4.2 Current Voltage Characteristics

Reading Assignment: *pp. 248-262*

A. MOSFET Circuit Symbols

HO: The Circuit Symbols of Enhancement MOSFETs

B. i_D Dependence on v_{DS} and v_{GS}

Q:

A: HO: A Mathematical Description of MOSFET Behavior

Q:

A: HO: Channel Resistance for Small v_{DS}

C. Drain Output Resistance

Q:

A:

HO: Drain Output Resistance

D. The Body Effect

Q:

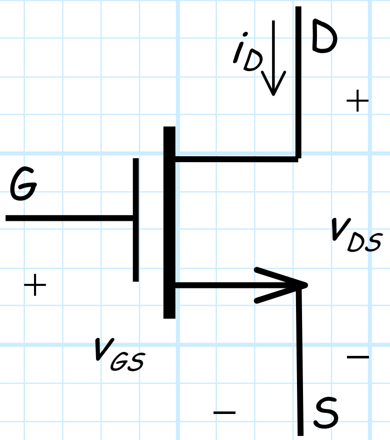
A:

HO: The Body Effect

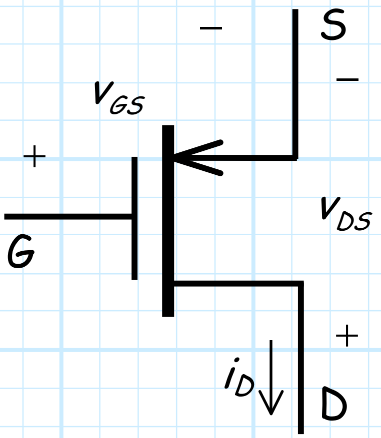
The Circuit Symbols of Enhancement MOSFETs

If we assume that the **body** and the **source** of a MOSFET are tied (i.e., **connected**) together, then our four-terminal device becomes a **three-terminal** device!

The circuit symbols for these three-terminal devices (NMOS and PMOS) are shown below:



Enhancement NMOS



Enhancement PMOS

Study these symbols carefully, so **you** can quickly identify the symbol and the name of each terminal (e.g., source S, gate G).

Likewise, make sure **you** can correctly label the relevant currents and voltages—including the polarity of the voltages and the direction of the current i_D!

A Mathematical Description of MOSFET Behavior

Q: *We've learned an awful lot about enhancement MOSFETs, but we still have **yet** to established a **mathematical** relationships between i_D, v_{GS}, or v_{DS}. How can we determine the correct **numeric** values for MOSFET voltages and currents?*

A: A **mathematical** description of enhancement MOSFET behavior is relatively straightforward ! We actually need to concern ourselves with just **3 equations**.

Specifically, we express the drain current i_D in terms of v_{GS} and v_{DS} for each of the **three MOSFET modes** (i.e., Cutoff, Triode, Saturation).

Additionally, we need to mathematically define the **boundaries** between each of these three modes!

But first, we need to examine some fundamental **physical parameters** that describe a MOSFET device. These parameters include:

$$k' \doteq \text{Process Transconductance Parameter} \left[A/V^2\right]$$

$$\frac{W}{L} = \text{Channel Aspect Ratio}$$

The Process Transconductance Parameter k' is a constant that depends on the process technology used to fabricate an integrated circuit. Therefore, all the transistors on a given substrate will typically have the **same value** of this parameter.

The Channel Aspect Ratio W/L is simply the ratio of channel width W to channel length L. This is the MOSFET device parameter that can be **altered** and **modified** by the circuit designer to satisfy the requirements of the given circuit or transistor.

We can likewise combine these parameter to form a **single** MOSFET device parameter ***K***:

$$\boxed{K = \frac{1}{2} k' \left(\frac{W}{L}\right) \qquad \left[{}^{A}\!/\!{}_{V^2}\right]}$$

Now we can mathematically describe the behavior of an enhancement MOSFET! Well do this **one mode at a time**.

CUTOFF

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

$$i_D = 0 \qquad \text{(CUTOFF mode)}$$

TRIODE

When in **triode** mode, the drain current is dependent on **both** v_{GS} and v_{DS}:

$$\begin{aligned} i_D &= k'\left(\frac{W}{L}\right)\left[\left(v_{GS}-V_t\right)v_{DS}-\frac{1}{2}\,v_{DS}^2\right] \qquad \text{(TRIODE mode)} \\ &= K\left[2\left(v_{GS}-V_t\right)v_{DS}-v_{DS}^2\right] \end{aligned}$$

This equation is valid for **both** NMOS and PMOS transistors (**if** in TRIODE mode). Recall that for **PMOS** devices, the values of v_{GS} and v_{DS} are **negative**, but note that this will result (correctly so) in a **positive** value of i_D.

