

T6 MODELO FUNCIONAL DEL TRANSISTOR MOS

T6.1. Funcionamiento del transistor NMOS

T6.2. Etapa en fuente común

T6.3. El transistor MOS real: efectos de segundo orden

T6.4. Modelos SPICE

El estudio de la microelectrónica digital, además de abarcar el proceso de diseño del circuito electrónico y las posibilidades de programación o fabricación del mismo, debe atender, también, a la tecnología, al «soporte material» sobre el cual se va a configurar el circuito diseñado.

Se trata de insertar el circuito digital resultante de un diseño específico en el interior de un circuito integrado, bien por configuración de un circuito integrado programable, bien por fabricación del circuito integrado según el diseño propio.

Es necesario conocer el soporte material que permite tal programación o fabricación, estudiar con un poco de detalle la tecnología que permite «materializar» nuestro diseño «particular». Disponer de un modelo básico conceptual y operativo de dicha tecnología permitirá conocer sus posibilidades, sus prestaciones, sus limitaciones y sus exigencias (los requisitos que la tecnología impone sobre el diseño).

Tal es el objeto de estudio de los capítulos T6, T7, T8 y T9. En éste, primero de ellos, se presenta un modelo funcional del comportamiento de los transistores MOS, empleando los «diagramas de Memelink» como herramienta gráfica muy útil para representar la distribución de carga en el canal de los transistores y para efectuar cálculos relativos a tensiones e intensidades (y, también, a tiempos de propagación y consumos).

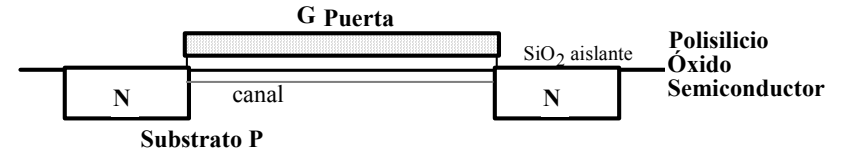
El comportamiento de los transistores MOS ya fue descrito, con cierto detalle, en el capítulo T1 (volumen 1). En los apartados que siguen se repetirá, ampliándola, tal descripción; se ha procurado resumir en lo posible lo ya dicho en T1.3., de manera que la consulta a dicho apartado puede ser útil para razonar y comprender aquellas cuestiones que, por resumidas, no hayan quedado claras al estudiar el presente capítulo. Asimismo, se recomienda la lectura de todo el capítulo T1 como introducción previa.

Junto con el modelo funcional (gráfico) del comportamiento MOS, se detallan aquellos aspectos de segundo orden que separan a los transistores reales de dicho modelo y que permiten comprender mejor y simular más fielmente su funcionamiento.

Respecto a este capítulo y a los siguientes, debe apreciarse que el estudio de la tecnología con la cual pueden fabricarse nuestros diseños (y con la cual se fabrican, también, los dispositivos en los que pueden programarse) no es una mera «opción cultural» no es una simple oportunidad de conocer cómo son «por dentro» los circuitos integrados reales. Sino que, de dicho estudio, se derivan muchas cuestiones prácticas, muchas limitaciones, condiciones o requisitos que han de tenerse en cuenta para hacer efectiva la integración de un sistema digital.

T6.1. Funcionamiento del transistor NMOS

La idea básica del transistor MOS fue anterior a la del bipolar (1947); la estructura MOS presenta una gran sencillez conceptual, por su analogía directa con un interruptor:



Una tensión positiva suficientemente alta sobre el terminal de puerta V_G da lugar, por atracción sobre los portadores minoritarios del sustrato, a un canal continuo entre las dos zonas de difusión N, que permite el flujo de corriente entre ellas:

$V_G < V_{TO}$ no existe canal \Rightarrow el transistor no conduce

$V_G > V_{TO}$ se forma canal \Rightarrow el transistor conduce

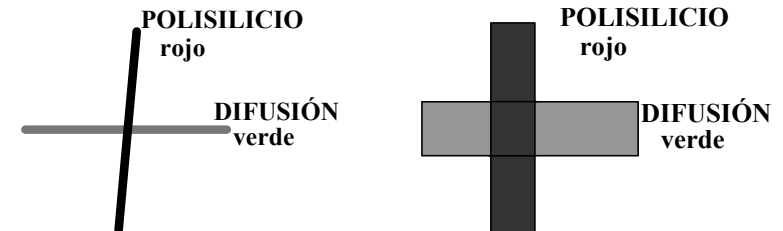
Siendo V_{TO} la tensión umbral de conducción.

La fabricación efectiva de transistores MOS se retrasó hasta los años 60 por problemas técnicos: el transistor es un efecto superficial, no volumétrico (caso de los transistores bipolares) y se requieren técnicas de extraordinaria limpieza en la superficie.

En el proceso de integración se forma primero la «banda» de óxido de puerta y polisilicio (cuya anchura determina la longitud del transistor) y, luego, en sentido perpendicular, la «banda» de difusión que forma la fuente y el sustrato; de esta forma, la puerta y las dos zonas de difusión N quedan autoalineadas (la difusión se realiza sobre todo el transistor pero no penetra debajo del polisilicio).

El transistor se forma físicamente por el cruce de dos regiones: polisilicio y difusión. Debajo del polisilicio la difusión se encuentra interrumpida, pero puede inducirse, por medio de la tensión de puerta, su continuidad dando lugar a la formación de un canal.

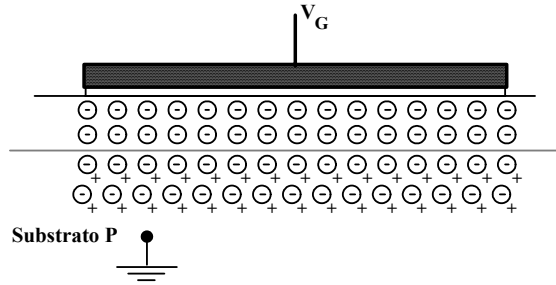
Abstrayendo las dimensiones superficiales, el transistor puede representarse en un diagrama lineal como un cruce entre dos líneas: difusión - polisilicio.



[Al dibujar las correspondientes máscaras, suele utilizarse color rojo para la de polisilicio y color verde para la difusión.]

a) Capa vacía

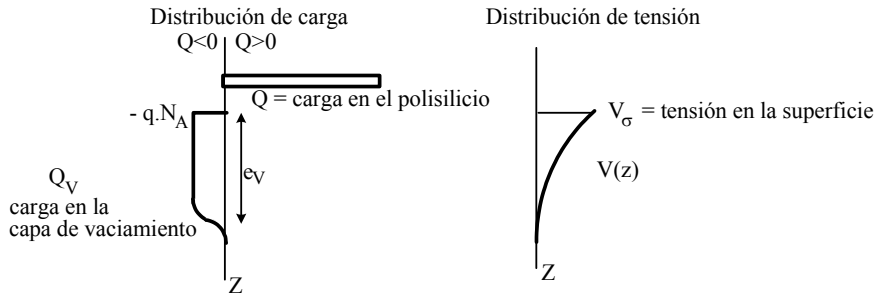
El sustrato de los transistores NMOS es de tipo P, dopado con átomos de Boro con sólo 3 electrones de valencia, que dejan un hueco en la estructura cristalina. Cuando se aplica una tensión positiva a la puerta, $V_G > 0$, se forma una capa de vaciamiento, desprovista de portadores, debida a la repulsión de la tensión de puerta sobre los huecos del sustrato; los átomos de boro quedan cargados negativamente.



