

Design of Analog CMOS Integrated Circuits

Francesco Gallone

Fundación Fulgor

19 de septiembre de 2025

Índice

1. Introducción	3
2. Marco Teórico	4
2.1. Estructura	4
2.2. Funcionamiento	4
2.3. Simbología	4
3. Zonas de operación	5
3.1. Voltaje Umbral	5
3.2. Zona del triodo	5
3.3. Región de Saturación	6
3.4. Trasconductancia en MOS	7
4. Efectos de Segundo orden	8
4.1. Body Effect	8
4.2. Modulación de longitud de canal	8
4.3. Conducción subumbral	9
5. Teoría de Escalado	9
5.1. Efectos de canal corto	9
6. Cálculos problema 2.1	11
6.1. Análisis para la comparativa NMOS/PMOS	11
7. Cálculos problema 2.2	12
7.1. LTSpice	13

1. Introducción

Repasaremos conceptos básicos de la estructura de dispositivos MOS, con sus respectivas regiones operativas más comunes y las principales diferencias entre ellas.

2. Marco Teórico

2.1. Estructura

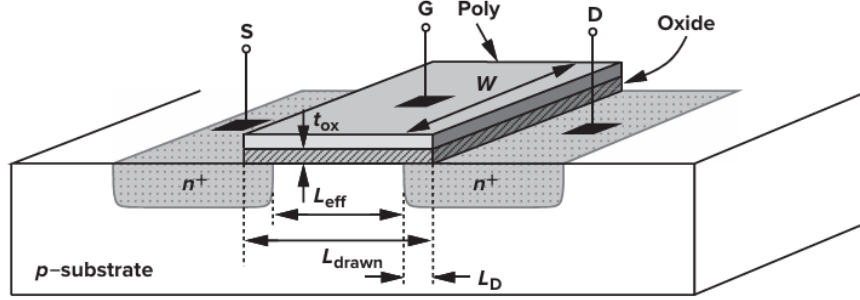


Figura 1: Estructura de un nMOS, un MOS tipo N.

La estructura del dispositivo se fabrica en un sustrato tipo p, el dispositivo consiste en dos regiones tipo n fuertemente dopadas que forman los terminales *source* y *drain*, una pieza de polisilicio fuertemente dopada llamada "*poly*" que trabaja como una *gate*, y una capa delgada de dióxido de Silicio (SiO₂) llamada simplemente "óxido" la cual tiene de función aislar la puerta del sustrato, la principal acción ocurre en la región bajo el óxido de la *gate*.

La distancia entre los terminales *source* y *drain* es llamada la Longitud, L y perpendicularmente a la longitud tenemos el ancho, W. Debido a que la junturas S/D tienen bordes difusos por el proceso de fabricación, la longitud real entre la *source* y el *drain* es ligeramente menor que la longitud total. Teniendo L_D como la longitud de la difusión, se puede conocer el valor de L_{eff} :

$$L_{eff} = L - 2L_D$$

2.2. Funcionamiento

Para entender de forma práctica el funcionamiento general, podemos definir a *source* como el terminal que provee portadores de carga (electrones en este caso) y *drain* como el terminal que los recolecta. El potencial del sustrato sobre el que dispositivo es fabricado influencia en gran medida las características del dispositivo. Por ello, es que el MOSFET es un dispositivo de *cuatro terminales*. Debido a que la juntura S/D de los diodos debe estar en inversa es que asumimos que los transistores NMOS están conectados a la fuente más negativa del sistema.

