Lecture-2

Date: 04.08.2016

- NMOS I/V Characteristics
- Discussion on I/V Characteristics
- MOSFET – Second Order Effect

NMOS I-V Characteristics

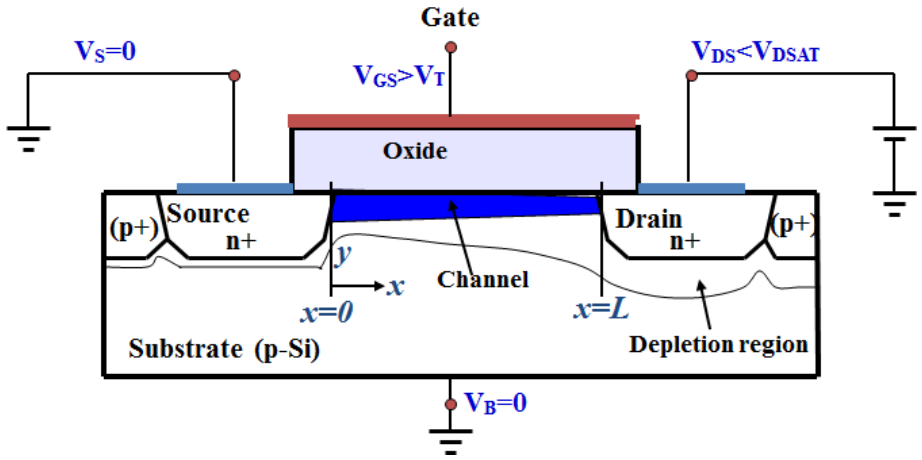
Gradual Channel Approximation: **Cut-off → Linear/Triode → Pinch-off/Saturation**

Assumptions: $V_{SB} = 0$; V_T is constant along the channel ; E_x dominates E_y → need to consider current flow only in the x −direction

- **Cutoff Mode:** $0 \le V_{GS} \le V_T$; $I_{DS(cutoff)} = 0$

This relationship is simple—if MOSFET is in **cutoff**, drain current is simply **zero** !

- **Linear Mode:** $V_{GS} \ge V_T, 0 \le V_{DS} \le V_{D(SAT)} \rightarrow V_{DS} \le V_{GS} - V_T$
- The channel reaches the drain.

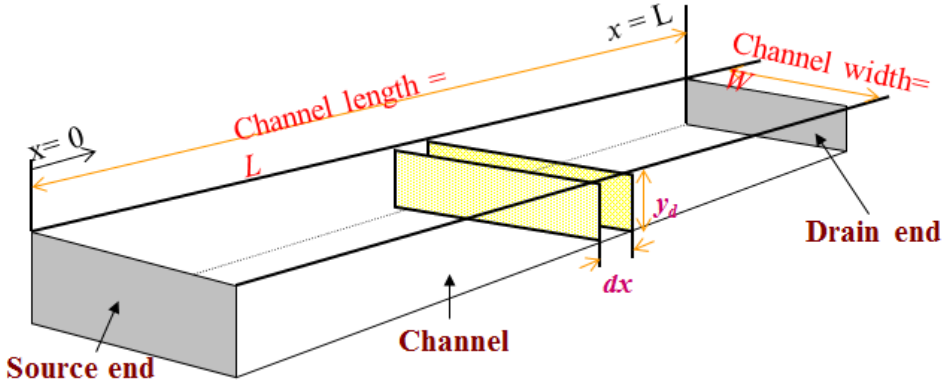


- $V_c(x)$: Channel voltage with respect to the source at position x.
- Boundary Conditions:
  $V_c(x = 0) = V_S = 0$; $V_c(x = L) = V_{DS}$

Linear Mode (Contd.)

Q_d: the charge density along the direction of current $= WC_{ox}[V_{GS} - V_T]$

where, **W** = width of the channel and **WC_{ox}** is the capacitance per unit length



- The channel potential varies from **0 at source** to **V_{DS} at the drain:**

$Q_d(x) = WC_{ox}[(V_{GS} - V_c(x)) - V_T]$,
where, $V_c(x)$ = channel potential at x.

- **Subsequently we can write:** $I_D(x) = Q_d(x).v$,
where, v = velocity of charge (m/s)

$v = \mu_n E$; where, μ_n = mobility of charge carriers (electron)

E = electric field in the channel given by: $E(x) = -\frac{dV}{dx}$

Therefore, $I_D(x) = WC_{ox}\left[V_{GS} - V_c(x) - V_T\right]\mu_n \frac{dV}{dx}$

Linear Mode (Contd.)

- Applying the boundary conditions for $V_c(x)$ we can write:

$$I_D(x) = I_D = \int_{x=0}^{x=L} I_D.dx = \int_{V=0}^{V=V_{DS}} WC_{ox}\left[V_{GS} - V(x) - V_T\right]\mu_n.dV$$

- **Simplification gives the drain current in linear mode as:**

$$I_D = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}\left[2(V_{GS} - V_T)V_{DS} - V_{DS}^2\right]$$

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

The $I_{D,max}$ occurs at, $V_{DS} = V_{GS}-V_T$, **[how/why?]** called overdrive voltage

Home Assignment # 0

Observations:

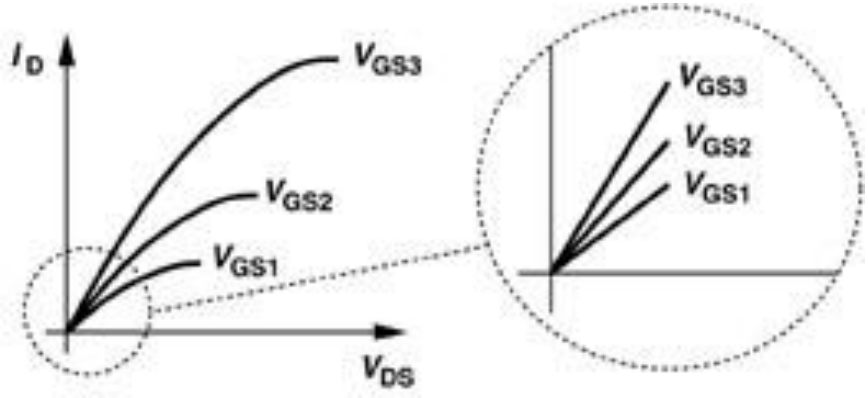
- I_D is dependent on constant of technology ($\mu_n C_{ox}$), the device dimensions (W and L), and the gate and drain potentials with respect to the source
- For $V_{DS} << 2(V_{GS}-V_T)$: $I_D \approx \mu_n C_{ox}\frac{W}{L}(V_{GS} - V_T)V_{DS}$ ⟵ Linear function of V_{DS}

Linear Mode (Contd.)

For small values of V_{DS}, the drain current can be thought of as a straight line → implying that the path from source to drain can be represented by a linear resistor → support of earlier assumption

$$R_{DS} = \frac{V_{DS}}{I_D} = \frac{1}{\mu_n C_{ox} \frac{W}{L}(V_{GS} - V_T)}$$

MOSFET transistor operate as a resistor whose value is controlled by overdrive voltage



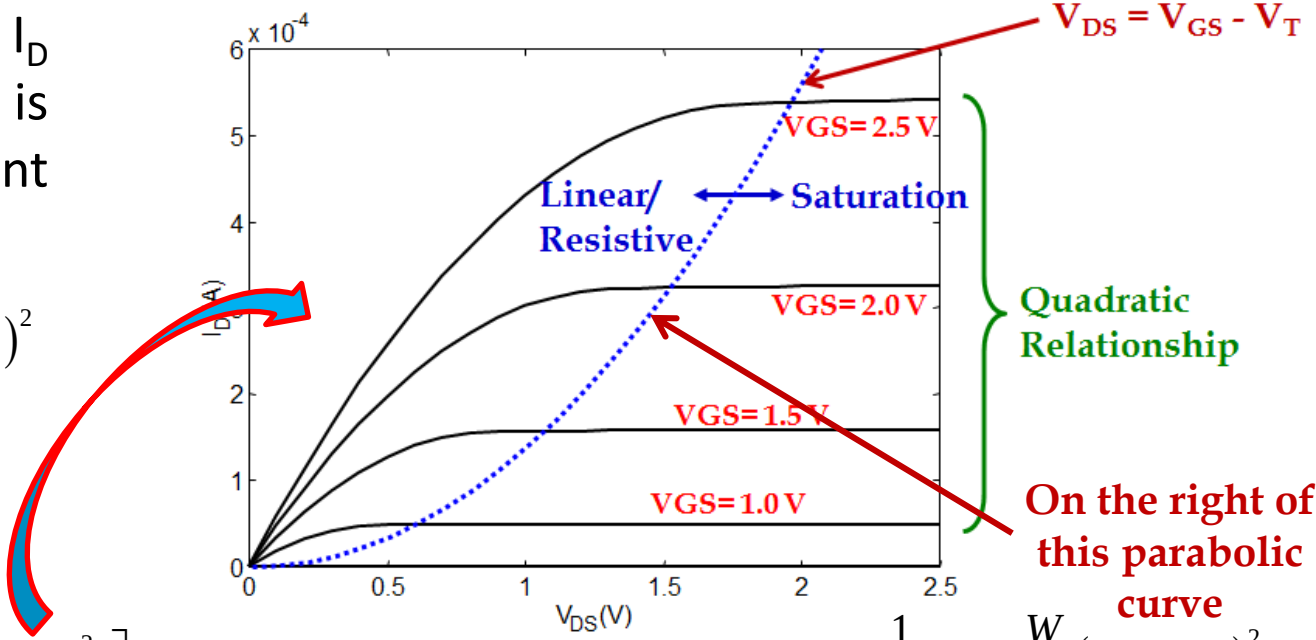
Pinch-off point (Edge of Saturation): $V_{GS} \geq V_T,\ V_{DS} = V_{D(SAT)}$

- The channel just reaches the drain but with zero inversion at the drain
- Electrons start to drift from the channel to the drain
- The drain current is given by: $I_D = I_{D,\max} = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}\left(V_{GS} - V_T\right)^2$

Saturation Mode: $V_{GS} \geq V_T,\ V_{DS} \geq V_{GS} - V_T$

- After pinch-off, the I_D saturates i.e, is relatively independent of V_{DS}:

$$I_{D,sat} = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}\left(V_{GS} - V_T\right)^2$$

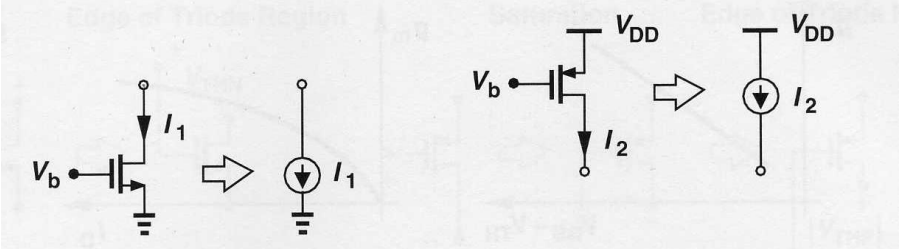


$$I_D = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}\left[2\left(V_{GS} - V_T\right)V_{DS} - V_{DS}^2\right]$$

$$I_{D,sat} = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}\left(V_{GS} - V_T\right)^2$$

Saturation Mode (Contd.)

