Gate-Soure/Gate-Drain Overlap provides overlap capacitance $C_{ox}$, provided through a per-unit-width parameter.

Birds Peak Additional Gate to bulk Capacitance, provided per-unit-length units (rather per-unit width).

Reduction of Drawn Channel

Gate-Soure/Gate-Drain Overlap provides overlap capacitance $C_{ox}$, provided through a per-unit-width parameter.

*note, actual encroachment into the channel and into the junctions may be different

Direction of Current Flow

Vertical Contacts

Gate

Source and Drain

Substrate

Source and Drain

Gate

Gate

Source and Drain

Gate

Source and Drain

Gate

Direction of Transistor Length and Current Flow

Direction of Transistor Length and Current Flow

Direction of Transistor Width

Direction of Transistor Length and Current Flow

Direction of Transistor Length and Current Flow

Direction of Transistor Length and Current Flow
Diffusion: under thermal excitation, carriers (holes or electrons) exhibit a net movement from areas of high concentration to areas of low concentration.

Drift: carriers (holes or electrons) move in a net direction under the influence of an electric field toward areas of lowest electrical potential.
Consider a cluster of particles in some vicinity in free space with a system temperature $T$ with no applied force (no electric field force on a charge and no gravitation force on a mass $m$). $T$ defines the average movement of the particles.

If the barrier is removed, and we wait...we expect a net outward movement, even without any net force compelling any particular particle to move in a particular direction. The longer we wait the further the spread we expect.
Another experiment: this time, an barrier box is introduced. After some time, we expect an even distribution in the box. At equilibrium, no net movement statistically, even though total individual particle movement characterized by $T$.
Yet another experiment: An adjacent empty box is added, and an opening in the barrier between them is removed. After some time... we expect a new equilibrium, established through an undriven "flow" with a rate through the opening proportional to the difference in concentration and temperature.

\[ \text{rate} : J = N \cdot v_{\text{diff}} \propto \frac{d\rho}{dx}, T ; \]

- \( N \) represents the number of particles in the volume
- \( v_{\text{diff}} \) : effective velocity of diffusion
- \( \rho \) : density of particles
- \( \frac{d\rho}{dx} \) : concentration gradient at pos. \( x \)

Note: Once equilibrium is achieved, motion continues even though net motion is statistically zero.
Diffusion: under thermal excitation, carriers (holes or electrons) make a net movement from areas of high concentration to areas of low concentration.

Diffusion Current: Fick's Law

\[ J_{\text{diff}}(x) = -D \frac{dp(x)}{dx} \] (minus sign indicates that particles flow from high concentration to low)
Question to answer: For a given applied voltage, what determines the drift current in a resistive material?

Drift current may be thought of as some number of carriers in a unit volume moving at some average speed

- a number (or density) of mobile charge carriers
- electric field to push (accelerate) the carriers in a net direction
- average amount of time a carrier will accelerate before a collision or recombination, which determines the average speed of carriers

### Drift Current

\[
J_{\text{drift}}(x) = \rho(x) \mu F(x) = -\rho(x) \mu \frac{dU(x)}{dx}
\]

Note: a force defines a potential gradient in the opposite direction of the force
\(\mu\): ratio of force to average velocity (factors to be discussed later)
F: force (gravity, e-field)
U: potential energy
Drift and Diffusion (Gravity)

More boxes: introduce gravity and particle mass with vertical boxes:

At "equilibrium", at every height, $h$, $v_{\text{drift}} = v_{\text{diffusion}}$

Concentration between points is related by

$N = N_0 e^{-w \cdot h / kT}$; where $w \cdot h$ is the energy with respect to the reference height $h_0$ where the concentration is $N_0$
Now turn the boxes back, introduce e-field and particle charge & ignore gravity (assume small)

- Potential Energy is defined by $qV$
  - $\rho(x) = \rho(0)e^{-qV(x)/kT}$ Maxwell-Boltzmann Distribution
  - $V(x) = \ln \left( \frac{\rho(0)}{\rho(x)} \right) e^{-kT/q}$ Nerst Potential
  - $k$: Boltzman’s constant
  - $q$: fundamental charge
Relationship of Drift and Diffusion

A relationship between diffusion and drift can be found by studying the aforementioned equilibrium condition.