2.3. Simbología

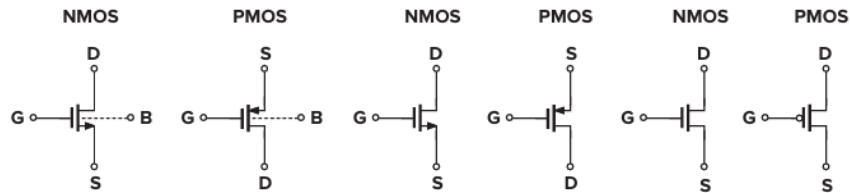


Figura 2: Simbología MOS

3. Zonas de operación

3.1. Voltaje Umbral

Cuando el voltaje de *gate* V_G incrementa desde cero, el dielectrico y el sustrato forman un capacitor, como V_G se vuelve más positivo, las lagunas del sustrato p son repelidas del área de la gate dejando iones negativos detrás como una carga reflejada en la gate. En otras palabras se crea una región de agotamiento, por la cual no hay flujo de corriente ya que no hay portadores de carga disponibles. Así como V_G incrementa, tambien lo hace el ancho de la región de agotamiento y el potencial en la interfaz oxido-silicio. La estructura se vuelve similar a un divisor de voltaje formado por dos capacitores en serie: el capactio de óxido de la gate y el capacitor de la region de agotamiento. Cuando el potencial de la interfaz alcanza un valor suficientemente positivo ($V_G = V_{TH}$), los electrones fluyen desde source a la interfaz y eventualmente al drain. De esta manera, se forma un canal de portadores de carga debajo del oxido de la gate, encendiendo el transistor.

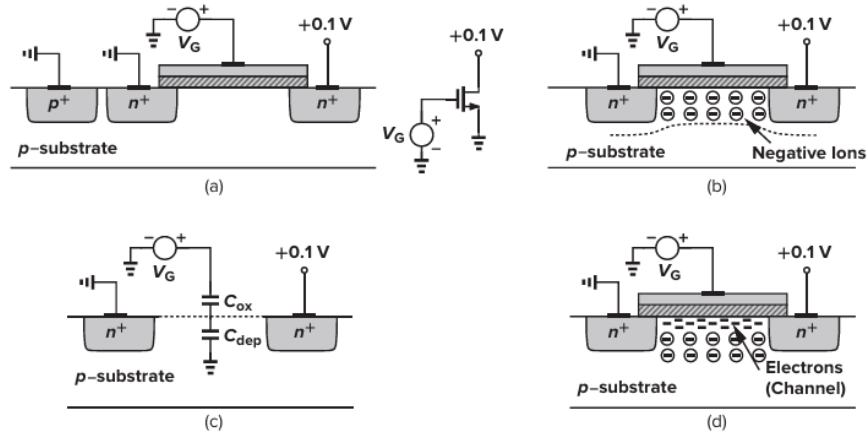


Figura 3: (a) MOSFET activado por tension de gate; (b) formación de la región de agotamiento ;(c) inicio de la inversión;(d) formación de la capa de inversión

3.2. Zona del triodo

Llamamos a $V_{GS} - V_{TH}$ como voltaje de sobreexcitación y W/L como relación de aspecto. Si se cumple $V_{DS} \leq V_{GS} - V_{TH}$ entonces el transistor opera en la región del triodo.

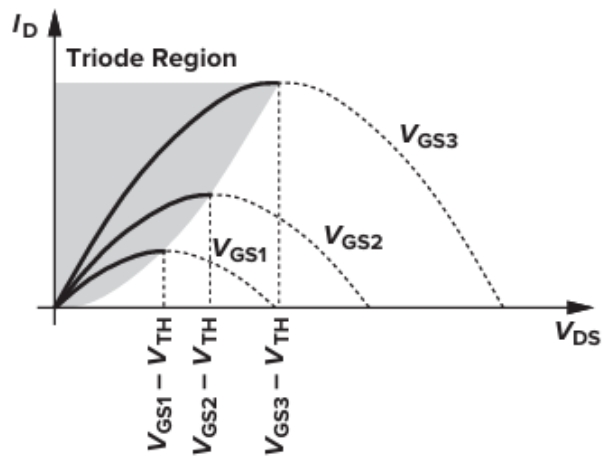


Figura 4

La corriente I_D responde a la ecuación:

$$I_{D, max} = \frac{1}{2} \mu_n C_{ox} \frac{W}{L} (V_{GS} - V_{TH})^2$$

Figura 5

Esta última describe la dependencia de I_D respecto a la constante $\mu_n C_{ox}$, las dimensiones del dispositivo, W y L , y los potenciales de gate y drain con respecto a source.

3.3. Región de Saturación

¿Qué ocurre si $V_{DS} > V_{GS} - V_{TH}$? Se podría creer que la corriente de drain I_D seguiría el comportamiento de la parábola pero en realidad lo que ocurre es que se vuelve relativamente constante, por lo que decimos que el dispositivo opera en la *Región de "Saturación"*. Para comprender este fenómeno debemos recordar que la densidad local de al carga de la capa de inversión es proporcional a $V_{GS} - V(x) - V_{TH}$. Si $V(x)$ se aproxima a $V_{GS} - V_{TH}$ entonces, $Q_d(x)$ cae a cero. En síntesis si V_{DS} es ligeramente mayor que $V_{GS} - V_{TH}$ la capa de inversión se detiene en $x \leq L$, y decimos que el canal está *estrangulado* (*pinch-off*).