SATURATION

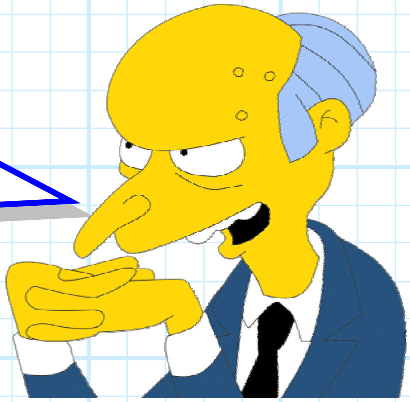
When in **saturation** mode, the drain current is (approximately) dependent on v_{GS} **only**:

$$\begin{aligned} i_D &= \frac{1}{2} k' \left(\frac{W}{L}\right) (v_{GS} - V_t)^2 \qquad \text{(SATURATION mode)} \\ &= K (v_{GS} - V_t)^2 \end{aligned}$$

Thus, we see that the drain current in saturation is proportional to **excess gate voltage squared**!

This equation is likewise valid for **both** NMOS and PMOS transistors (**if** in SATURATION mode).

Q: *OK, so know we know the expression for **drain current** i_D in each of the **three** MOSFET modes, but how will we know **what mode** the MOSFET is in?*



A: We must determine the **mathematical boundaries** of each mode. Just as before, we will do this **one mode at a time**!

CUTOFF

A MOSFET is in **cutoff** when **no channel** has been induced. Thus, for an enhancement **NMOS** device:

$$\text{if} \quad v_{GS} - V_t < 0 \quad \text{then NMOS in CUTOFF}$$

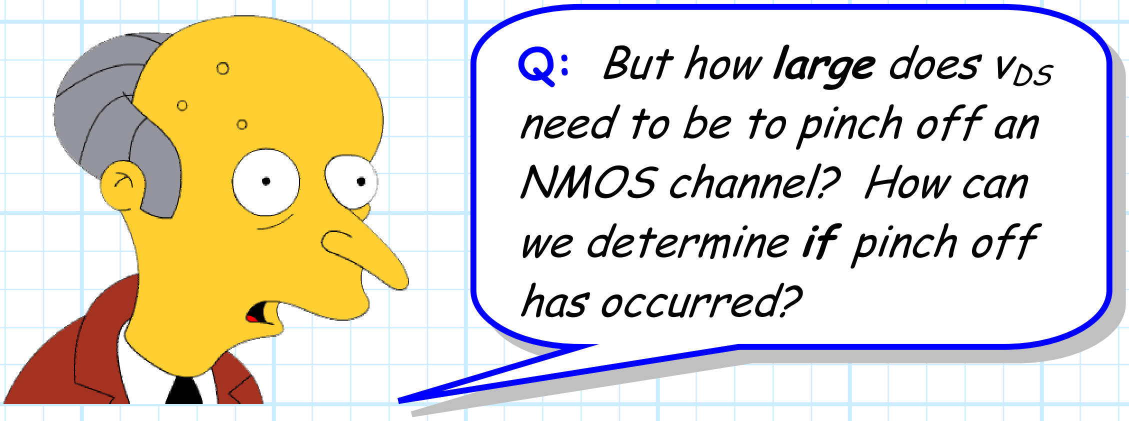
Like wise, for an enhancement **PMOS** device:

$$\boxed{\text{if} \quad v_{GS} - V_t > 0 \quad \text{then PMOS in CUTOFF}}$$

TRIODE

For triode mode, we know that a channel **is** induced (i.e., an inversion layer is present).

Additionally, we know that when in triode mode, the voltage v_{DS} is not sufficiently large for NMOS, or sufficiently small (i.e., sufficiently negative) for PMOS, to pinch off this induced channel.



Q: *But how **large** does v_{DS} need to be to pinch off an NMOS channel? How can we determine **if** pinch off has occurred?*

A: The answer to that question is surprisingly **simple**. The induced channel of an NMOS device is pinched off **if** the voltage v_{DS} is greater than the **excess gate voltage**! I.E.:

$$\boxed{\text{if} \quad v_{DS} > v_{GS} - V_t \quad \text{then NMOS channel is "pinched off"}}$$

Conversely, for PMOS devices, we find that:

if $v_{DS} < v_{GS} - V_t$ then PMOS channel is "pinched off"

These statements of course mean that an NMOS channel is **not** pinched off if $v_{DS} < v_{GS} - V_t$, and a PMOS channel is **not** pinched off if $v_{DS} > v_{GS} - V_t$. Thus, we can say that an **NMOS** device is in the **TRIODE** mode:

if $v_{GS} - V_t > 0$ and $v_{DS} < v_{GS} - V_t$ then NMOS in TRIODE

Similarly, for **PMOS**:

if $v_{GS} - V_t < 0$ and $v_{DS} > v_{GS} - V_t$ then PMOS in TRIODE

SATURATION

Recall for SATURATION mode that a channel **is** induced, and that channel **is** pinched off.