La densidad superficial de carga (carga por unidad de superficie) Q será:

$$Q_V = q \cdot N_A \cdot e \cdot v$$

q = carga del electrón,
 N_A = dopado del sustrato = nº de huecos por cm^3 ,
 $e \cdot v$ = profundidad de la capa de vaciamiento.



Integrando la ecuación de Poisson para la tensión, resulta:

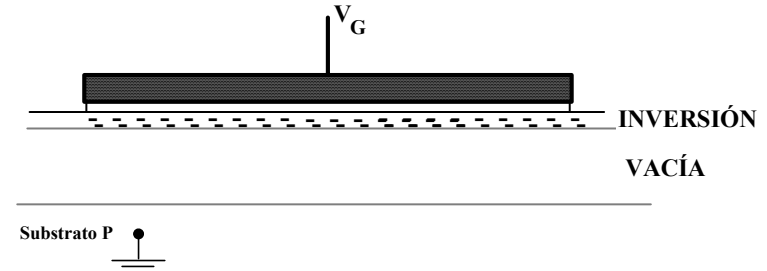
$$\frac{d^2V}{dz^2} = \frac{\rho(z)}{\epsilon} \Rightarrow V_\sigma = \frac{q \cdot N_A}{2 \cdot \epsilon_{Si}} \cdot e \cdot v^2$$

siendo V_σ la tensión en la superficie del semiconductor; sustituyendo $e \cdot v$ en Q_V se obtiene la relación entre la carga de la capa de vaciamiento y la tensión en la superficie:

$$Q_V = \sqrt{2 \cdot \epsilon_{Si} \cdot q \cdot N_A} \cdot \sqrt{V_\sigma} = K_1 \sqrt{V_\sigma}, \text{ donde } K_1 = \sqrt{2 \cdot \epsilon_{Si} \cdot q \cdot N_A}.$$

b) Capa de inversión: canal conductor

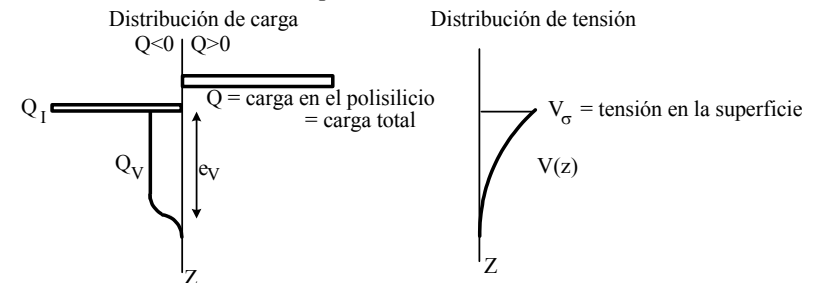
Cuando la tensión de puerta es suficientemente positiva, se forma una capa superficial de electrones libres, arrastrados por el gradiente de tensión de la capa de vaciamiento: capa de inversión o canal; tales electrones libres proceden de los pares electrón-hueco del propio silicio semiconductor y su número es reducido.



Para formar una capa de inversión o canal conductor se necesita una tensión en la superficie del semiconductor Φ_B capaz de mantener una concentración de electrones equivalente al dopado (concentración de huecos) del sustrato, es decir, el sustrato P se vuelve N en la zona superficial; el valor de la tensión superficial necesaria para un «dopado inverso» equivalente al del sustrato N_A , *condición de fuerte inversión*, es:

$$\phi_B = 2 \frac{KT}{q} \ln \frac{N_A}{N_i}$$

siendo K la constante de Boltzman, T la temperatura absoluta y N_i la concentración de pares electrón-hueco del silicio a esa temperatura.



La carga almacenada en el «condensador» puerta-sustrato tiene dos componentes: carga correspondiente a la capa vacía + carga debida al canal o capa de inversión, $Q = Q_V + Q_I$; la carga total Q será proporcional a la diferencia de potencial existente entre puerta y sustrato y el factor de proporcionalidad corresponderá a la capacidad de dicho condensador c_{ox} . En las siguientes expresiones las diversas cargas están expresadas por unidad de superficie y lo mismo sucede con la capacidad de puerta c_{ox} .

$$\left. \begin{aligned} Q &= Q_I + Q_V \\ Q &= c_{ox} (V_G - \Phi_B) \\ Q_V &= K_1 \sqrt{V_\sigma} \end{aligned} \right\} \begin{aligned} Q &= \text{carga por unidad de superficie} \\ c_{ox} &= \text{capacidad por unidad de superficie} \end{aligned}$$

$$Q_I = Q - Q_V = c_{ox}(V_G - \phi_B) - K_1 \sqrt{\phi_B} = c_{ox} [V_G - (\phi_B + \gamma \sqrt{\phi_B})]$$

$$Q_I = C_{ox} (V_G - V_{TO}) \quad \text{donde} \quad V_{TO} = \phi_B + \gamma \sqrt{\phi_B}$$

$$\text{y} \quad \gamma = \frac{K_1}{c_{ox}} = \frac{\sqrt{2 \cdot \epsilon_{Si} \cdot q \cdot N_A}}{c_{ox}}$$

V_{TO} es la tensión umbral de conducción del MOS (tensión de puerta necesaria para que se forme la capa de inversión, en condiciones de fuerte inversión) y γ es el parámetro de efecto sustrato (expresa el efecto de la capa de vaciamiento sobre V_{TO}).

Comentario: sobre V_{TO} inciden otros efectos de segundo orden (ver apartado 3); en la práctica, dicho valor lo fija el fabricante, mediante un proceso de implantación iónica en la superficie del sustrato.

Valores de los diversos parámetros (tecnología Ψ):

Para tener una idea más concreta y una referencia cuantitativa de los resultados, en este capítulo y en los siguientes se incluyen valores de los diversos parámetros y se efectúan los cálculos correspondientes, en relación a una tecnología hipotética pero realista a la que denominaremos tecnología Ψ .

La tecnología Ψ es de 1 micra; sus transistores tendrán las siguientes dimensiones: 1 μm de longitud mínima y 1,5 μm de anchura mínima y el espesor del óxido de puerta t_{ox} será de 0,05 μm .

La dimensión mínima de una tecnología de integración CMOS viene dada por la anchura mínima del polisilicio, que corresponde a la longitud del transistor: $L_{\text{mín}}$. En general, la anchura de la difusión (que coincide con la del transistor W) no puede ser tan reducida como la del polisilicio; lo habitual en la mayoría de las tecnologías es que $W_{\text{mín}} = 1,5 \cdot L_{\text{mín}}$ y, por ello, los transistores de tamaño mínimo no serán cuadrados sino que su factor de forma será $ff = W/L = 1,5$.

A temperatura ambiente ($27^\circ\text{C} = 300^\circ\text{K}$):

N (densidad átomos de Si) $\sim 10^{22}$ átomos Si/cm³ = 10^{10} ats/ μm^3

N_i (pares e-h Si puro) $\sim 10^{10}$ pares e-h/cm³

1 par e-h intrínseco por cada billón de átomos de Si

N_A (dopado sustrato) $\sim 10^{15}$ ats B/cm³ = 10^3 ats/ μm^3

1 átomo de Boro por cada 10 millones de átomos de Si

1 hueco extrínseco por cada 10 millones de átomos de Si

100.000 veces más portadores extrínsecos que intrínsecos

La capacidad de puerta, por unidad de superficie, será:

$$c_{ox} = \frac{\epsilon_{SiO_2}}{t_{ox}} \quad \left| \begin{array}{l} \epsilon_{SiO_2} = 3,97 \cdot \epsilon_0 \approx 0,035 \text{ fF}/\mu\text{m}; \epsilon_0 = 8,85 \text{ pF/m} \\ t_{ox} \text{ (espesor óxido)} \approx 0,05 \mu\text{m} \end{array} \right. \quad c_{ox} \approx 0,7 \text{ fF}/\mu\text{m}^2$$

Un condensador MOS de 1 mm² presentará una capacidad ~ 1 nF; dicho efecto capacitivo se refiere al cociente $Q_I / (V_G - V_{TO})$: es una capacidad de tipo incremental $C = \Delta Q / \Delta V_G$.