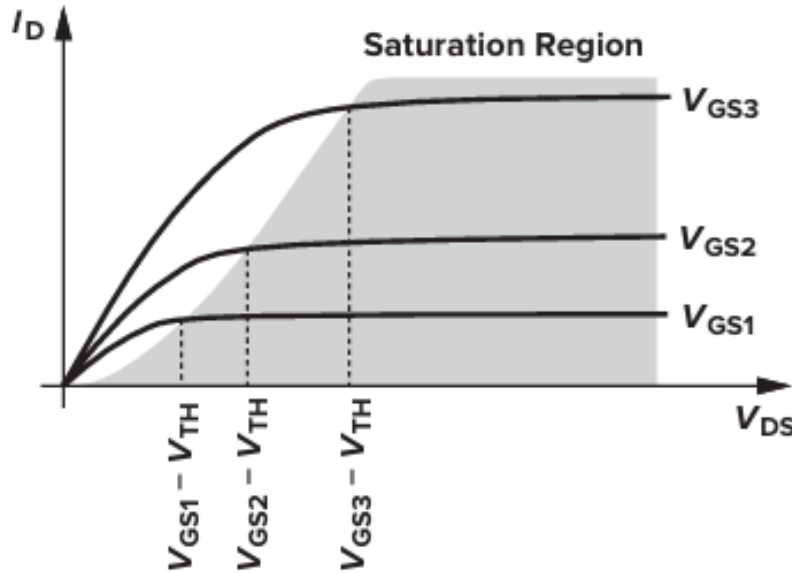


Figura 6

¿Cómo es que el dispositivo conduce corriente en presencia de este estrangulamiento? A medida que los electrones se acercan al *pinch-off point* (donde $Q_d \rightarrow 0$) su velocidad aumenta tremendamente ($v = I/Q_d$). Una vez que pasa ese punto, los electrones simplemente atraviesan la región de agotamiento cerca de la juntura de drain y llegan a la terminal drain.

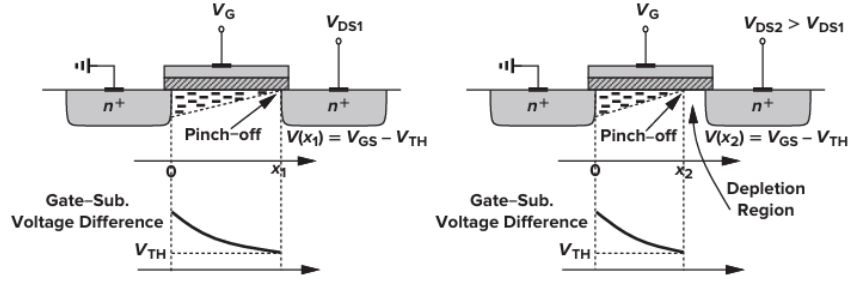


Figura 7

3.4. Trasconductancia en MOS

La trasconductancia g_m es una figura de mérito que indica que tan bien un dispositivo convierte un voltaje en una corriente. Es decir, representa la *sensibilidad*.

$$g_m = \left. \frac{\partial I_D}{\partial V_{GS}} \right|_{V_{DS} \text{ const.}}$$

$$= \mu_n C_{ox} \frac{W}{L} (V_{GS} - V_{TH})$$

Figura 8

Como se puede apreciar, para un g_m alto, un pequeño cambio en V_{GS} genera un gran cambio en I_D . Expresamos g_m en $1/\Omega$ o en *Siemens*. En diseño analógico decimos que un MOSFET opera como un *V/I Converter* para indicar que convierte un cambio de voltaje en un cambio de corriente. g_m también puede ser expresada de la forma:

$$g_m = \sqrt{2\mu_n C_{ox} \frac{W}{L} I_D}$$

$$= \frac{2I_D}{V_{GS} - V_{TH}}$$

Figura 9

En esta última ecuación, cada una de las expresiones resulta útil para estudiar el comportamiento de g_m en función de un parámetro mientras que los otros parámetros permanecen constantes.