- MOSFET can be used as current source connected between the drain and the source, controlled by V_{GS}.



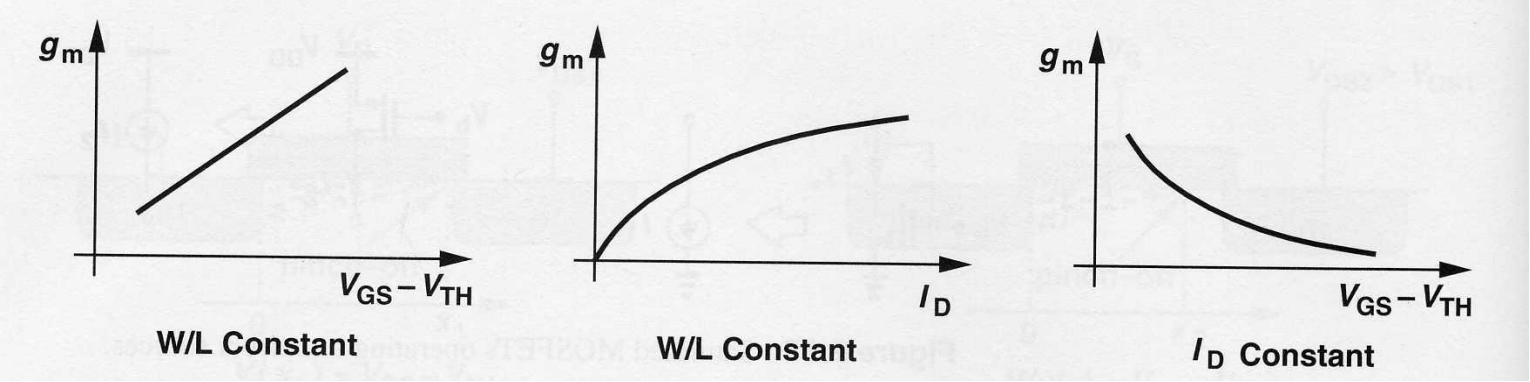
- MOSFET in saturation mode → produces a current regulated by V_{GS} → imperative to define a figure of merit (FOM) that identifies the effectiveness with which the MOSFET can convert voltages in currents → the FOM in this scenario is called "transcodunctance (g_m)".
- Defined as the change in the drain current divided by the change in the gate-source voltage.

$$g_m = \left.\frac{\partial I_D}{\partial V_{GS}}\right|_{V_{DS},const.} = \mu_n C_{ox}\frac{W}{L}(V_{GS}-V_T)$$

- In essence, g_m represents the sensitivity of the device. For a high g_m, a small change in V_{GS} results in a large change in I_D.
- Other formulations of g_m: $g_m = \sqrt{2.\mu_n C_{ox}\frac{W}{L}I_D}$ ⇒ $g_m = \frac{2I_D}{V_{GS}-V_T}$

Transconductance (g_m)

- Behavior of g_m as a function of one parameter while other parameters remain fixed.



Home Assignment # 0

Can we define g_m in the triode/linear region?

Channel Resistance for Small V_{DS}

- Voltage V_{DS} will be **directly proportional** to I_D, provided that:
  1. A conducting channel has been **induced**.
  2. The value of $\boldsymbol{V_{DS}}$ **is small**.

Note for this situation, the MOSFET will be in **triode** region.

Channel Resistance for Small V_{DS}

→ As we **increase** the value of V_{DS}, the conducting channel will begin to **pinch off**—the current will **no longer** be directly proportional to V_{DS}.

- Specifically, there are **two phenomena** at work as we **increase V_{DS}** while in the **triode** region:
  1. Increasing V_{DS} will increase the potential difference across the conducting channel → leads to proportional increase in I_D.
  2. Increasing V_{DS} will decrease the conductivity of the induced channel → leads to decrease in I_D.

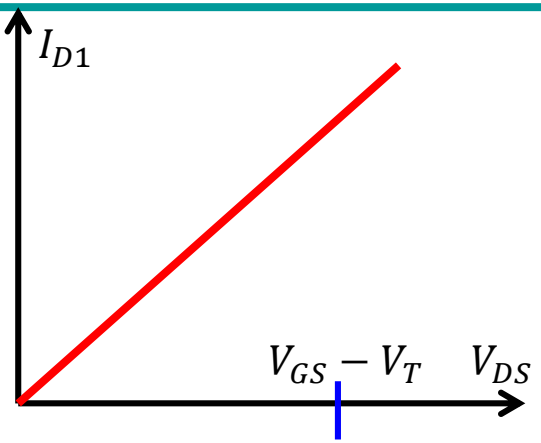
- There are **two** physical phenomena at work as we increase V_{DS}, and there are **two** terms in the triode drain current equation!