Thus, we can state that for **NMOS**:

$$\text{if} \quad v_{GS} - V_t > 0 \quad \text{and} \quad v_{DS} > v_{GS} - V_t \quad \text{then NMOS in SAT.}$$

And for **PMOS**:

$$\text{if} \quad v_{GS} - V_t < 0 \quad \text{and} \quad v_{DS} < v_{GS} - V_t \quad \text{then PMOS in SAT.}$$

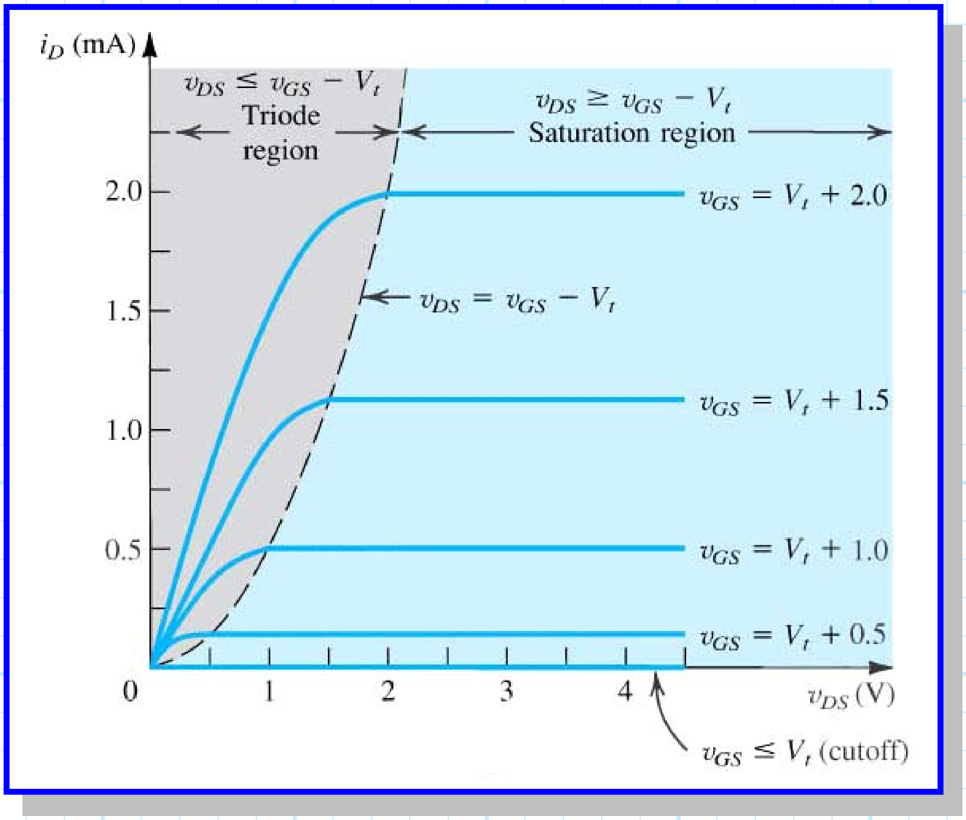
We now can construct a **complete** (continuous) expression relating drain current i_D to voltages v_{DS} and v_{GS}. For an **NMOS** device, this expression is:

$$i_D = \begin{cases} 0 & \text{if} \quad v_{GS} - V_t < 0 \\ K\left[2(v_{GS} - V_t)v_{DS} - v_{DS}^2\right] & \text{if} \quad v_{GS} - V_t > 0 \text{ and } \mathrm{v}_{\mathrm{DS}} < v_{GS} - V_t \\ K\left(v_{GS} - V_t\right)^2 & \text{if} \quad v_{GS} - V_t > 0 \text{ and } \mathrm{v}_{\mathrm{DS}} > v_{GS} - V_t \end{cases}$$

Likewise, for a **PMOS** device we find:

$$i_D = \begin{cases} 0 & \text{if} \quad v_{GS} - V_t > 0 \\ K\left[2(v_{GS} - V_t)v_{DS} - v_{DS}^2\right] & \text{if} \quad v_{GS} - V_t < 0 \text{ and } v_{DS} > v_{GS} - V_t \\ K(v_{GS} - V_t)^2 & \text{if} \quad v_{GS} - V_t < 0 \text{ and } v_{DS} < v_{GS} - V_t \end{cases}$$

Let's take a look at what these expressions look like when we **plot** them. Specifically, for an NMOS device let's plot i_D versus v_{DS} for different values of v_{GS}:



Channel Resistance for Small V_{DS}

Recall voltage v_{DS} will be **directly proportional** to i_D, provided that:

1. A conducting channel has been **induced**.
2. The value of v_{DS} **is small**.

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

Recall also that 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, we have previously determined that 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, an effect that works to proportionally increase the drain current* i_D
2. *Increasing* v_{DS} *will decrease the conductivity of the induced channel, an effect that works to decrease the drain current* i_D.