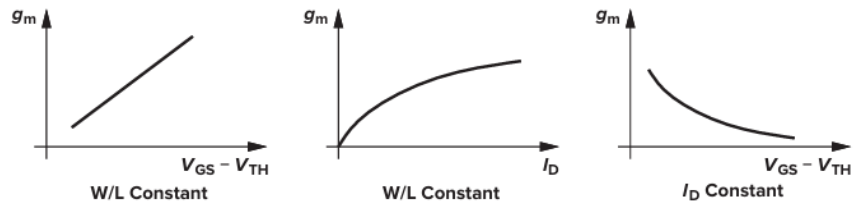


Figura 10

4. Efectos de Segundo orden

En nuestro análisis de la estructura MOS realizamos varias simplificaciones, algunas de las cuales no son validad en muchos circuitos analógicos. En esta sección vamos a describir tres de ellos que son esenciales para el subsecuente análisis de circuitos.

4.1. Body Effect

En nuestro análisis asumimos que la masa y source del transistor estaban ancladas a tierra. ¿Qué ocurre si el voltaje de la masa en un NFET cae por debajo del voltaje de la ecuación de g_m ? Cuando el voltaje es mayor que cero, se incrementa la región de agotamiento bajo el canal, esto requiere más voltaje en la compuerta para invertir el canal, lo que eleva el voltaje umbral V_{TH}

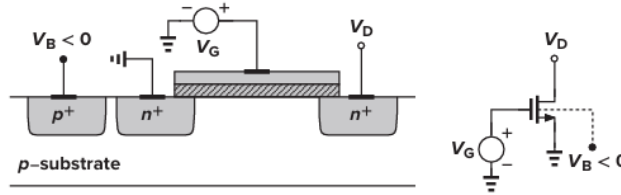


Figure 2.22 NMOS device with negative bulk voltage.

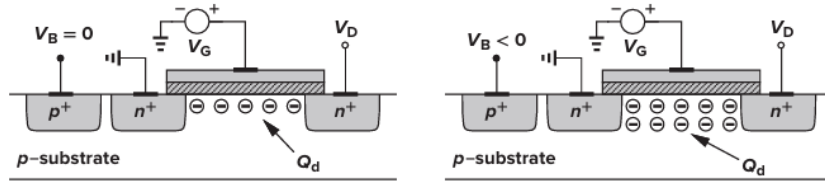


Figure 2.23 Variation of depletion region charge with bulk voltage.

Figura 11

4.2. Modulación de longitud de canal

En el análisis de pinch-off observamos que la longitud del canal decrece a medida que el diferencial de potencial entre la gate y drain decrece, este efecto se conoce como *Modulación de longitud de canal*.

$$I_D \approx \frac{1}{2} \mu_n C_{ox} \frac{W}{L} (V_{GS} - V_{TH})^2 (1 + \lambda V_{DS})$$

Figura 12

En donde λ es el *coeficiente* de modulación de longitud de canal. Este parámetro representa la variación relativa en longitud por un incremento en V_{DS} dado. Así, para canales más largos, λ es más pequeño.

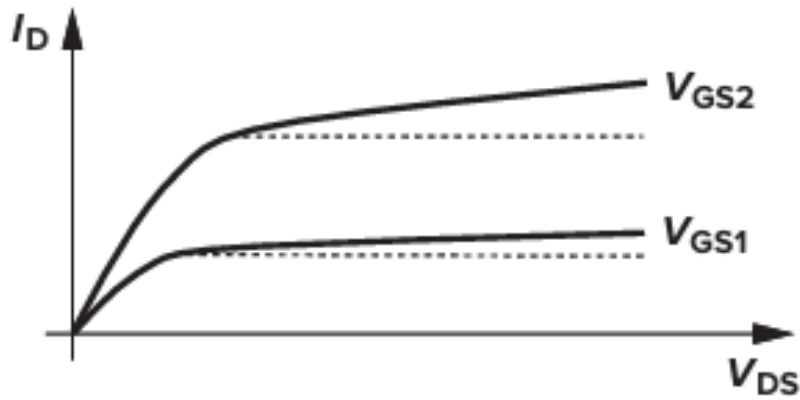


Figura 13

4.3. Conducción subumbral

En nuestro análisis del MOSFET asumimos que el dispositivo se apaga abruptamente cuando V_{GS} cae por debajo de V_{TH} . Pero, en realidad para $V_{GS} \approx V_{TH}$ sigue habiendo una capa de inversión débil que permite un pequeño flujo de corriente de D a S. Incluso para $V_{GS} < V_{TH}$, I_D es finito, pero exhibe una dependencia exponencial a V_{GS} . Esta corriente es llamada *corriente subumbral*. Esto ocurre porque los portadores todavía pueden moverse por *difusión*, en lugar de conducción.