Un transistor MOS de tamaño mínimo (1 $\mu\text{m} \times 1,5 \mu\text{m}$) tendrá una capacidad ~ 1 fF en su zona de puerta.

Respecto a la capa de inversión o canal conductor:

$$\Phi_B = 2 \frac{KT}{q} \ln \frac{N_A}{N_i}; \quad \frac{KT}{q} \approx 0,026 \text{ V (a } 25^\circ\text{C)} \quad \Phi_B \approx 0,6 \text{ V}$$

$$\gamma = \frac{\sqrt{2 \cdot \epsilon_{Si} \cdot q \cdot N_A}}{c_{ox}} \quad \left| \begin{array}{l} \epsilon_{Si} = 11,7 \cdot \epsilon_0 \approx 0,1 \text{ fF}/\mu\text{m} \\ q \text{ (carga electrón)} = 16 \times 10^{-20} \text{ C} \end{array} \right. \quad \gamma = 0,25 \text{ V}^{1/2}$$

$$V_{TO} = \Phi_B + \gamma \cdot \sqrt{\Phi_B} \quad V_{TO} \approx 1 \text{ V}$$

También puede calcularse la profundidad de la capa de vaciamiento:

$$\Phi_B = \frac{q \cdot N_A}{2 \cdot \epsilon_{Si}} \cdot e_v^2 \quad e_v = \sqrt{\frac{2 \cdot \epsilon_{Si}}{q \cdot N_A}} \cdot \sqrt{\Phi_B} \approx 1 \mu\text{m}$$

c) Tensión de referencia (en el canal) $\neq 0$

Si sobre el sustrato en su zona superficial se induce, por algún medio, una tensión V' , dicha tensión actúa como referencia o nivel 0 de tensiones, de forma que la tensión efectiva sobre el condensador MOS será $V_G - V'$:

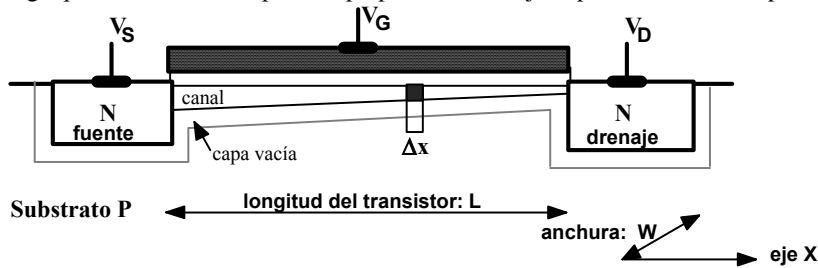
$$Q_I = c_{ox} \cdot (V_G - V' - V_{TO}).$$

Esto ocurre al polarizar la fuente y el drenaje para que circule una corriente: las tensiones V_D y V_S establecen una distribución de tensiones en la zona superficial del sustrato, de V_S (en el extremo de la fuente) hasta V_D (en el del drenaje). Orientando el eje X en la dirección del canal, la concentración de portadores de carga en el canal será:

$$Q_I = c_{ox} \cdot (V_G - V_{TO} - V(x)) \quad \text{donde } V(x) \text{ varía entre } V_S \text{ y } V_D.$$

d) Intensidad de corriente a través del canal

La diferencia de tensión V_{DS} da lugar a una intensidad de corriente que corresponde a la carga que atraviesa una superficie perpendicular al eje X por unidad de tiempo.



La distribución de portadores libres en la capa de inversión variará a lo largo del canal: será mayor al lado de la fuente que en el del drenaje ($V_{DS} > 0, V_G - V_S > V_G - V_D$). En un elemento diferencial Δx la cantidad de carga que corresponde a los portadores libres, a todo lo ancho W del transistor será:

$$\Delta Q_I = [\text{densidad de carga}] \times \text{volumen} = Q_I \cdot W \cdot \Delta x = c_{ox} \cdot (V_G - V_{TO} - V(x)) \cdot W \cdot \Delta x$$

$$I = \frac{dQ_I}{dt} = \frac{\Delta Q_I}{\Delta x} \cdot \frac{\Delta x}{\Delta t} = \frac{\Delta Q_I}{\Delta x} \cdot v \quad \frac{\Delta Q_I}{\Delta x} : \text{carga de portadores por unidad de longitud}$$

v es la velocidad de los portadores, que es proporcional al campo eléctrico: la constante de proporcionalidad es la movilidad de los portadores de carga μ

$$\left. \begin{aligned} \frac{\Delta Q_I}{\Delta x} &= c_{ox} \cdot W \cdot (V_G - V_{TO} - V(x)) \\ v &= \mu \cdot E = \mu \cdot \frac{dV}{dx} \end{aligned} \right\} I = \mu \cdot c_{ox} \cdot W \cdot (V_G - V_{TO} - V(x)) \cdot \frac{dV(x)}{dx}$$

Integrando la anterior ecuación diferencial a lo largo del canal resulta

$$\int_0^L I dx = \int_{V_S}^{V_D} \mu \cdot c_{ox} \cdot W \cdot (V_G - V_{TO} - V(x)) dV$$

$$I = \mu \cdot c_{ox} \cdot \frac{W}{L} \cdot \int_{V_S}^{V_D} (V_G - (V_{TO} + V(x))) dV = \alpha \cdot A$$

donde $A = \int_{V_S}^{V_D} (V_G - (V_{TO} + V(x))) dV$ expresa el efecto de las tensiones aplicadas sobre

el transistor y $\alpha = \mu \cdot c_{ox} \cdot W/L$ incluye parámetros tecnológicos y geométricos:

K_p = coeficiente de transconductancia = $\mu \cdot c_{ox}$ (parámetros tecnológicos)

ff = factor de forma = W/L , cociente entre anchura W y longitud L del transistor.

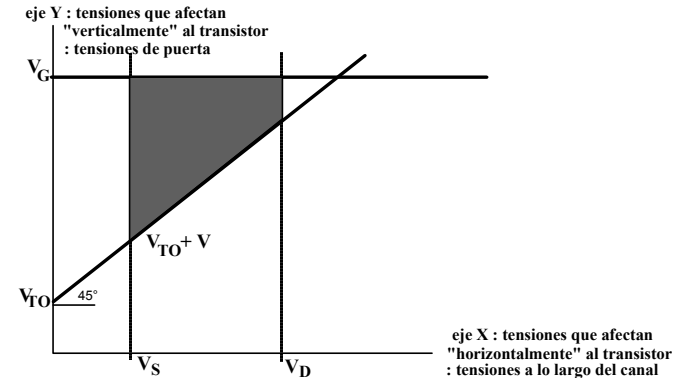
Ambos influyen en proporcionalidad directa sobre la intensidad que conduce el transistor.

