en el canal disminuye con respecto a la del semiconductor extrínseco con la misma cantidad de impurezas. En Fig. 16 se muestra la medida de la movilidad en función del campo eléctrico transversal ε_x. La movilidad para el silicio dopado con N_A=10¹⁵cm⁻³ tiene un valor de 1500cm²/V·s. En un MOSFET, el transporte de portadores está confinado dentro de una estrecha región de inversión (la capa de inversión es de entre 10 y 100Å) y las colisiones con la superficie (mostradas en la misma figura) provocan una reducción de la movilidad. La movilidad promedio del canal es de aproximadamente la mitad de la del semiconductor dopado con la misma cantidad de impurezas.

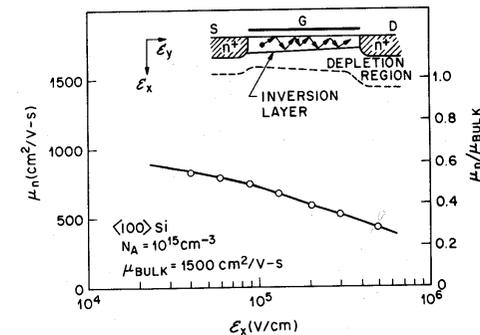


Fig. 16: Movilidad en función del campo eléctrico longitudinal en el canal de inversión de un MOSFET de canal N.

Fig. 17 muestra la medida de la velocidad de arrastre de los electrones en un dispositivo de canal N para un determinado valor del campo transversal ε_x y en función del campo longitudinal ε_y. Para bajos valores de ε_y (del orden de 10³V/cm), la velocidad de arrastre varía de forma lineal con ε_y; dando lugar a una movilidad constante. Este valor de la movilidad corresponde al mostrado en Fig. 16. Conforme ε_y se incrementa, la velocidad de arrastre tiende a incrementarse de forma más lenta; y cuando ε_y alcanza el valor de 10⁵V/cm la velocidad de arrastre de los electrones llega a un valor de saturación llamado v_s (v_s = 9·10⁶cm/s).

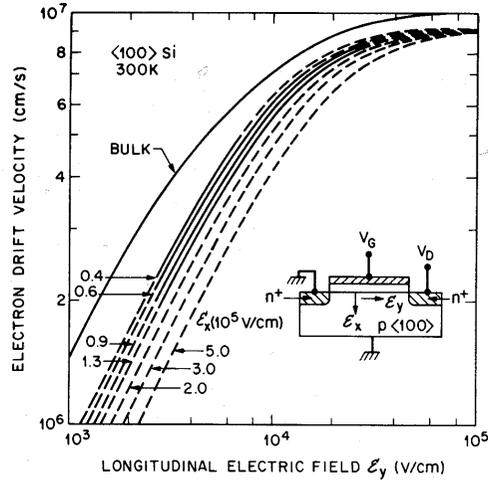


Fig. 17: Velocidad de arrastre de los electrones en el canal de inversión en función del campo eléctrico longitudinal para un campo transversal determinado.

En esta situación se llega a una corriente de saturación dada por el producto del número de portadores por su velocidad y por q. Con $v = v_s$ tenemos:

$$I_{Dsat} = Z \cdot q \cdot v_s \cdot \int_0^{x_i} n(x) \cdot dx$$

donde x_i es el grosor del canal. La integral es precisamente la carga del canal por unidad de área, Q_n . En el extremo del canal próximo a la fuente, dicha carga vale $(V_G - V_T) \cdot C_o$ y en consecuencia:

$$I_{Dsat} \cong Z \cdot C_o \cdot v_s \cdot (V_G - V_T)$$

donde se ha despreciado la variación de Q_n a lo largo del canal. La transconductancia permanece constante:

$$g_m = \frac{\partial I_{Dsat}}{\partial V_G} = Z \cdot C_o \cdot v_s$$

Cuando se llega a la velocidad de saturación debido al hecho de tener un dispositivo de canal corto, tanto la corriente (I_{Dsat}) como la transconductancia del canal (g_m) se reducen sustancialmente con respecto a lo que sucedería si el dispositivo fuese de canal largo.

Los efectos de canal corto, por tanto, complican el funcionamiento del dispositivo y degradan sus prestaciones por lo que deberían ser eliminados o al menos minimizados de manera que los dispositivos de canal corto conserven las características eléctricas de los dispositivos de canal largo.

5.2.4.2. Miniaturización de dispositivos.

Uno de los métodos empleados para evitar los efectos de canal corto es realizar una reducción a escala de todas las dimensiones del dispositivo y tensiones en un determinado factor de escala $K > 1$, de manera que los campos eléctricos internos son los mismos que para un MOSFET de canal largo. Las nuevas dimensiones geométricas serán:

$$L' = \frac{L}{K}, d' = \frac{d}{K} \text{ y } Z' = \frac{Z}{K}$$

Para un campo eléctrico constante, las tensiones de funcionamiento variarán de la forma:

$$V' = \frac{V}{K}$$

Las nuevas magnitudes físicas relativas al transistor MOSFET ya escaladas serán:

$$C'_o = \frac{\epsilon_{ox}}{d/K} = K \cdot \frac{\epsilon_{ox}}{d} = K \cdot C_o \quad (F/cm^2)$$

$$(C_o \cdot A) = (K \cdot C_o) \cdot \frac{Z}{K} \cdot \frac{L}{K} = \frac{C_o \cdot A}{K} \quad (F)$$

$$I'_{Dsat} = \frac{Z}{K} \cdot \frac{(K \cdot C_o) \cdot v_s \cdot (V_G - V_T)}{K} = \frac{I_{Dsat}}{K} \quad (A)$$

$$J'_{Dsat} = \frac{I'_{Dsat}}{A'} = \frac{I_{Dsat}}{K} \cdot \frac{K^2}{A} = K \cdot J_{Dsat} \quad (A/cm^2)$$