$$I_D = I_0 \exp \frac{V_{GS}}{\epsilon V_T}$$

5. Teoría de Escalado

La tecnología CMOS domina la industria de semiconductores por dos principales razones, la no disipación de potencia estática en la logica CMOS y la escalabilidad de los MOSFETs. En un paper publicado en 1974 *Design of Ion-Implanted MOSFET's with Very Small Physical Dimensions* se introdujo formalmente la *teoría de escalado* o también conocida *Dennard Scaling*, en la que el Dennard reconoció el tremendo potencial de escalado de transistores MOS, haciendo predicciones sobre su velocidad y disipación de potencia de circuitos CMOS digitales a medida que los dispositivos se reducen.

La teoría de escalado sigue tres reglas:

- 1- Reducir las dimensiones laterales y verticales por un factor α (≥ 1).
- 2- Reducir la tensión umbral y la tensión de alimentación. Esto mantiene las proporciones eléctricas evitando campos eléctricos excesivos (escalado de cmapo constante).
- 3- Aumentar la concentración de niveles de dopaje por un factor α .

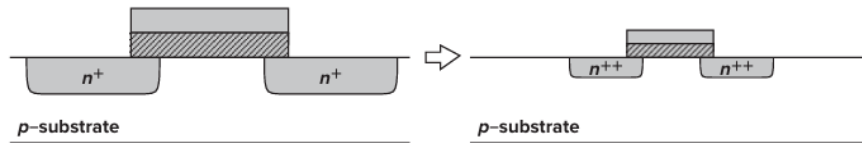


Figura 14: Escalado ideal de transistor MOS

5.1. Efectos de canal corto

los efectos de geometría pequeña surgen por cinco factores que desvian el escalado del escenario ideal:

1 - Los campos eléctricos tienden a aumentar porque el voltaje de alimentación no se ha escalado proporcionalmente.

2 - El término de potencial interno en la ecuación para el ancho total de la region de agotamiento no es escalable ni tampoco despreciable

$$W = \sqrt{\frac{2\epsilon_s N_A + N_D}{q N_A N_D}} (V_{bi} + V) \quad (1)$$

3- La profundidad de las uniones S/D no se puede escalar fácilmente.

4- La movilidad disminuye a medida que aumenta el dopaje del sustrato

5- La pendiente subumbral no es escalable

6. Cálculos problema 2.1

Table 2.1 Level 1 SPICE models for NMOS and PMOS devices.

NMOS Model			
LEVEL = 1	VTO = 0.7	GAMMA = 0.45	PHI = 0.9
NSUB = 9e+14	LD = 0.08e-6	UO = 350	LAMBDA = 0.1
TOX = 9e-9	PB = 0.9	CJ = 0.56e-3	CJSW = 0.35e-11
MJ = 0.45	MJSW = 0.2	CGDO = 0.4e-9	JS = 1.0e-8
PMOS Model			
LEVEL = 1	VTO = -0.8	GAMMA = 0.4	PHI = 0.8
NSUB = 5e+14	LD = 0.09e-6	UO = 100	LAMBDA = 0.2
TOX = 9e-9	PB = 0.9	CJ = 0.94e-3	CJSW = 0.32e-11
MJ = 0.5	MJSW = 0.3	CGDO = 0.3e-9	JS = 0.5e-8

Figura 15

2.1. For $W/L = 50/0.5$, plot the drain current of an NFET and a PFET as a function of $-V_{GS}$ as $-V_{GS}$ varies from 0 to 3 V. Assume that $-V_{DS} = 3$ V.