Q: *That's quite a **coincidence**! There are **two** physical phenomena at work as we increase v_{DS}, and there are **two** terms in the triode drain current equation!*

$$
\begin{aligned}
i_D &= K\left[2\left(v_{GS}-V_t\right)v_{DS}-v_{DS}^2\right] \\
&= 2K\left(v_{GS}-V_t\right)v_{DS}-K\,v_{DS}^2
\end{aligned}
$$

A: This is **no** coincidence! **Each** term of the triode current equation effectively describes **one** of these two physical phenomena.

We can thus **separate** the triode drain current equation into **two components**:

$$i_D = i_{D1} + i_{D2}$$

where:

$$i_{D1} = 2K\left(v_{GS}-V_t\right)v_{DS}$$

and:

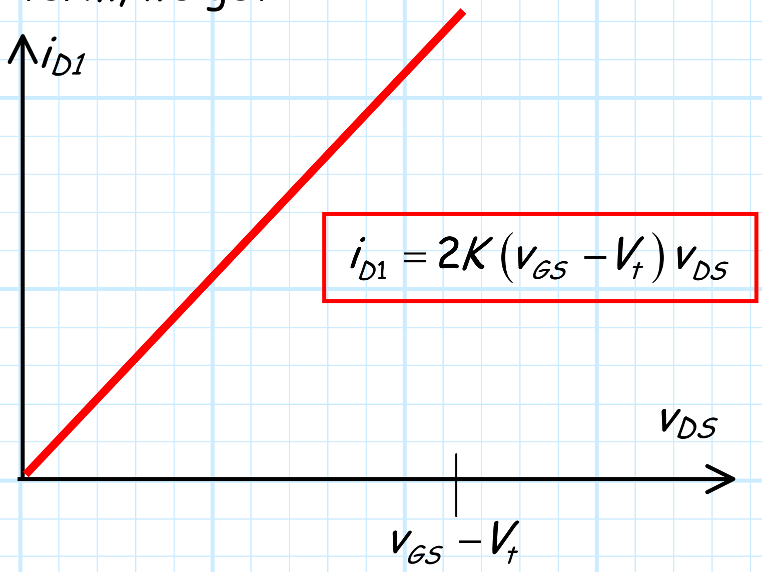
$$i_{D2} = -K\,v_{DS}^2$$

Let's look at each term **individually**.

$$i_{D1} = 2K\left(v_{GS} - V_t\right)v_{DS}$$

We first note that this term is **directly proportional** to v_{DS}— if v_{DS} increases 10%, the value of this term will increase 10%. Note that this is true **regardless** of the magnitude of v_{DS}!

Plotting this term, we get:



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

1. *Increasing v_{DS} will increase the potential difference across the conducting channel, an effect that works to proportionally increase the drain current 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**!

But of course, it does **not** behave like a resistor! The **second** term i_{D2} describes this very **nonresistor-like** behavior.

$$i_{D2} = -K\, v_{DS}^2$$

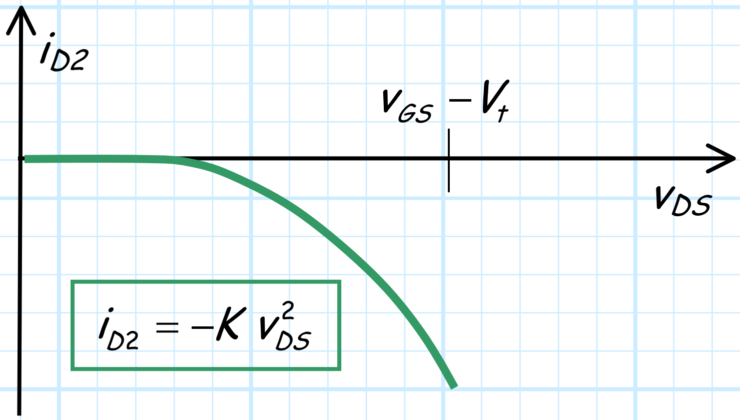
Q: *My Gosh! It is apparent that i_{D2} is **not** directly proportional to v_{DS}, but instead proportional to v_{DS} **squared**!!*