$$I_D = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}\left[2(V_{GS}-V_T)V_{DS}-V_{DS}^2\right] \quad \Rightarrow \quad I_D = \mu_n C_{ox}\frac{W}{L}(V_{GS}-V_T)V_{DS}-\frac{1}{2}\mu_n C_{ox}\frac{W}{L}V_{DS}^2$$

$$I_D = I_{D1} + I_{D2}$$

Where: $I_{D1} = \mu_n C_{ox}\frac{W}{L}(V_{GS}-V_T)V_{DS}$ $\qquad$ $I_{D2} = -\frac{1}{2}\mu_n C_{ox}\frac{W}{L}V_{DS}^2$

$$I_{D1} = \mu_n C_{ox} \frac{W}{L}(V_{GS} - V_T)V_{DS}$$



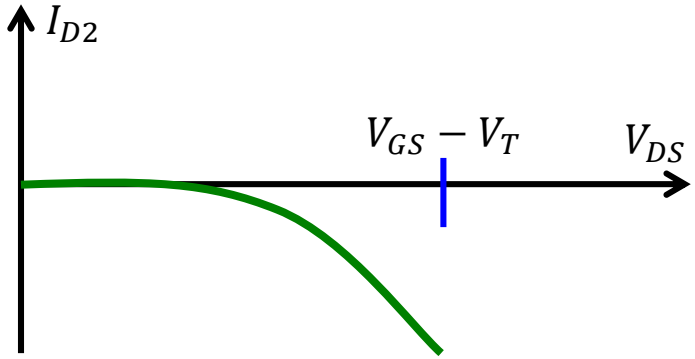
→ It is evident that this term describes the **first** phenomenon:

1. Increasing V_{DS} will increase the potential difference across the conducting channel → leads to proportional increase in I_D.

In other words, this first term would accurately describe the relationship between I_D and V_{DS} **if** the MOSFET induced channel behaved like a **resistor**! ↔ it means the second term doesn't allow it to behave like a perfect resistor.

$$I_{D2} = -\frac{1}{2}\mu_n C_{ox} \frac{W}{L} V_{DS}^2$$

It is apparent that I_{D2} is proportional to V_{DS} **squared**!!

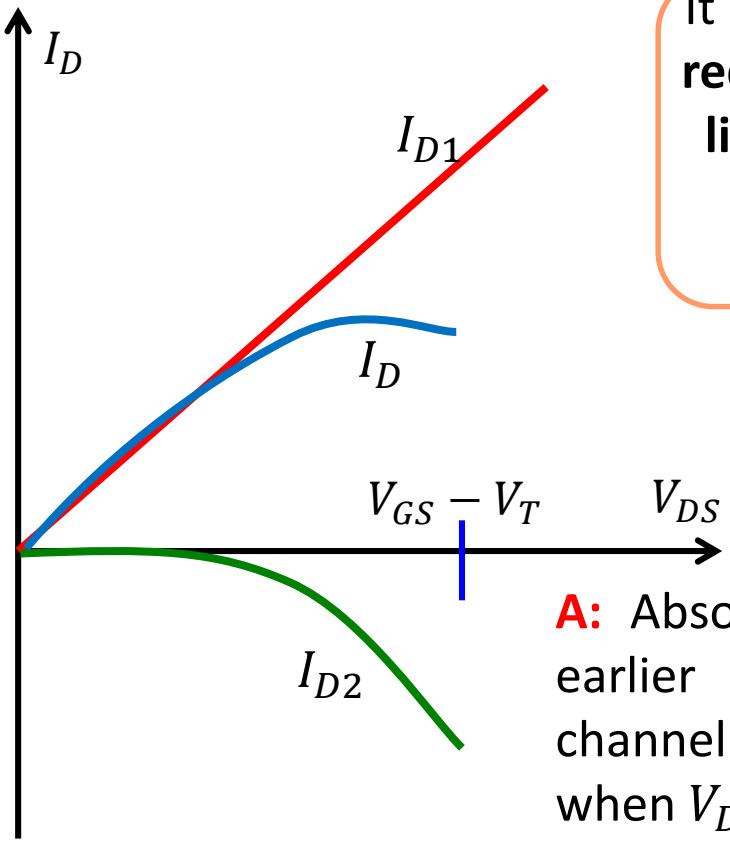


Moreover, the minus sign means that as V_{DS} increases, I_{D2} will actually **decrease**! This behavior is **nothing** like a resistor—what the heck is going on here??

→ This **second term** essentially describes the result of the **second** phenomena:

2. Increasing V_{DS} will decrease the conductivity of the induced channel → leads to decrease in I_D.

- Now let's **add** the two terms I_{D1} and I_{D2} together to get the **total triode drain current** I_D:



It is apparent that the second term I_{D2} works to **reduce** the total drain current from its "**resistor-like**" value I_{D1}. This of course is physically due to the **reduction in channel conductivity** as V_{DS} increases.

Q: But look! It appears to me that for **small** values of V_{DS}, the term I_{D2} is **very small**, and thus $I_D \approx I_{D1}$ (**when** V_{DS} is small)!

A: Absolutely **true**! Recall this is **consistent** with our earlier discussion about the induced channel—the channel conductivity begins to significantly **degrade** only when V_{DS} becomes **sufficiently large**!

Channel Resistance (contd.)