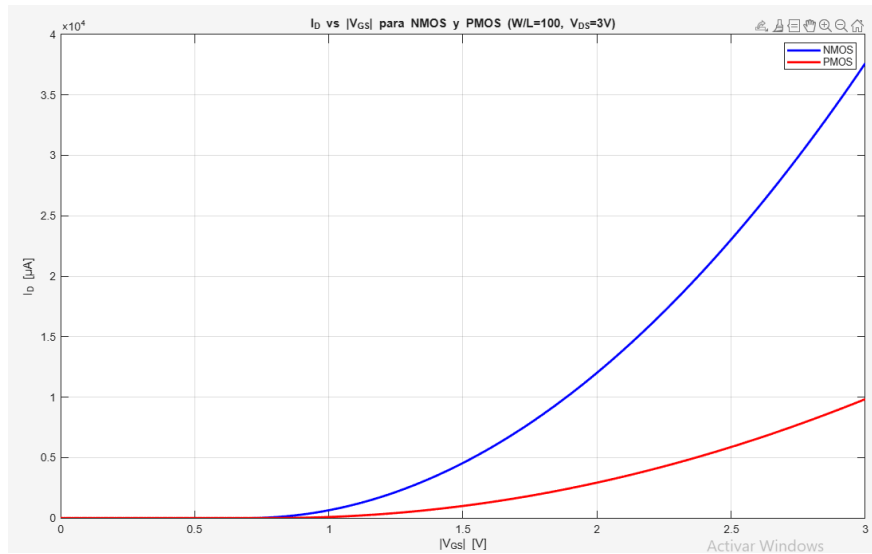


Figura 16

6.1. Análisis para la comparativa NMOS/PMOS

El NMOS no conduce hasta que $V_{GS} > 0.7$, a partir de ahí la corriente crece rápido porque su movilidad es bastante mas alta que la del PMOS, por eso para el mismo V_{GS} el NMOS conduce mas corriente que el PMOS. Para el PMOS recién parece comenzar a conducir para $V_{GS} > 0.8$. La pendiente es menor ya que la movilidad de huecos es menor que la de electrones. Eso hace que su I_D sea bastante menor que la del NMOS para un mismo V_{GS} . Para corregir este desbalance en un inversor CMOS suele hacerse el PMOS más ancho (w más grande) que el NMOS. Eso compensa la movilidad más baja de los huecos y ayuda a que los tiempos de subida y bajada de la señal sean similares. Consigna 1

$L_{eff,nmos} = L - 2L_{Dn} = 0,5\mu m - 2(0,08\mu m) = 0,340\mu m$ El valor de la longitud de canal total siempre va a ser mayor por la fabricación "side-difussion". Una L_{eff} más corta implica un canal más débilmente controlado por la gate, por lo que se incrementa el efecto de canal corto

Consigna 2

$$L_{eff,pmos} = L - 2L_{Dn} = 0,5\mu m - 2(0,09\mu m) = 0,320\mu m$$

Consigna 3

$$\frac{W}{L_{eff,PMOS}} = \frac{50\mu m}{0,340\mu m} = 147,06$$

Consigna 4

$$\frac{W}{L_{eff,PMOS}} = \frac{50\mu m}{0,320\mu m} = 156,25$$

Esta relación determina su capacidad de conducción relativa.

Consigna 5

$C_{ox} = \frac{\epsilon_{ox}}{t_{ox}} = \frac{3,9\epsilon_0}{9 \times 10^{-9}} = 3836,73 \mu F/m^2$ Este parámetro depende del espesor del óxido y de la constante dieléctrica del material, a mayor valor mejor control del canal por parte de gate, por lo que mejora g_m Consigna 6

$$1/2\mu_n C_{ox} \frac{W}{L_{eff}} (1 + \lambda V_{DS}) = 1/2(350 \times 10^{-4})(C_{ox})(147,06)(1 + 0,1 * 3) = 12,84 mA/V^2$$

El NMOS muestra mejor ganancia de corriente en saturación ya que tiene mayor movilidad de los electrones a comparación de los huecos del PMOS Consigna 7

$$1/2\mu_n C_{ox} \frac{W}{L_{eff}} (1 + \lambda V_{DS}) = 1/2(100 \times 10^{-4})(C_{ox})(156,25)(1 + 0,2 * 3) = 4,80 mA/V^2$$

Consigna 8

$$I_{D,n} = K_n (V_{GS} - V_{Tn})^2 = (12,84 \times 10^{-3})(3 - 0,7)^2 = 67,90 mA$$

Consigna 9

$$I_{D,p} = K_p (V_{GS} - V_{Tp})^2 = (4,80 \times 10^{-3})(3 - 0,8)^2 = 23,21 mA$$