*Moreover, the minus sign out front 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??*

A: This **second term** i_{D2} essentially describes the result of the **second** phenomena:

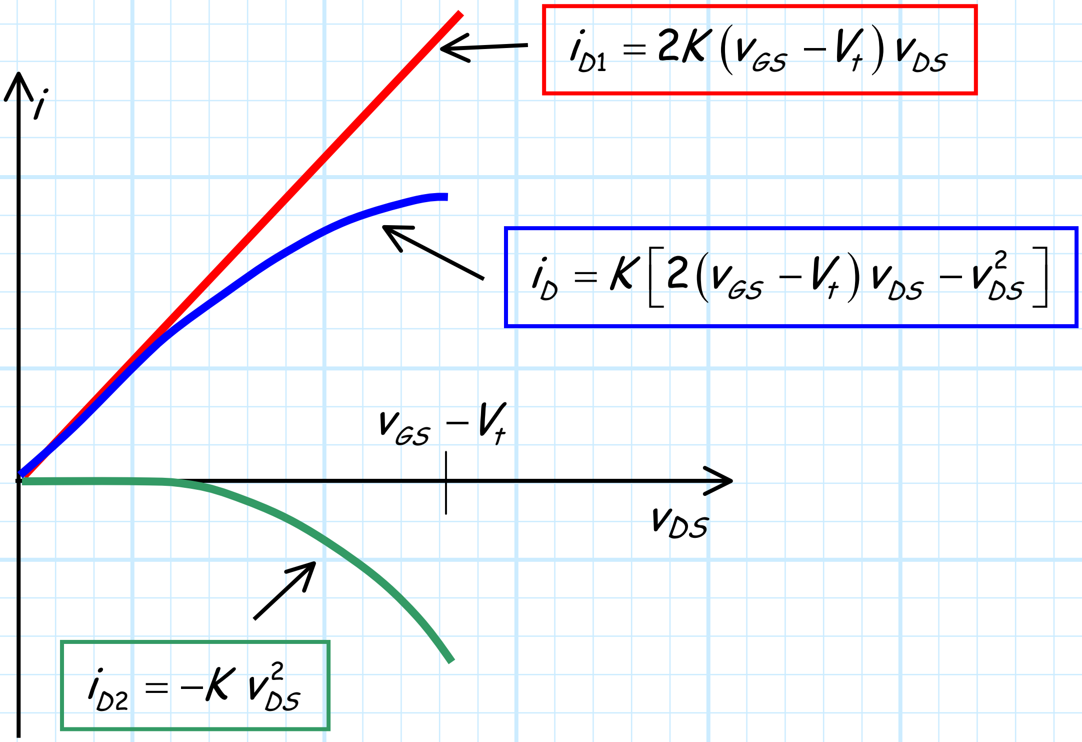
> **2.** *Increasing v_{DS} will decrease the conductivity of the induced channel, an effect that works to decrease the drain current i_D.*

Plotting this term, we get:



A **very** unresistor-like behavior !

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**!

Thus, we can conclude:

$$\boxed{\begin{aligned} i_D &\approx i_{D1} \\ &= 2K\left(v_{GS} - V_t\right)v_{DS} \\ &= k'\left(\frac{W}{L}\right)\left(v_{GS} - V_t\right)v_{DS} \qquad \text{for small } v_{DS} \end{aligned}}$$

Moreover, we can (for small v_{DS}) approximate the induced channel as a resistor r_{DS} of value $r_{DS} = v_{DS}/i_{DS}$:

$$
\begin{aligned}
r_{DS} &\doteq \frac{v_{DS}}{i_D} \\
&= \frac{v_{DS}}{2K\left(v_{GS}-V_t\right)v_{DS}} \\
&= \frac{1}{2K\left(v_{GS}-V_t\right)} \qquad \text{for small } v_{DS}
\end{aligned}
$$

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 this as:

$$
\begin{aligned}
\left|i_{D2}\right| &\ll \left|i_{D1}\right| \\
K\, v_{DS}^2 &\ll 2K\left(v_{GS}-V_t\right)v_{DS} \\
v_{DS} &\ll 2\left(v_{GS}-V_t\right)
\end{aligned}
$$

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**:

$$\begin{aligned} i_D &\approx i_{D1} \\ &= 2K\left(v_{GS}-V_t\right)v_{DS} \\ &= k'\left(\frac{W}{L}\right)\left(v_{GS}-V_t\right)v_{DS} \qquad \text{for } v_{DS} \ll 2\left(v_{GS}-V_t\right) \end{aligned}$$

and:

$$\begin{aligned} r_{DS} &= \frac{1}{2K\left(v_{GS}-V_t\right)} \\ &= \frac{1}{k'\left(W/L\right)\left(v_{GS}-V_t\right)} \qquad \text{for } v_{DS} \ll 2\left(v_{GS}-V_t\right) \end{aligned}$$

Q: *There you go* ***again!*** *The statement* $v_{DS} \ll 2\left(v_{GS}-V_t\right)$ *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 $i_D \approx i_{D1} = 2K(v_{GS}-V_t)v_{DS}$ 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}:

$$i_{D2} < \frac{i_{D1}}{10}.$$

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 $v_{DS} < (v_{GS}-V_t)/10$.