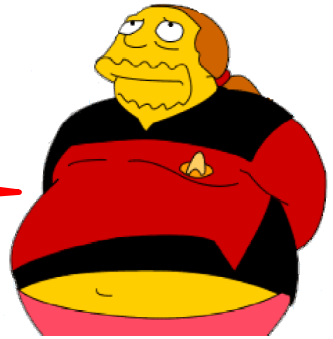
- Thus, we can conclude: $I_D \approx I_{D1} = \mu_n C_{ox} \frac{W}{L}(V_{GS} - V_T)V_{DS}$ For small V_{DS}

- Moreover, we can (for small V_{DS}) approximate the induced channel as a resistor R_{DS} of value $R_{DS} = \frac{V_{DS}}{I_D}$:

$$R_{DS} = \frac{1}{\mu_n C_{ox} \frac{W}{L}(V_{GS} - V_T)}$$

For small V_{DS}

Q: I've just about had it with this "for small V_{DS}" nonsense! Just **how** small **is** small? How can we know **numerically** when this approximation is valid?



A: Well, we **can** say that this approximation is valid when I_{D2} is much smaller than I_{D1} (i.e., I_{D2} is **insignificant**).

- **Mathematically**, we can state as: $|I_{D2}| \ll |I_{D1}|$

$$\frac{1}{2}\mu_n C_{ox} \frac{W}{L} V_{DS}^2 \ll \mu_n C_{ox} \frac{W}{L}(V_{GS} - V_T)V_{DS} \quad \Rightarrow \quad V_{DS} \ll 2(V_{GS} - V_T)$$

Channel Resistance (contd.)

Thus, we can **approximate** the induced channel as a **resistor** R_{DS} when V_{DS} is **much less** than the **twice the excess gate voltage**.

$$R_{DS} = \frac{1}{\mu_n C_{ox} \frac{W}{L}(V_{GS} - V_T)} \qquad \text{For} \qquad V_{DS} \ll 2(V_{GS} - V_T)$$

Q: There you go **again!** The statement $V_{DS} \ll 2(V_{GS} - V_T)$ is only **slightly** more helpful than the statement "when V_{DS} is small". Precisely how **much** smaller than **twice the excess gate voltage** must V_{DS} be in order for our approximation to be **accurate**?

A: We cannot say **precisely** how much smaller V_{DS} needs to be in relation to $2(V_{GS} - V_T)$ unless we state **precisely** how **accurate** we require our approximation to be!

- For example, if we want the **error** associated with the approximation $I_D \approx I_{D1}$ to be **less than 10%**, we find that we require the voltage V_{DS} to be **less than 1/10** the value $2(V_{GS} - V_T)$.

In other words, **if**: $V_{DS} < \frac{2(V_{GS} - V_T)}{10} = \frac{V_{GS} - V_T}{5}$

- we find then that I_{D2} is less than 10% of I_{D1}: $I_{D2} < \frac{I_{D1}}{10}$

Channel Resistance (contd.)

This **10% error criteria** is a **typical** "rule-of thumb" for many approximations in electronics. However, this does **not** mean that it is the "correct" criteria for determining the validity of this (or other) approximation.

- For some applications, we might require **better** accuracy. For **example**, if we require less than **5% error**, we would find that:

$$V_{DS} < \frac{V_{GS} - V_T}{10}$$

It is important to note that we should use these approximations when we can—it can make our **circuit analysis much easier**!

See, the thing is, you should use these approximations whenever they are **valid**. They often make your **circuit analysis** task **much** simpler

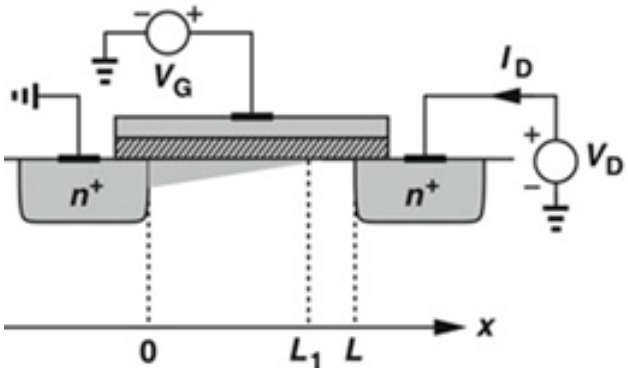
Second Order Effect - Channel Length Modulation

- I have been saying that for a MOSFET in **saturation**, the drain current is **independent** of the drain-to-source voltage V_{DS} i.e.

$$I_D = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{GS} - V_T)^2$$

In reality, this is only **approximately** true!

- Let us look at operation of NMOS in saturation mode:



Observations:

- The pinch-off point moves towards the source with the increase in V_{DS}
- Channel length reduces
- Channel resistance decreases

This modulation of channel length (L) by V_{DS} is known as **channel-length modulation**, and leads to slight dependence of I_D on V_{DS}.