Para los transistores NMOS de la tecnología Ψ :

$$\mu_e \sim 600 \text{ cm}^2/V \cdot s, \quad c_{ox} \sim 0,7 \text{ fF}/\mu^2 = 0,07 \text{ }\mu\text{F}/\text{cm}^2, \quad K_p \sim 600 \cdot 0,7 \approx 40 \text{ }\mu\text{A}/V^2.$$

Diagramas de Memelink

La integral A puede resolverse en forma gráfica mediante los denominados *diagramas de Memelink* que representan, en el eje horizontal X , las tensiones aplicadas a fuente y drenaje (y por tanto la distribución de tensiones a lo largo del canal) y, en el eje vertical Y , la tensión aplicada a la puerta y la tensión umbral efectiva $V_{TO} + V$ (de modo que su diferencia $V_G - (V_{TO} + V)$ expresa la tensión efectiva generadora del canal):



$$A = \int_{V_S}^{V_D} (V_G - (V_{TO} + V(x))) dV = \text{área del trapecio}$$

$$= \left(\frac{(V_G - V_{TO} - V_S) + (V_G - V_{TO} - V_D)}{2} \right) \cdot (V_D - V_S)$$

$$= \left(\frac{(V_{GS} - V_{TO}) + (V_{GS} - V_{TO} - V_{DS})}{2} \right) \cdot V_{DS} = \left(V_{GS} - V_{TO} - \frac{V_{DS}}{2} \right) \cdot V_{DS}$$

En un *diagrama de Memelink* el área limitada por las cuatro rectas $x = V_S, x = V_D, y = V_G$ e $y = V_{TO} + V$ representa la distribución de carga a lo largo del canal; es, por tanto, una imagen del propio canal. Téngase en cuenta que el eje de abscisas no es espacial, sino un eje de tensiones; por ello, el diagrama representa la distribución de carga en el canal, en función de la distribución de tensiones a lo largo el mismo.

Integrando dicho área y multiplicándola por $\alpha = \mu \cdot c_{ox} \cdot W/L$ (que expresa los efectos tecnológicos y geométricos) se obtiene la ecuación de la corriente en un transistor MOS (canal no saturado):

$$I = K_p \cdot \frac{W}{L} \cdot \left(V_{GS} - V_{TO} - \frac{V_{DS}}{2} \right) \cdot V_{DS}$$

e) Diferencia de tensión drenaje-fuente pequeña

Para V_{DS} muy pequeño: $V_{DS} \ll V_{GS}$, $V_D \approx V_S$, la expresión $(V_G - V_{TO} - V(x))$ es prácticamente constante, $V_{GS} - V_{TO}$; la tensión aplicada es la misma a lo largo de todo el canal, que resulta plano, con una distribución de carga uniforme:

$$I_D = K_p \cdot \frac{W}{L} \cdot (V_{GS} - V_{TO}) \cdot V_{DS}$$

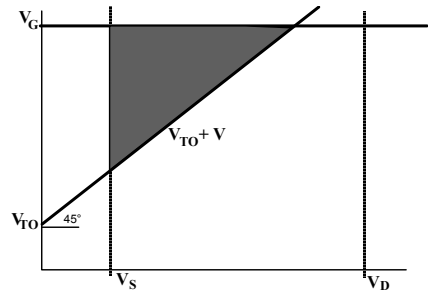
expresión que muestra una proporcionalidad directa entre la intensidad que pasa por el transistor y la tensión aplicada sobre el mismo. El transistor se comporta como una simple resistencia cuyo valor disminuye al aumentar la tensión de puerta V_{GS} :

$$R_{eq} = \frac{I_D}{V_{DS}} = \frac{1}{K_p \frac{W}{L} (V_{GS} - V_{TO})}$$

Esta región de funcionamiento del transistor se denomina *zona lineal* o *zona óhmica* y es la situación booleana que corresponde a un transistor MOS en conducción: equivale a una resistencia cuyo valor puede hacerse adecuadamente pequeño, a través del factor de forma W/L (la resistencia es inversamente proporcional a su anchura W).

f) Canal saturado

Al aumentar V_{DS} el canal se hace más estrecho (presenta menor número de portadores) en el lado del drenaje ($V_{GD} < V_{GS}$) y llega un momento en que se satura, es decir, la diferencia de tensiones $V_G - V(x)$ no supera la tensión umbral $V_G - V(x) < V_{TO}$: en tal zona no hay capa de inversión que contribuya a la conducción: si $V_{Gx} \leq V_{TO}$ se cierra el canal.



A = área del triángulo rectángulo isósceles

$$= \frac{1}{2} (V_G - (V_{TO} + V_S))^2 = \frac{1}{2} (V_{GS} - V_{TO})^2$$

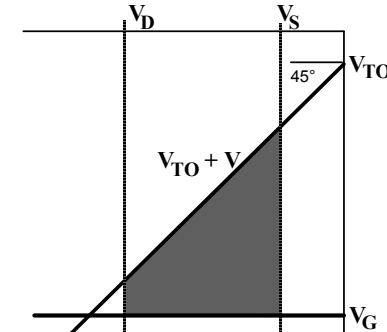
La situación de canal saturado se da si $V_G < V_{TO} + V_D$, $V_D > V_G - V_{TO}$, en cuyo caso la integral **A** se extiende solamente a la zona en que $V_G - V_{TO} - V(x)$ es positiva y viene dada por el área del triángulo formado por las rectas $x = V_S$, $y = V_G$ e $y = V_{TO} + V$:

$$I_{D,sat} = \frac{K_p}{2} \cdot \frac{W}{L} (V_{GS} - V_{TO})^2$$

expresión que representa la intensidad máxima que el transistor puede conducir para una tensión de puerta V_{GS} determinada; esta intensidad no depende de la tensión de drenaje y aumenta fuertemente con la tensión de puerta.

Transistores PMOS

En el transistor MOS de canal P son los huecos (en lugar de los electrones) los portadores que forman el canal; por ello, requiere tensiones de alimentación V_{DD} y de puerta V_G negativas respecto al sustrato y a la fuente, siendo también negativa su tensión umbral V_{TO} . Por ello, su *diagrama de Memelink* se sitúa en el tercer cuadrante:



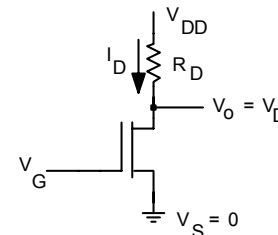
Además, la movilidad de los huecos es unas tres veces inferior a la de los electrones:

- Los huecos no se mueven por sí mismos sino como resultado de que un electrón ligado, que se encontraba en un enlace, pasa a cubrir el hueco y deja un nuevo hueco en su enlace anterior; de manera que el hueco se mueve en sentido contrario a como lo hacen los electrones ligados que lo rellenan.
- Siempre es más costoso movilizar un electrón ligado (que forma parte de un enlace) que un electrón libre (que se encuentra ya suelto en la banda de conducción).
- A igualdad de dimensiones (a igualdad de factor de forma W/L) la intensidad conducida por un transistor PMOS será muy inferior a la de un transistor NMOS.