La potencia de conmutación P_{ac} y la de continua P_{dc} son también reducidas en K^2 :

$$P'_{ac} = \frac{P_{ac}}{K^2} \quad (W)$$

$$P'_{dc} = I' \cdot V' = \left(\frac{I}{K}\right) \cdot \left(\frac{V}{K}\right) = \frac{P_{dc}}{K^2} \quad (W)$$

Y la energía de conmutación E' :

$$E' = \frac{1}{2} \cdot (C_o \cdot A) \cdot V'^2 = \frac{1}{2} \cdot \frac{C_o \cdot A}{K} \cdot \left(\frac{V}{K}\right)^2 = \frac{E}{K^3} \quad (J)$$

Por lo tanto, cuando el dispositivo es escalado, todos los valores cambian favorablemente. Se incrementa la velocidad de funcionamiento y la densidad de componentes, y la densidad de potencia permanece constante. Sin embargo, la densidad de corriente se incrementa en el factor de escala. En metales conductores la densidad de corriente está limitada a un valor máximo impuesto por electromigración, es decir, el movimiento de los átomos de un sitio a otro bajo la influencia de una fuerza de tipo eléctrico. Este efecto limita la máxima densidad de corriente para el caso del aluminio en alrededor de $10^5 A/cm^2$. Por tanto, Por tanto el proceso de reducción a escala viene limitado por la conductividad de los metales empleados como conductores.

Fig. 18 recoge las características de salida de un dispositivo reducido y otro normal. Se observa que si bien la tensión umbral ha sido reducida en un factor K, la capacidad de corriente es la misma para ambos dispositivos y, en consecuencia, la densidad de corriente es mayor en el dispositivo reducido.

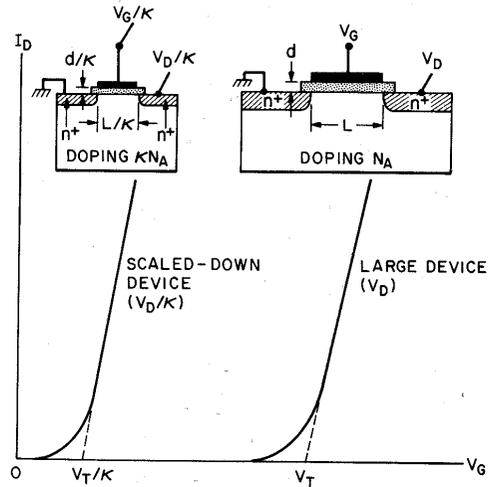


Fig. 18: Escalado de un MOSFET con factor de escala K.

Hay también otro método alternativo al de la reducción a escala y que permite un diseño mucho más flexible de las distintas variables del dispositivo. Este método es más empírico que analítico y consiste en encontrar un valor mínimo de la longitud del canal L_{min} en el cual es posible miniaturizar el dispositivo conservando las propiedades características de un MOS de "grandes dimensiones".

Llamando L_{min} a la mínima longitud del canal para la cual puede ser observado un comportamiento de "canal largo", y $\gamma = r_j \cdot d \cdot (W_s + W_D)^2$ donde r_j es el grosor de la unión (en μm), d es el grosor del óxido (en \AA) y W_s y W_D la anchura de las regiones espaciales de carga de fuente y drenador (en μm); puede ser obtenida, tras un laborioso estudio experimental y simulaciones bidimensionales por ordenador la siguiente ecuación empírica:

$$L_{min} \cong 0.4 \cdot [r_j \cdot d \cdot (W_s + W_D)^2]^{1/3} = 0.4 \cdot \gamma^{1/3}$$

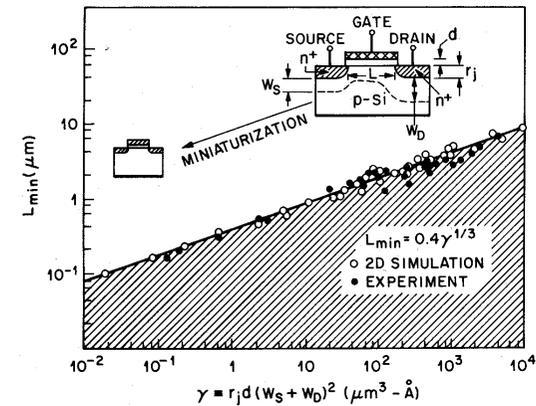


Fig. 19: L_{min} en función de γ , donde L_{min} es el valor mínimo de la longitud del canal para la cual puede ser observado un comportamiento de canal largo.

Por ejemplo, para diseñar un MOSFET con un canal de $0.5 \mu m$ con un comportamiento adecuado de canal largo, el parámetro γ debería ser 2. Una vez determinado γ se seleccionan r_j , d , W_s y W_D de manera que el γ resultante no sea mayor que 2.

La expresión de L_{min} puede ser utilizada para hacer un escalado mucho más flexible, ya que permite que varios parámetros del dispositivo sean ajustados de forma independiente siempre y cuando el valor de γ se mantenga constante. En consecuencia, no todos los parámetros del dispositivo han de ser escalados en el mismo factor K. Esta flexibilidad nos permite escoger nuevas geometrías que sean fáciles de realizar o optimizar otros aspectos del funcionamiento del dispositivo, lo cual sería difícil con el anterior método en que todos los parámetros eran escalados en un mismo factor K.