- The drain current in saturation mode is: $I_{D,sat} = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{GS} - V_T)^2$

L actually varies with V_{DS}

The decrease in channel length with increase in V_{DS} essentially increases the drain current I_D

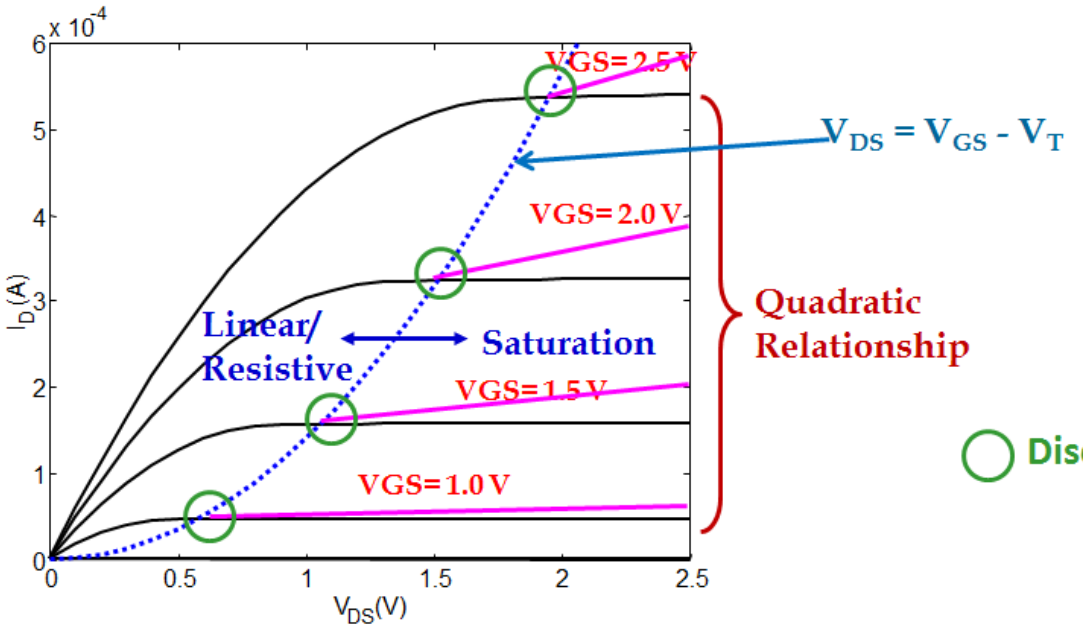
Second Order Effect - Channel Length Modulation (contd.)

If $\Delta L = L - L_1$ then: $$\frac{1}{L_1} = \frac{1}{L-\Delta L} = \frac{1}{L}.\frac{1}{1-\frac{\Delta L}{L}} = \frac{1}{L}.\frac{1}{1-\lambda V_{DS}} \approx \frac{1}{L}.(1+\lambda V_{DS})$$

λ: channel length modulation coefficient (usually less than 0.1)

- Therefore the drain current in saturation mode becomes:

$$I_{D,sat} = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{GS}-V_T)^2(1+\lambda V_{DS})$$

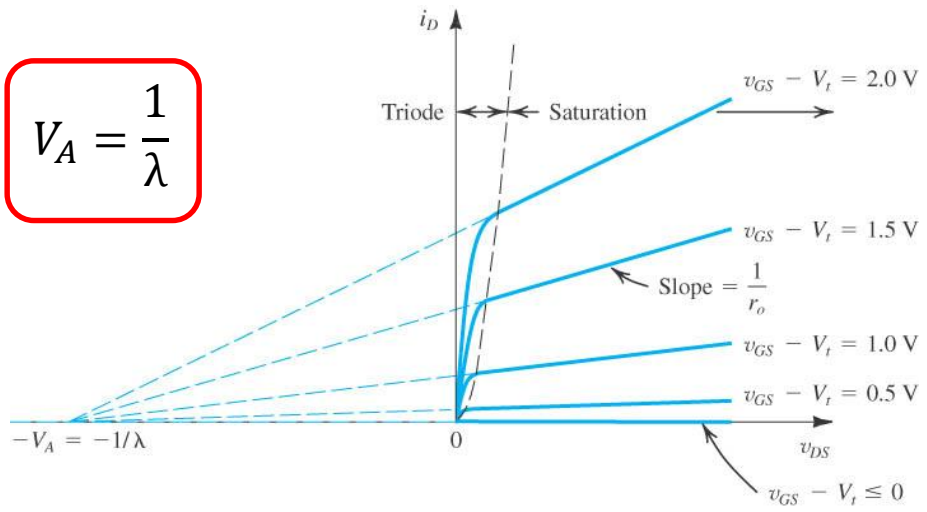


λ is a MOSFET **device parameter** with units of 1/V (i.e., V^{-1}). Typically, this value is small (thus the dependence on V_{DS} is slight), ranging from 0.005 to 0.02 V^{-1}.

Second Order Effect - Channel Length Modulation (contd.)

- Often, the channel-length modulation parameter λ is expressed as the **Early Voltage** V_A, which is simply the inverse value of λ.
- The parameter V_A is set at the time of fabrication and hence the circuit designers can't alter it at circuit/system design stage.

$$V_A = \frac{1}{\lambda}$$



- The drain current for a MOSFET in **saturation** can **likewise** be expressed as:

$$I_D = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{GS}-V_T)^2(1+\lambda V_{DS}) \quad \Rightarrow \quad I_D = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{GS}-V_T)^2\left(1+\frac{V_{DS}}{V_A}\right)$$

- Now, let's **define** a value I_{DI}, which is simply the drain current in saturation **if** no channel-length modulation actually occurred—in other words, the **ideal** value of the drain current:

$$I_{DI} = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{GS}-V_T)^2$$

Second Order Effect - Channel Length Modulation (contd.)