Para los transistores PMOS de la tecnología Ψ :

$$\mu_h \sim 200 \text{ cm}^2/\text{V}\cdot\text{s}, \quad c_{ox} \sim 0,7 \text{ fF}/\mu^2 = 0,07 \text{ }\mu\text{F}/\text{cm}^2, \quad K_p \sim 200 \cdot 0,7 \approx 15 \text{ }\mu\text{A}/\text{V}^2.$$

T6.2. Etapa en fuente común



$$V_S = 0$$

zona lineal

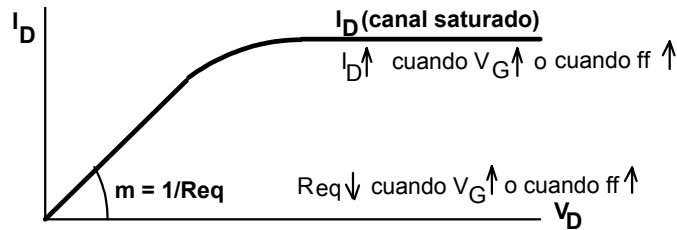
$$R_{eq} = \frac{1}{K_p \frac{W}{L} (V_G - V_{TO})}$$

canal saturado

$$I_{D,sat} = \frac{K_p}{2} \cdot \frac{W}{L} (V_G - V_{TO})^2$$

Para canal no saturado:
$$I_D = K_p \frac{W}{L} \left(V_G - V_{TO} - \frac{V_D}{2} \right) V_D$$

Representando la intensidad que circula por el transistor (drenaje-fuente) I_D en función de la tensión entre sus terminales V_D , se obtiene la curva característica de la conducción del transistor:



Para pequeños valores de V_D , $V_D \ll V_G - V_{TO}$, zona lineal, el transistor se comporta como una resistencia cuyo valor disminuye al aumentar la tensión de puerta y, también, si aumenta el factor de forma ($ff = W/L$):

$$R_{eq} = \frac{1}{K_p \frac{W}{L} (V_G - V_{TO})}$$

Para canal saturado $V_D > V_G - V_{TO}$,
$$I_D = \frac{K_p}{2} \frac{W}{L} (V_G - V_{TO})^2$$

la intensidad que atraviesa el transistor no depende de la tensión de drenaje y aumenta fuertemente con la tensión de puerta y linealmente con el factor de forma.

Actuando sobre las dimensiones geométricas del transistor ($ff = W/L$), se modifica su resistencia en zona lineal y, en sentido contrario, la intensidad con canal saturado.

$$W \uparrow \text{ ó } L \downarrow \text{ (en ambos casos } ff \uparrow) \Rightarrow R_{eq} \downarrow ; I_D \text{ (canal saturado)} \uparrow ;$$

la magnitud de ambas variaciones es proporcional a la modificación del factor de forma.

Valores de los diversos parámetros (tecnología Ψ):

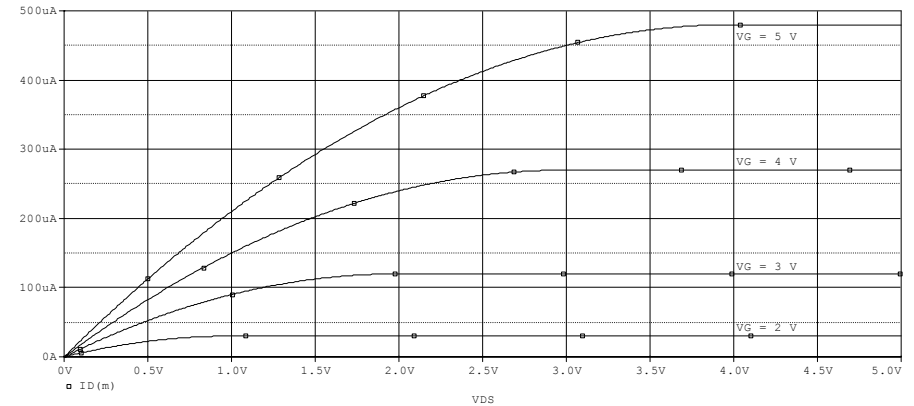
Para un transistor cuadrado ($W=L$; $ff=1$) con tensión umbral de 1 V y tensión de puerta de 5 V los valores de la resistencia equivalente en zona lineal y de la intensidad de canal saturado serán:

NMOS	PMOS
$K_p \approx 40 \mu A/V^2$	$K_p \approx 15 \mu A/V^2$
$R_{eq} = 1 / (40 \cdot 10^{-6} \cdot (5-1)) \approx 6 \text{ k}\Omega$	$R_{eq} = 1 / (15 \cdot 10^{-6} \cdot (5-1)) \approx 17 \text{ k}\Omega$
$I_D = (40/2)(5-1)^2 \approx 0,3 \text{ mA}$	$I_D = (15/2)(5-1)^2 \approx 0,1 \text{ mA}$

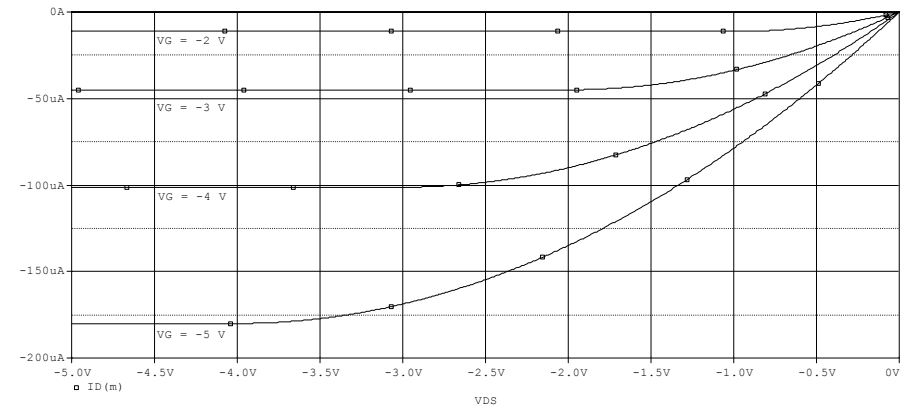
Para un transistor de tamaño mínimo ($ff=1,5$):

$R_{eq} = 1 / (40 \cdot 10^{-6} \cdot 1,5 \cdot (5-1)) \approx 4 \text{ k}\Omega$	$R_{eq} = 1 / (15 \cdot 10^{-6} \cdot 1,5 \cdot (5-1)) \approx 11 \text{ k}\Omega$
$I_D = (40/2) \cdot 1,5 \cdot (5-1)^2 \approx 0,5 \text{ mA}$	$I_D = (15/2) \cdot 1,5 \cdot (5-1)^2 \approx 0,2 \text{ mA}$

Las gráficas siguientes muestran las curvas características de los transistores MOS de tamaño mínimo ($ff = 1,5$) de la tecnología Ψ para diversos valores de la tensión de puerta; han sido obtenidas mediante simulación SPICE para $L = 1 \mu m$, $W = 1,5 \mu m$, $V_{TO,NMOS} = 1 \text{ V}$, $V_{TO,PMOS} = -1 \text{ V}$, $K_{p,NMOS} = 40 \mu A/V^2$ y $K_{p,PMOS} = 15 \mu A/V^2$.