7. Cálculos problema 2.2

Resultados para $W/L = 50/0,5y I_D = 0,5mA$:

Consigna 10

$$\frac{W}{L_{NMOS}} = \frac{50\mu m}{0,5\mu m} = 100$$

$$\frac{W}{L_{PMOS}} = \frac{50\mu m}{0,5\mu m} = 100$$

Consigna 11

$$C_{ox} = \frac{\epsilon_{ox}}{t_{ox}} = \frac{3,9\epsilon_0}{9 \times 10^{-9}} = 3836,73 \mu F/m^2$$

Consigna 12

$$g_{m,NMOS} = \mu C_{ox} \frac{W}{L} (V_{GS} - V_{Tn}) = 350 \times 10^{-4} C_{ox} 100 (V_{GS} - 0,7) = 1,97 mA/V$$

Consigna 13

$$g_{m,PMOS} = \mu C_{ox} \frac{W}{L} (V_{GS} - V_{Tp}) = 100 \times 10^{-4} C_{ox} 100 (V_{GS} - 0,8) = 0,99 mA/V$$

Estos valores de transconductancia nos dicen la capacidad de los transistores para convertir un cambio de voltaje en la gate en un cambio de corriente en el drain, mientras mayor sea mejor es la conversión, por los valores vistos de NMOS y PMOS podemos ver que NMOS tiene mayor movilidad de electrones que PMOS y por lo tanto genera más corriente por cada voltio aplicado en comparación a PMOS Consigna 14

$$r_{o,NMOS} = \frac{1}{\lambda_n I_D} = \frac{1}{0,1 * 0,5 * 10^{-3}} = 20K\Omega$$

Consigna 15

$$r_{o,NMOS} = \frac{1}{\lambda_n I_D} = \frac{1}{0,2 * 0,5 * 10^{-3}} = 10K\Omega$$

Se puede representar en circuitos la modulación de longitud de canal con un resistor r_o esta representa la variación de la corriente de drenaje ante cambios en V_{DS} debido a la modulación de longitud de canal. Y λ como constante de modulación de longitud de canal. Un r_o alto implica que la corriente I_D dependa menos de V_{DS} , o sea mejor linealidad y mayor capacidad de amplificación del transistor. A mayor λ la curva I_D sube rápido con V_{DS} -> Mayor dependencia de V_{DS} por lo que hay menor resistencia de salida, reflejando menor ganancia intrínseca. Consigna 16

Ganancia intrínseca NMOS: $g_m * r_o = 1,97 * 10^{-3} * 20 * 10^3 = 39,4(A\Omega)/V$

Consigna 17

Ganancia intrínseca NMOS: $g_m * r_o = 0,99 * 10^{-3} * 10 * 10^3 = 9,9(A\Omega)/V$

7.1. LTSpice

From the simulated traces determine the value of the collector current at $V_1=10V$ and $I_1=100\mu A$, and enter it in the answer box below in units of mA.

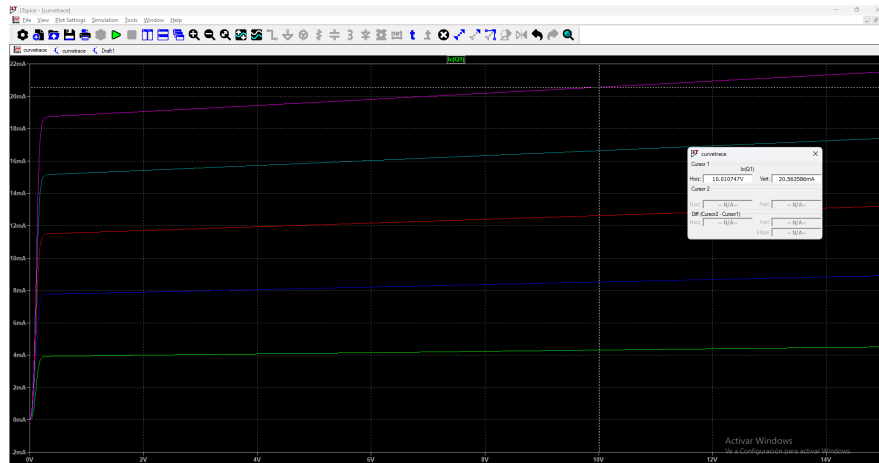


Figura 17

Corriente de colector:

$$I_C = 20,56mA$$