- Thus, we can **alternatively** write the drain current in saturation as:

$$I_D = I_{DI}\left(1+\frac{V_{DS}}{V_A}\right)$$

This **explicitly** shows how the drain current behaves as a function of voltage V_{DS}.

$$I_D = I_{DI}\left(1+\frac{V_{DS}}{V_A}\right)$$

We can interpret the value V_{DS}/V_A as the **percent increase** in drain current I_D over its ideal (i.e., no channel length modulation) saturation value

- Now, let's introduce a **third** way (i.e. in addition to, λ and V_A) to describe the "extra" current created by channel-length modulation. Define the **Drain Output Resistance** r_o:

$$r_o = \frac{1}{\lambda I_{DI}} = \frac{V_A}{I_{DI}}$$

- Using this definition, we can write the **saturation** drain current expression as:

$$I_D = I_{DI}\left(1+\frac{V_{DS}}{V_A}\right) \quad \Rightarrow \quad I_D = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{GS}-V_T)+\frac{V_{DS}}{r_o}$$

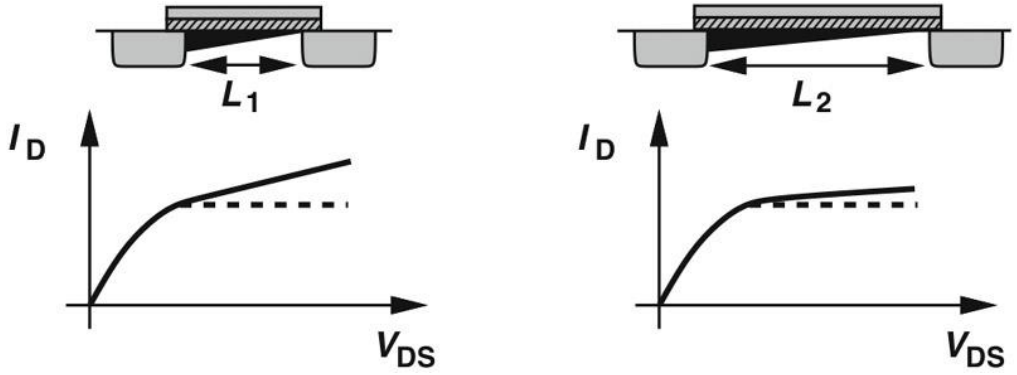
Second Order Effect - Channel Length Modulation (contd.)

Thus, we **interpret** the "extra" drain current (due to channel length modulation) as the current flowing through a **drain output resistor** r_o.

Finally, there are **three** important things to remember about channel-length modulation:

- The values λ and V_A **are** MOSFET **device parameters**, but drain output resistance r_o is **not** (r_o is dependent on I_{DI}).
- Often, we "**neglect** the effect of channel-length modulation", meaning that we use the **ideal** case for saturation: $I_D = I_{DI} = \mu_n C_{ox}(V_{GS} - V_T)^2$. Effectively, we assume that λ = 0, meaning that $V_A = \infty$ and $r_o = \infty$ (i.e., **not** $V_A = 0$ and $r_o = 0$).
- The drain output resistance r_o is **not** the same as channel resistance R_{DS}. The two are different in **many, many** ways.

Second Order Effect - Channel Length Modulation (contd.)



- For a longer channel length, the relative change in L and therefore in I_D for a given change in V_{DS} is smaller.
- To minimize channel length modulation, smaller length transistors should be avoided.

Q: Any idea about limitation of long channel devices?

Home Assignment # 0

Second Order Effect – Body Effect

- In discrete circuit usually there is no body effect as the body is connected to the source terminal.
- In integrated circuit, there are thousands or millions of MOSFET **source terminals** and there is only one Body (B) – the silicon Substrate.
- Thus, if we were to tie (connect) **all** the MOSFET source terminals to the single body terminal, we would be connecting **all** the MOSFET source terminals to each other!

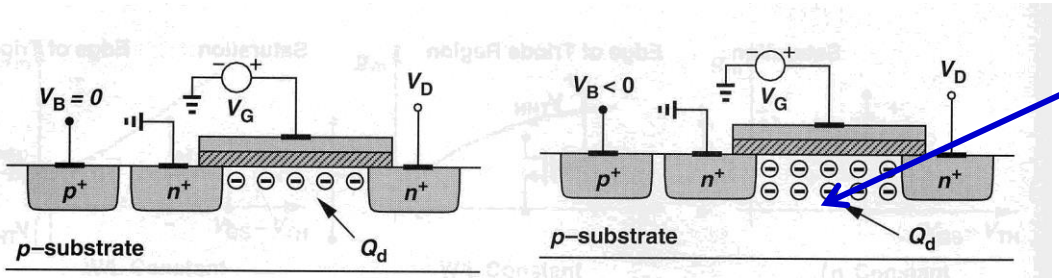
→ This would almost certainly result in a **useless** circuit!

Therefore, for integrated circuits, the MOSFET source terminals are **not** connected to the substrate body.

- Actually, the substrate is connected to the most negative power supply for NMOS circuit for achieving the desired functionality from the device.
- In such a scenario, what happens if the bulk voltage drops below the source voltage?
- Now the voltage $\boldsymbol{V_{SB}}$(voltage source-to-body) is **not** necessarily equal to zero (i.e., $V_{SB} \neq 0$). Thus, we are back to a **four-terminal** MOSFET device.
- There are **many** ramifications of this body effect; perhaps the most significant is with regard to the **threshold voltage** V_T.