Curvas características $I_D - V_{DS}$ del transistor NMOS para varios valores de V_{GS}



Curvas características $I_D - V_{DS}$ del transistor PMOS para varios valores de V_{GS}

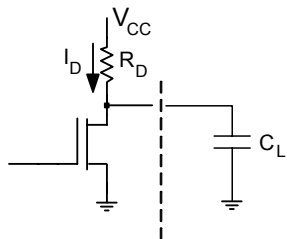
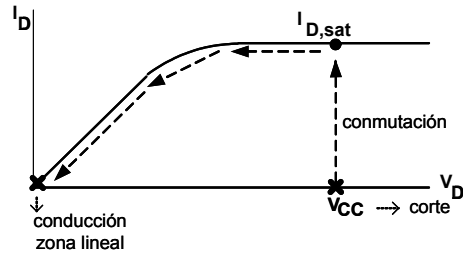
Comportamiento booleano

Las situaciones booleanas de un transistor MOS corresponden a:

- transistor en corte $V_{GS} < V_{TO}$ $I = 0$

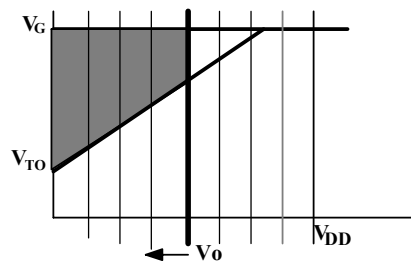
- transistor en zona lineal $V_{GS} > V_{TO}$ $V_{DS} = 0$
$$R_{eq} = \frac{1}{K_p \frac{W}{L} (V_G - V_{TO})}$$

En la conmutación, en la puesta en conducción del transistor, éste recorre la curva característica correspondiente a $V_{GS} = V(1)$ desde la situación de canal saturado hasta alcanzar la zona lineal con $V_{DS} = 0$ V.



La $I_{D,sat}$ señala la máxima intensidad disponible por el transistor, con ella se inicia el proceso de conmutación, para descargar las capacidades equivalentes conectadas a su salida (debidas a otras puertas o dispositivos a los que el transistor comunica el valor booleano 0).

En condiciones funcionales la salida de una etapa se encontrará conectada a una o más etapas siguientes (o a otro tipo de carga), de forma que deberá cargar y descargar la capacidad equivalente que suponen las puertas de dichas etapas: C_L .



Al pasar la tensión de puerta V_G de 0 V a un valor $V(1)$ el transistor conduce y descarga la capacidad C_L que pasa de $V_o = V_{CC}$ (situación anterior, para $V_G = 0$ V) a $V_o = 0$ V: la intensidad, al principio, corresponde a la de canal saturado (área triangular en el diagrama de Memelink) y se reduce a medida que V_o disminuye (área trapezoidal), de forma que la intensidad decrece fuertemente al realizarse la descarga.

En la etapa representada en la figura, al pasar V_G de $V(1)$ a 0 V la capacidad C_L se carga a través de la resistencia R_D : proceso exponencial con constante de tiempo $R_D \cdot C_L$; resulta una situación sin interés respecto a las puertas lógicas CMOS, ya que tales puertas no se construyen con resistencias de carga.

T6.3. El transistor MOS real: efectos de segundo orden

Las principales diferencias de un transistor MOS real con el modelo desarrollado en los apartados anteriores son las siguientes:

1 La placa superior del «condensador» puerta-canal no es metálica, sino semiconductor, construida con silicio policristalino fuertemente dopado N^+ ; ello motiva que la distribución de carga en dicha placa no sea puramente superficial sino volumétrica: se distribuye en la zona superficial del polisilicio abarcando una cierta profundidad no nula. Tal distribución volumétrica de carga introduce un pequeño efecto resistivo, dando lugar a la correspondiente caída de tensión.

El alto dopado del polisilicio ($\sim 10^{20}$ ats/cm³, cien mil veces superior al sustrato, 10^{15}) reduce tal efecto volumétrico, permitiendo despreciarlo respecto al funcionamiento del transistor MOS. En cambio, deberá tenerse en cuenta que la resistividad de las conexiones realizadas en polisilicio es relativamente alta (unas 300 veces la de un metal) y su espesor es inferior al de las líneas de metal ($\sim 0,2$ frente a $0,5 \mu m$); por ello, la resistencia que presenta una línea de polisilicio es bastante superior (resistencia por cuadro: $30 \Omega/\square$ el polisilicio, frente a $50 m\Omega/\square$ en las conexiones metálicas).

2 La tensión umbral V_{TO} se encuentra afectada por los siguientes efectos, todos ellos superficiales:

- la terminación abrupta de la red cristalina del sustrato,
- la existencia de iones positivos, principalmente de sodio, atrapados en el óxido de silicio durante los procesos de oxidación,
- la aparición de cargas fijas en las superficies de contacto entre el silicio y el óxido, que provienen, principalmente, de una oxidación incompleta del silicio.

Estos tres efectos se suman y pueden ser contabilizados conjuntamente en forma de una distribución de carga positiva $Q_{ox} = q \cdot N_{ox}$ que induce sobre el sustrato una carga negativa, es decir, crea una pequeña capa vacía inicial en ausencia de tensión exterior aplicada, reduciendo la tensión umbral V_{TO} (en el caso de un transistor de canal N; aumentándola, si es de canal P):

$$\Delta V_{TO} = - Q_{ox} / C_{ox} = - q \cdot N_{ox} / C_{ox}$$

La concentración equivalente de iones N_{ox} suele ser del orden de:

$$N_{ox} \approx 5 \times 10^{10} \text{ iones / cm}^2 = 500 \text{ iones / } \mu m^2 \quad Q_{ox} \sim 0,08 \text{ fC / } \mu m^2$$

$$\Delta V_{TO} \sim - 0,08 / 0,7 \approx -0,1 \text{ V.}$$

Un cuarto efecto superficial se debe a las interfases o potenciales de contacto polisilicio-óxido y óxido-sustrato y es consecuencia del diferente potencial o energía de extracción de los electrones en las capas en contacto. Tal potencial de contacto origina, asimismo, un desplazamiento de la tensión umbral

$$\Delta V_{TO} = \phi_{\text{sustrato}} - \phi_{\text{polisilicio}} \sim 0,8 \text{ V.}$$

con signo + para el sustrato N, canal P y con signo - para el sustrato P, canal N.

La suma de ambos desplazamientos (- 0,9 canal N y + 0,7 canal P) se traduce en la práctica cancelación de la tensión umbral V_{TO} en el caso del transistor N, mientras que para el transistor P dicha tensión umbral se reduce en forma importante (a 0,3 V).

Ahora bien, estos efectos que desplazan la tensión umbral V_{TO} expresan también la posibilidad de controlar tecnológicamente dicha tensión (en el proceso de fabricación de los transistores) mediante la implantación de iones en la superficie del sustrato.

La implantación de iones, a través del óxido de puerta en la zona semiconductor contigua al mismo, modifica la carga equivalente Q_{ox} y permite ajustar la tensión umbral de los transistores, situándola en valores adecuados que suelen estar en el entorno de 1 V en los transistores NMOS y en valores un poco inferiores (para compensar la menor movilidad de sus portadores) en los transistores PMOS.

Mediante la implantación de iones negativos (Bor^-) se aumenta la tensión umbral del transistor N, acercándola a 1 V y, recíprocamente, la implantación de iones positivos ($Fósforo^+$ o $Arsenio^+$) permite situar la tensión umbral del transistor P en valores absolutos ligeramente superiores a 0,6 voltios. De esta forma, se utiliza la desigualdad $V_{TO} (PMOS) < V_{TO} (NMOS)$ para compensar un poco la asimetría entre ambos transistores, causada por la menor movilidad de los huecos.

Todos estos efectos superficiales inciden en el valor de la tensión umbral V_{TO} pero no afectan al modelo de transistor desarrollado anteriormente.

3 La reducción de las dimensiones del transistor MOS (longitud L o anchura W) afecta también a la tensión umbral V_{TO} , en forma de desplazamiento de la misma.