However, **using the 10% error criteria**, we arrive at the conclusion that:

$$
\boxed{
\begin{aligned}
i_D &\approx i_{D1}\\
&= 2K\left(v_{GS}-V_t\right)v_{DS}\\
&= k'\left(\frac{W}{L}\right)\left(v_{GS}-V_t\right)v_{DS} \qquad \text{for} \quad v_{DS} < \left(v_{GS}-V_t\right)/5
\end{aligned}
}
$$

and:

$$
\boxed{
\begin{aligned}
r_{DS} &= \frac{1}{2K\left(v_{GS}-V_t\right)}\\
&= \frac{1}{k'\left(W/L\right)\left(v_{GS}-V_t\right)} \qquad \text{for } v_{DS} < \left(v_{GS}-V_t\right)/5
\end{aligned}
}
$$

We find 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.*

Drain Output Resistance

I **fibbed**!

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.:*

$$i_D = K\left(v_{GS} - V_t\right)^2$$

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

Due to a phenomenon known as **channel-length modulation**, we find that drain current i_D is **slightly dependent** on v_{DS}. We find that a **more accurate** expression for drain current for a MOSFET in **saturation** is:

$$\boxed{i_D = K\left(v_{GS} - V_t\right)^2\left(1 + \lambda v_{DS}\right)}$$

Where the value 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}.

Often, the channel-length modulation parameter λ is expressed as the **Early Voltage** V_A, which is simply the inverse value of λ:

$$\boxed{V_A = \frac{1}{\lambda} \quad [\mathrm{V}]}$$

Thus, the drain current for a MOSFET in **saturation** can **likewise** be expressed as:

$$\boxed{i_D = K\left(v_{GS} - V_t\right)^2 \left(1 + \frac{v_{DS}}{V_A}\right)}$$

Now, let's **define** a value I_D, 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:

$$I_D \doteq K\left(v_{GS} - V_t\right)^2$$

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

$$i_D = I_D \left(1 + \frac{v_{DS}}{V_A}\right)$$

This **explicitly** shows how the drain current behaves as a function of voltage v_{DS}. For example, consider a **typical** case case where v_{DS}=5.0 V and V_A= 50.0 V. We find that:

$$\begin{aligned} i_D &= I_D\left(1+\frac{v_{DS}}{V_A}\right) \\ &= I_D\left(1+\frac{5.0}{50.0}\right) \\ &= I_D\left(1+0.1\right) \\ &= 1.1\, I_D \end{aligned}$$

In other words, the drain current is **10% larger** than its "**ideal**" value I_D.

We can thus 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 $I_D = K\left(v_{GS} - V_t\right)^2$.

Thus, as v_{DS} increases, the drain current i_D will **increase slightly**.