Second Order Effect – Body Effect

- To understand, let us assume $V_S = V_D = V_B = 0$, and V_G is somewhat less than V_T. A depletion region forms but no inversion layer exists.
- As V_B becomes negative, more holes get attracted to the substrate which leaves a larger negative charge behind and as a result the depletion region becomes wider.



- The wider depletion region leads to increase in threshold voltage given by:

$$V_T = V_{T0} + \gamma\left(\sqrt{|2\Phi_F + V_{SB}|} - \sqrt{|2\Phi_F|}\right)$$

where γ and Φ_F are MOSFET **device parameters** and are essentially process dependent.

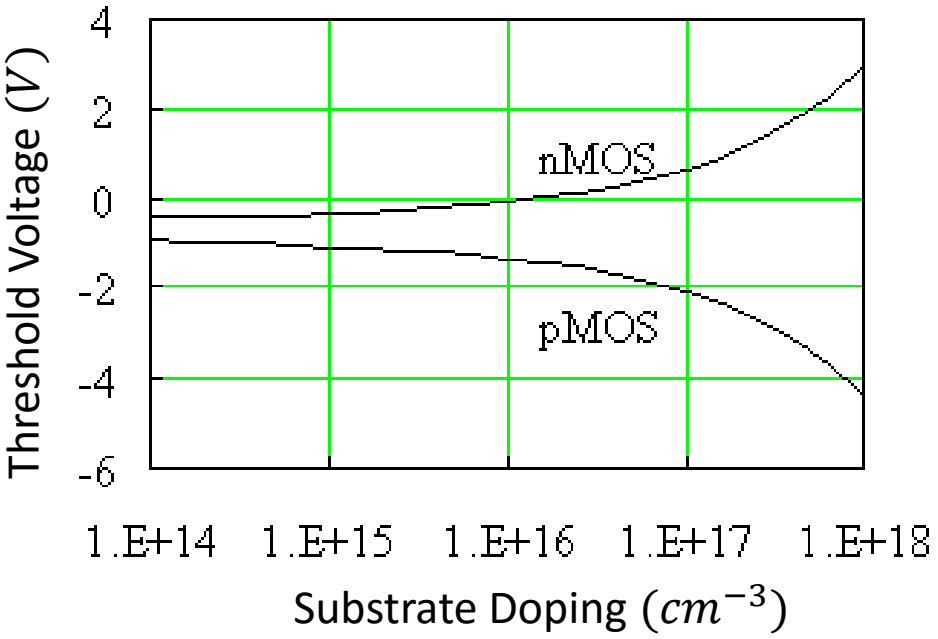
- Note the value V_{T0} is the value of the threshold voltage **when** $V_{SB} = 0 \leftrightarrow$ the value V_{T0} is simply the value of the device parameter V_T that we have been calling the threshold voltage up till now, i.e., without body effect.

Second Order Effect – Body Effect

- It is thus evident that the term $\gamma\left(\sqrt{|2\Phi_F| + V_{SB}} - \sqrt{|2\Phi_F|}\right)$ simply expresses an **extra** value added to the "ideal" threshold voltage V_{T0} when $V_{SB} \neq 0$.

Questions

- **Can a circuit designer control threshold voltage?**
- **Body effect is desirable or undesirable?**



Home Assignment # 0

Second Order Effect – Body Effect (contd.)

- Effect of body effect on Transconductance (g_m).
- Let us check the sensitivity of I_{DS} to V_{SB}

g_m

Simplification Yields:

$$g_{mb} = \mu_n C_{ox} \frac{W}{L}(V_{GS} - V_{T0})\left(-\frac{\partial V_T}{\partial V_{BS}}\right)$$

$$\left(\frac{\partial V_T}{\partial V_{BS}}\right) = -\left(\frac{\partial V_T}{\partial V_{SB}}\right) = -\frac{\gamma}{2}(2\Phi_F + V_{SB})^{-\frac{1}{2}} \quad \Rightarrow g_{mb} = g_m \frac{\gamma}{2\sqrt{2\Phi_F + V_{SB}}} = \eta g_m$$

$\eta = g_{mb}/g_m$ η=1/3 to 1/4, bias dependent

Thus, body effect has potential to alter (reduce) transconductance and therefore can impact device performance adversely

For many cases, we find that this Body Effect is relatively insignificant, so we will (unless **otherwise** stated) **ignore the Body Effect**.

However, do **not** conclude that the Body Effect is **always** insignificant—it can in some cases have a tremendous impact on MOSFET circuit performance!

Second Order Effect – Subthreshold Conduction

- For $V_{GS} \approx V_T$, a "week" inversion layer still exists and some current flows from D to S.
- Even for $V_{GS}<V_T$, I_D is finite → subthreshold conduction
- For V_{DS} greater than roughly 200 mV: $I_D = I_0 \exp \frac{V_{GS}}{\zeta V_{TM}}$

$V_{TM} = kT/q$

Nonideality Factor

- When V_{GS} is brought to zero, there will be some drain current in the channel. This is clearly unwanted situation as this small current will cause significant power consumption in large circuits

Second Order Effect – Voltage Limitations

- Various breakdown occurs if their terminal voltage differences exceed certain limits
- At high V_{GS}, the gate oxide breaks down irreversibly
- In short-channel devices, an excessively large V_{DS} can widen the depletion region around the drain so much that it touches that around the source, creating a very large drain current → punch through

Question

- Can channel length modulation affect NMOS/PMOS performance in linear/triode operation mode?