En las proximidades de las zonas de difusión (fuente o drenaje, ricas en portadores) el potencial de extracción de portadores es inferior al correspondiente al semiconductor que forma el sustrato. Por ello, cuando la longitud del canal es corta ($L \sim 1 \mu m$) la tensión umbral resulta afectada por la presencia de las difusiones (cercanas a todo el canal) y disminuye en forma apreciable, tanto más cuanto menor es la longitud del canal.

Por otra parte, la distribución de carga espacial de la zona vacía se extiende (en el sentido de la anchura del transistor) más allá de la longitud transversal definida por el óxido de puerta (más allá de la zona de puerta determinada por el polisilicio al cruzar a la difusión). El efecto de esta ampliación de la capa vacía es apreciable cuando el transistor es muy estrecho ($W < 1 \mu m$) y puede contabilizarse como una distribución de carga de vaciamiento Q_V superior a la considerada; como dicho término actúa en términos de resta, respecto a la tensión de puerta, se produce un aumento de la tensión umbral.

Estos dos efectos son de signo contrario y se compensan parcialmente entre sí, supuesto que ambas dimensiones L y W sean del mismo orden de tamaño. Sin embargo, en tecnologías «submicrónicas» ($< 1 \mu m$) el comportamiento real de los transistores difiere en gran medida del teórico (de las ecuaciones funcionales desarrolladas anteriormente), debido a este tipo de efectos relacionados con las dimensiones del transistor (influencia relativa de los bordes del transistor sobre su globalidad).

4 La longitud del transistor MOS no corresponde exactamente al tamaño transversal (anchura) del polisilicio que le cruza ya que la difusión penetra ligeramente debajo de la zona de puerta, por sus bordes, ocupando una franja de anchura L_D (difusión lateral); de forma que la longitud efectiva del canal será $L - 2L_D$.

Además, cuando el canal se encuentra saturado, se cierra antes de alcanzar al drenaje y su longitud efectiva es menor que la separación entre las dos difusiones (fuente y drenaje). Este efecto aumenta al hacerlo la tensión de drenaje y da lugar a que las curvas $I_D - V_{DS}$ no sean planas en la zona de canal saturado, sino que presenten una ligera pendiente; se conoce con el nombre de modulación de la longitud del canal y se contabiliza mediante un parámetro λ :

$$L_{\text{efectiva}} = (L - 2L_D) / (1 + \lambda V_{DS}).$$

En ocasiones, por analogía con los transistores bipolares, se utiliza como parámetro para expresar este efecto de aumento de I_D con V_D en la zona de canal saturado la denominada *tensión early* que corresponde a la tensión en que la prolongación de la recta I_D correspondiente a canal saturado corta al eje de tensiones.

5 La movilidad de los portadores μ_e μ_h y, por consiguiente, el parámetro de transconductancia K_p , depende fuertemente de la temperatura y de los campos eléctricos.

La temperatura disminuye la movilidad de los portadores y en el mismo sentido actúa el campo vertical producido por la tensión de puerta V_G ya que tiende a impulsar a los portadores hacia la superficie (interfase óxido-sustrato). En ambos casos aumenta el efecto de dispersión (los choques) que produce la red cristalina sobre los portadores (colisiones portadores - núcleos de la red); es el mismo efecto que sufren los materiales conductores al aumentar la temperatura y que da lugar a un aumento de su resistencia.

También afecta negativamente a la movilidad de los portadores el aumento del campo eléctrico horizontal producido por la diferencia de tensión entre fuente y drenaje; para valores de campo suficientemente intensos ($E > E_{\text{crítico}}$) se produce una saturación de la velocidad de los portadores: existe una velocidad máxima por encima de la cual la aceleración producida por el campo eléctrico no tiene efecto.

Por otra parte, el parámetro μ (movilidad de los portadores), además de ser sumamente sensible a todo tipo de magnitudes físicas, resulta de muy difícil medida; por ello, en muchas ocasiones, se utiliza directamente el parámetro de transconductancia $K_p = \mu \cdot C_{ox}$, cuya dependencia de la temperatura y de los campos eléctricos se establece mediante coeficientes de tipo experimental.

6 La distribución de tensión $V(x)$ a lo largo del canal (debida a las tensiones de fuente y drenaje) no solo actúa reduciendo la tensión efectiva de cada elemento Δx del condensador puerta-canal, $V_G - V_{TO} - V(x)$; también incrementa la carga de la capa vacía Q_V , reduciendo en igual cantidad la carga de la capa de inversión Q_I y su contribución a la corriente que circula por ella. Se denomina *efecto sustrato* al que producen las tensiones aplicadas horizontalmente (V_S, V_D) sobre la carga de la capa de vaciamiento.

Debido a la distribución de tensión que inducen V_S y V_D , la tensión en la superficie de un elemento $\Delta(x)$ será $\phi_B + V' > \phi_B$ y como la carga de la capa de vaciamiento es $Q_V = K_1 \sqrt{\phi_B + V'}$ (apartado T6.1.a), la carga de la capa de inversión será:

$$Q_I = Q - Q_V = c_{ox} (V_G - \phi_B - V') - K_1 \sqrt{\phi_B + V'}$$

$$= c_{ox} (V_G - \phi_B - V' - \gamma \sqrt{\phi_B + V'}) = c_{ox} (V_G - (\phi_B + V' + \gamma \sqrt{\phi_B + V'}))$$

Como $V_{TO} = \phi_B + \gamma \sqrt{\phi_B}$, $\phi_B = V_{TO} - \gamma \sqrt{\phi_B}$, podemos escribir

$$Q_I = c_{ox} (V_G - (\phi_B + V' + \gamma \sqrt{\phi_B + V'}))$$

$$= c_{ox} (V_G - (V_{TO} - \gamma \sqrt{\phi_B} + V' + \gamma \sqrt{\phi_B + V'}))$$

$$= c_{ox} (V_G - (V_{TO} + V' + \gamma \sqrt{\phi_B + V'} - \gamma \sqrt{\phi_B}))$$

y, con ello, la recta $V_{TO} + V$ del *diagrama de Memelink* se transforma en un curva de mayor pendiente: $V_{TO} + V + \gamma \sqrt{\phi_B + V} - \gamma \sqrt{\phi_B}$.

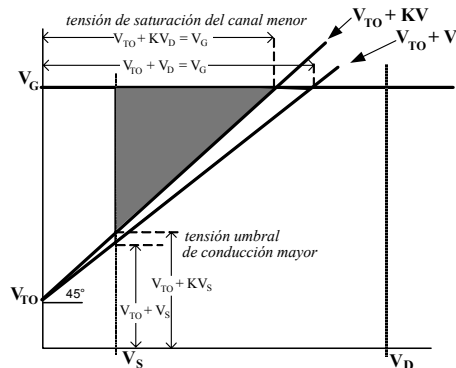
Resulta útil emplear una aproximación lineal de esta curva; para ello, en la expresión $Q_I = c_{ox} (V_G - \phi_B - V' - \gamma \sqrt{\phi_B + V'})$, el término de «efecto substrato» $\gamma \sqrt{\phi_B + V'}$ puede aproximarse mediante su desarrollo en serie de Taylor (primer término de la serie):

$$\gamma \sqrt{\phi_B + V'} = \gamma \sqrt{\phi_B} + aV' ; a = \frac{d}{dV'} (\gamma \sqrt{\phi_B + V'}) \Big|_{V'=0} = \frac{\gamma}{2\sqrt{\phi_B}}$$

$$\gamma \sqrt{\phi_B + V'} - \gamma \sqrt{\phi_B} = aV'$$

$$Q_I = c_{ox} (V_G - (V_{TO} + V' + aV')) = c_{ox} (V_G - (V_{TO} + KV')) \quad \text{con } K = 1 + a \approx 1,15$$