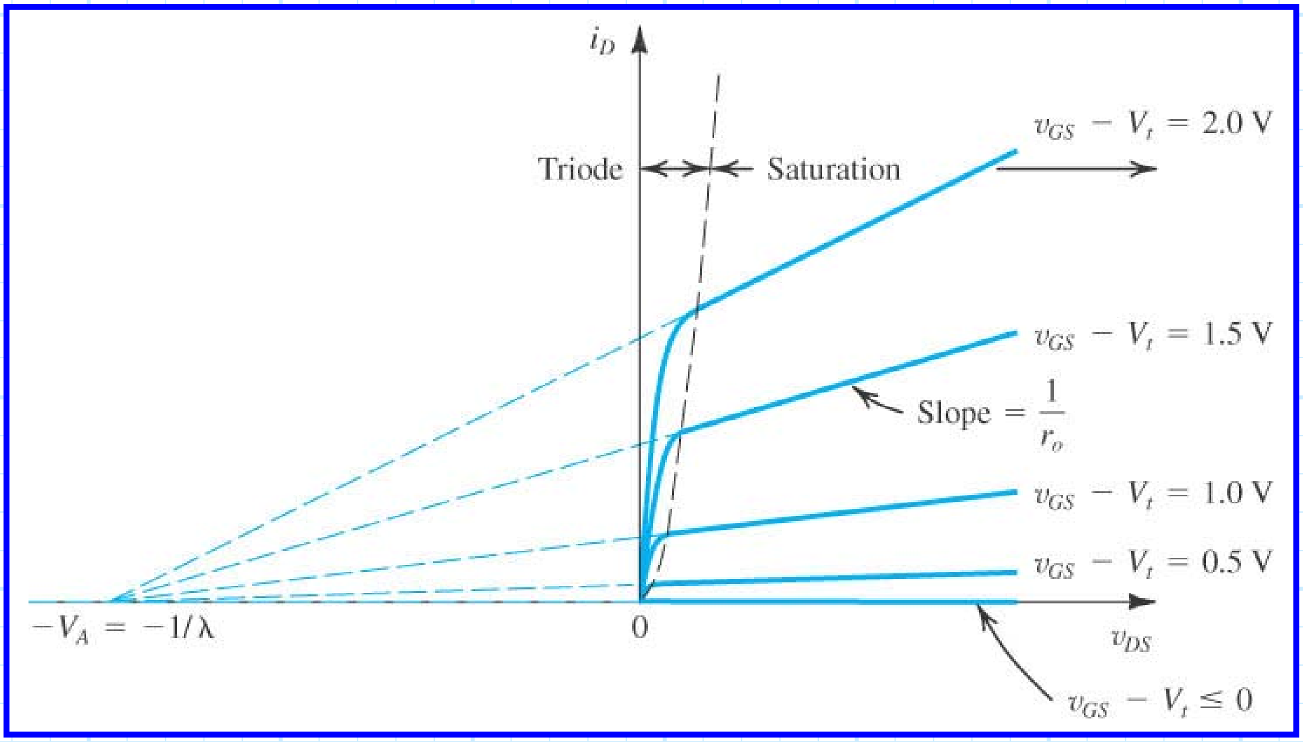
Now, let's introduce a **third** way (i.e. in addition to λ, V_A) to describe the "extra" current created by channel-length modulation. Define the **Drain Output Resistance** r_o:

$$\boxed{r_o \doteq \frac{V_A}{I_D} = \frac{1}{\lambda I_D}}$$

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

$$
\begin{aligned}
i_D &= I_D\left(1+\frac{v_{DS}}{V_A}\right)\\
&= I_D+\frac{I_D}{V_A}v_{DS}\\
&= I_D+\frac{v_{DS}}{r_o}\\
&= K\left(v_{GS}-V_t\right)^2+\frac{v_{DS}}{r_o}
\end{aligned}
$$

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_D!).

* Often, we "**neglect** the effect of channel-length modulation", meaning that we use the **ideal** case for saturation--$i_D=K(v_{GS}-V_t)^2$. Effectively, we assume that $\lambda = 0$, 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:

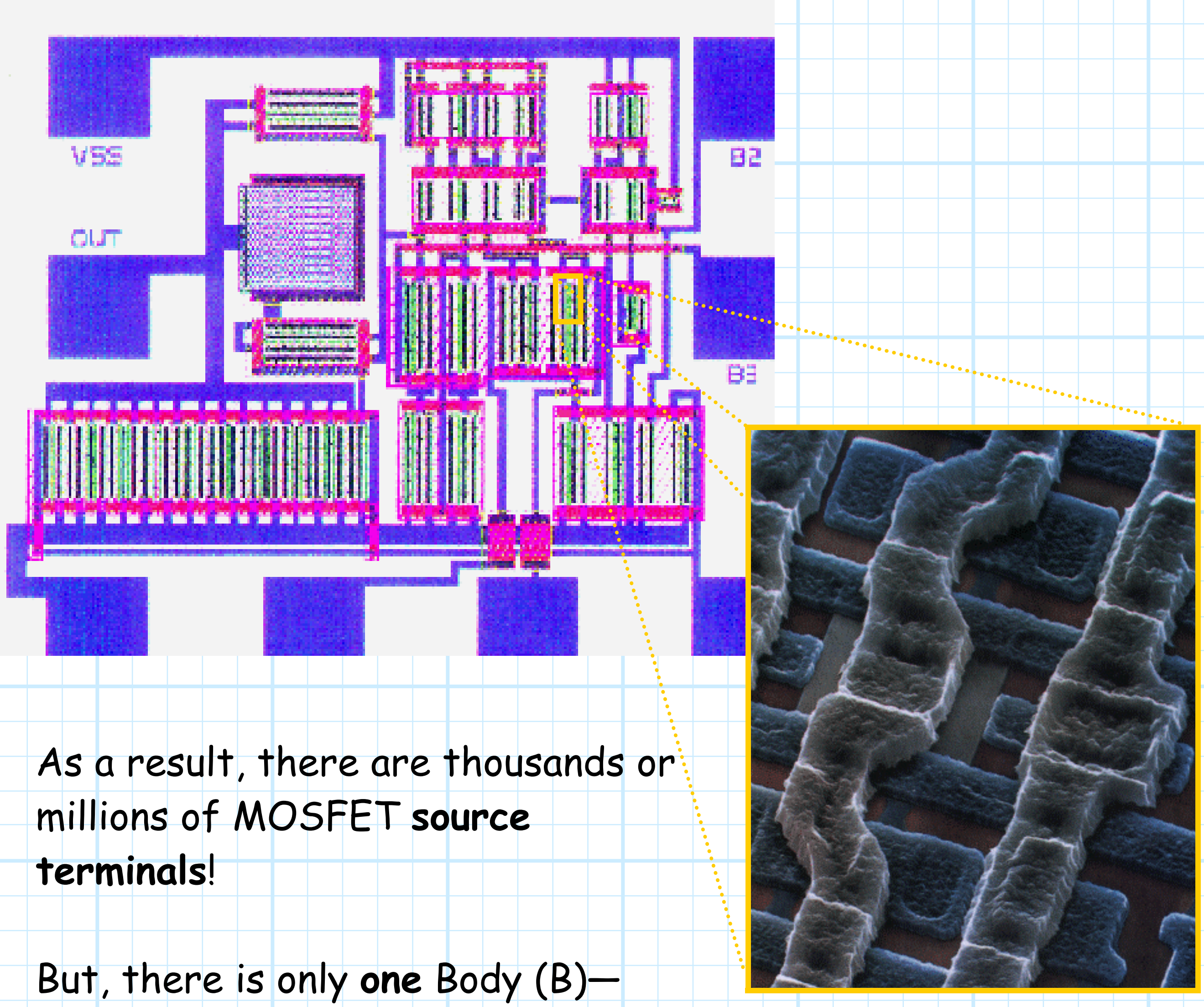
$$i_D = K\left(v_{GS} - V_t\right)^2 + \frac{v_{DS}}{r_o}$$ for a MOSFET in **saturation**

$$i_D = \frac{v_{DS}}{r_{DS}}$$ for a MOSFET in **triode** and v_{DS} small

$$\therefore \quad r_o \neq r_{DS} \qquad !!!!!!!$$

The Body Effect

In an integrated circuit using MOSFET devices, there can be **thousands** or **millions** of transistors.



As a result, there are thousands or millions of MOSFET **source terminals**!

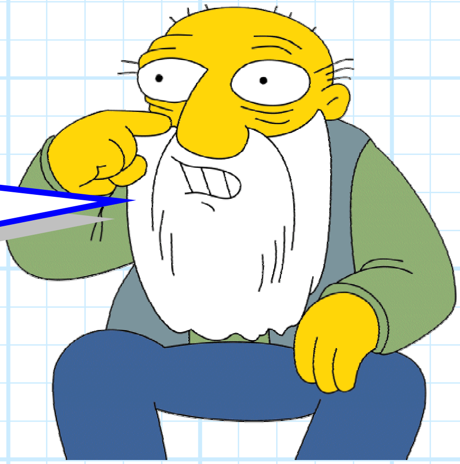
But, there is only **one** Body (B)—the Silicon **substrate**.

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!

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

Q: *Yikes! What happens to MOSFET behavior if the source is* ***not*** *attached to the body ??*



A: We must consider the MOSFET **Body Effect**!

We note that the voltage v_{SB} (voltage source-to-body) is **not** necessarily equal to zero (i.e., $v_{SB} \neq 0$)! Thus, were 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.

We find that when $v_{SB} \neq 0$, a more **accurate** expression of the threshold voltage is:

$$\boxed{V_t = V_{t0} + \gamma\sqrt{2\phi_f + v_{SB}} - \gamma\sqrt{2\phi_f}}$$

where γ and ϕ_f are MOSFET **device parameters**.

Note the value V_{t0} is the value of the threshold voltage **when** $v_{SB} = 0$, i.e.:

$$V_t = V_{t0} \quad \text{when} \quad v_{SB} = 0.0$$

Thus, 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!

In other words, V_{t0} is the value of the threshold voltage when we **ignored** the Body Effect, or when $v_{SB} = 0$.

It is thus evident that the term:

$$\gamma\sqrt{2\phi_f + v_{SB}} - \gamma\sqrt{2\phi_f}$$

simply expresses an **extra** value added to the "ideal" threshold voltage V_{t0} when $v_{SB} \neq 0$.

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!