$I = \mu \cdot c_{ox} \cdot \frac{W}{L} \cdot \int_{V_S}^{V_D} (V_G - (V_{TO} + K \cdot V(x))) \cdot dV$: la recta $V_{TO} + V$ del *diagrama de Memelink* pasa a $V_{TO} + K \cdot V$ ($K \sim 1,15 > 1$), es decir, aumenta su pendiente:



Habida cuenta de que la recta $V_{TO} + V$ actúa substractivamente, las consecuencias de su mayor pendiente son las siguientes:

➤ la intensidad que conduce el transistor es menor:

$$\begin{cases} I_D = K_p \frac{W}{L} \left(V_G - V_{TO} - K \frac{V_S + V_D}{2} \right) \cdot V_{DS} & \text{canal no saturado} \\ I_D = \frac{K_p}{2K} \frac{W}{L} (V_G - V_{TO} - KV_S)^2 & \text{canal saturado} \end{cases}$$

➤ cuando la tensión de fuente V_S es mayor que 0 V la tensión umbral aumenta: para $V_S > 0$ V el transistor no conducirá hasta que

$$V_G \geq V_{TO} + K \cdot V_S = V_{TO} + V_S + a \cdot V_S ; V_{GS} \geq V_T = V_{TO} + a \cdot V_S;$$

siendo la nueva tensión umbral $V_T = V_{TO} + a \cdot V_S$ ($a \sim 0,15$).

Este efecto tiene importancia cuando la fuente del transistor no se encuentra unida al terminal de polarización (por ejemplo, en cuando los transistores se encuentran en serie) y reduce en forma apreciable la conductividad del transistor.

➤ la saturación del canal se produce para tensiones de drenaje inferiores:

la saturación se produce cuando $V_{TO} + K \cdot V_D = V_G$

o sea, para una tensión de drenaje $V_D = (V_G - V_{TO}) / K$

$$V_D \sim 3.5 V_G \text{ para } V_G = 5 V \text{ y } V_{TO} = 1 V$$

y dicho valor limita la tensión máxima de salida de un transistor de paso (como se verá en su momento).

7 Al analizar el comportamiento del transistor MOS se han detallado dos situaciones excluyentes: ausencia/presencia de canal (capa de inversión). Las condiciones establecidas para que exista canal ($V_G > V_T$) corresponden a una «inversión fuerte»: concentración de portadores en el canal análogo al dopado del substrato ($n \sim N_A$).

Existen otras dos situaciones posibles: inversión débil $n \ll N_A$ e inversión moderada $n < N_A$ que es preciso tener en cuenta cuando interesa el comportamiento del transistor para tensiones de puerta en el entorno de la tensión umbral V_T ; en principio, tales situaciones no afectan al caso digital pero han de ser tenidas en cuenta en etapas de tipo analógico.

T6.4. Modelos SPICE

El modelo de nivel 1 (modelo de Shichman – Hodges) corresponde a las funciones desarrolladas en el apartado T6.1:

$$\begin{cases} I_D = K_p \frac{W}{L} \left(V_{GS} - V_{TO} - \frac{V_{DS}}{2} \right) V_{DS} & \text{canal no saturado} \\ I_D = \frac{K_p}{2} (V_{GS} - V_{TO})^2 & \text{canal saturado} \end{cases}$$

En este modelo, la descripción de un transistor MOS requiere tres parámetros:

KP	K_p parámetro de transconductancia
VTO	V_{TO} tensión umbral del transistor
TOX	t_{ox} espesor del óxido de puerta

TOX es necesario para tener en cuenta la capacidad de puerta; caso de no utilizarlo, se supone capacidad de puerta nula.

En lugar de KP puede utilizarse como parámetro la movilidad de los portadores μ , calculándose KP a partir del valor de dicha movilidad UO:

$$UO \quad K_p = \mu c_{ox} = \mu \frac{\epsilon_{SiO_2}}{t_{ox}}; \quad \mu = UO$$

En todo caso, es necesario indicar la dimensión de los transistores **L**, **W**, bien dentro del modelo o en la declaración de cada transistor; por defecto se asigna a L y W el valor de 1 metro, con lo cual se configuran transistores enormes, de 1 m² (que son tan grandes que dan lugar a resultados de simulación erróneos).

El modelo de nivel 1 admite los siguientes efectos de segundo orden:

- disminución de la longitud del canal debida a la difusión lateral **LD**: $L' = L - 2LD$
- modulación de la longitud de canal saturado **LAMBDA**: $L_{efectiva} = L / (1 + \lambda V_{DS})$
- efecto substrato: la recta $V_{TO} + V$ pasa a ser $V_{TO} + V + \gamma \sqrt{\phi_B + V} - \gamma \sqrt{\phi_B}$
que requiere dos parámetros

GAMMA	γ coeficiente de efecto substrato
PHI	ϕ_B potencial superficial para inversión fuerte,
o, en su defecto:	NSUB N_A dopado del substrato:

$$\phi_B = 2 \frac{KT}{q} \ln \frac{N_A}{N_i}$$

Ejemplo de modelo de transistores MOS (nivel 1):

- a) MODEL "nombre" NMOS LEVEL = 1
+ KP = 40U VTO = 1 TOX = 50E-9 L = 1U W = 1.5U
+ LD = 0.05E-6 LAMBDA = 0.01 GAMMA = 0.25 NSUB = 1E15
- b) MODEL "nombre" PMOS LEVEL = 1
+ KP = 15U VTO = -1 TOX = 50E-9 L = 1U W = 1.5U
+ LD = 0.05E-6 LAMBDA = 0.01 GAMMA = 0.25 NSUB = 5E15

Además, pueden incluirse los efectos resistivos y capacitivos de las regiones físicas que componen el transistor mediante los correspondientes parámetros: **RSH** (R_{\square} resistencia por cuadro –cuadrado– de la difusión), **CJ** (capacidad de la difusión respecto al substrato), **CJSW** (capacidad perimetral de la difusión respecto al substrato), **CGSO** y **CGDO** (capacidades de solapamiento puerta-fuente y puerta-drenaje), **RS**, **RD** y **RG** (resistencias óhmicas de los contactos con fuente, substrato y drenaje), etc.

El modelo de nivel 2 permite incluir todos los efectos de segundo orden comentados en el apartado T6.3 y, en particular, los relativos a los transistores submicrónicos; utiliza, en principio, los mismos parámetros que el de nivel 1, a los que añade parámetros complementarios para contabilizar los efectos sobre la movilidad de los portadores μ (y, por tanto, sobre K_p), sobre la tensión umbral V_T , etc. y para incluir situaciones de inversión débil. Los diversos efectos de segundo orden se tienen en cuenta o no según se asignen o no valores a los correspondientes parámetros.

El modelo de nivel 3 corresponde a un ajuste semiempírico de coeficientes a través de aproximaciones de las características del transistor obtenidas a través de la medida de las mismas. Ha sido desarrollado principalmente para aproximar los resultados de la simulación a los valores experimentales en los transistores de canal ultracorto.

Además de los anteriores, existen otros modelos SPICE de los transistores MOS: versiones avanzadas de SPICE suelen incluir una docena de tales modelos. Sin embargo, en muchas ocasiones, en particular en la simulación de circuitos digitales con transistores no muy pequeños ($L \geq 1\mu m$), basta el modelo de nivel 1 para conseguir resultados muy aceptables.