- At every point: 
  \[ 0 = J_{\text{drift}} + J_{\text{diffusion}} = -\rho(x)\mu \frac{dU}{dx} - D \frac{d\rho}{dx} \]

- Boltzmann distribution relates statistical average number of carriers to local potential at position \( x \): 
  \[ \rho(x) = Ae^{-\frac{U(x)}{kT}} ; \quad A \text{ is some constant.} \]

\[ \frac{d\rho(x)}{dx} = -\frac{1}{kT} \frac{dU(x)}{dx} \rho(x) \]

- Since this equation holds everywhere a non-trivial solution requires: 
  \[ \mu = \frac{D}{kT} \]

**Einstein relation general form**

\[ D = \mu kT \]

- \( D \) is the diffusion constant
- \( \mu = v_{\text{drift}}/F \): "mobility", ratio of the particle’s terminal drift velocity to an applied force
- \( k \) is Boltzmann’s constant, \( T \) is the absolute temperature
Gate-Soure/Gate-Drain Overlap provides overlap capacitance $C_{ox}$, provided through a per-unit-width parameter.

Birds Peak: Additional Gate to bulk Capacitance, provided per-unit-length units (rather per-unit width).

Reduction of Drawn Channel

Gate-Soure/Gate-Drain Overlap provides overlap capacitance $C_{ox}$, provided through a per-unit-width parameter.

*note, actual encroachment into the channel and into the junctions may be different.
FET Cross Sections

Top View Drawing

Select Mask
Substrate

Weff = W_{Drawn} - 2W_{FOE}

Cross Section
Cut Along Width

W_{Drawn}
Select Mask
Substrate

Cut Along Length

Top View Drawing

Select Mask (n-select or p-select)
diffusion opening
polysilicon

Cross Section Cut Along Length

Doping
Select Mask
Silicon Dopant
Implant

annealing

Diffusion, Post-Annealing

L_{Drawn}
LD
LD
LD
LD
Substrate

L_{eff} = L_{_{Drawn}} - 2LD
Annealing and Graded Junctions

Ideal Junction Depictions

Graded Junction Depiction

Sidewall Grading

Bottom Grading

LD $L_{\text{eff}}$ LD

Direction of Current Flow
Source, Gate, and Drain Potentials referenced to BULK potential

ENHANCEMENT-MODE MOSFET, very little current passes from the source to drain with only reversed-biased diodes in the path
Diode

Band Diagram

Conduction Band

Valence Band

Electric Field

Diffusion of high energy carriers

Force (on e−)

Free/Mobile Charge Carrier

Atom-Bound Potential Carrier

Larger V

More Diffusion

Lesser V

Less Diffusion

I = Is \left( e^{\frac{V}{n\phi T}} - 1 \right)

I = Is \left( e^{\frac{V}{n\phi T}} \right) \text{ for } V \gg 0

I = -I_s \text{ for } V \ll 0

The drawings to the left help visualize the probability of carriers overcoming the barrier and diffusing across the junction (electron distribution determined by Fermi distribution: with a moderate forward bias may be mentally approximated by a Boltzmann distribution for molecules)

With 0V applied, an equilibrium is established: a built-in potential develops along with equal but opposite drift and diffusion currents

We the applied voltage is modified we observe significant changes in the diffusion current

Dr. Ryan Robucci
Lecture 1
Surface Coupling for Gating (1)

Accumulation: Charge at Surface  Depletion: Charge modified at a depth (few mobile minority carriers)  Further Depletion: Weak Inversion: Charge modified at surface (depletion charge reaching ~max)  Strong Inversion: Charge modified at surface (depletion charge ~max)

Critical Point

Each is charge-neutral looking from the top down

\[
\begin{align*}
-Q'_G &= Q'_C \\
Q'_C &= Q'_B + Q'_I \\
\Delta Q'_C &= \Delta Q'_B + \Delta Q'_I
\end{align*}
\]

\[V_G \quad V_B \quad V_{ox} \quad \psi_s \quad C'_{ox} \quad C'_g \quad C'_C \quad C'_B \quad \psi_s \]

+: Neutral Charge Atom, with easily lost hole/easily gained electron  
⊕: Free Hole  
○: Free Electron  
□: Negative Ion, Uncovered/Location-Bound/Fixed Charge

Accumulation  Depletion  Strong Inversion
Gate Capacitance Per Unit Area looking into the top is plotted – determined by materials, inversely proportional to separation of additional charge collected on top plate from the opposing charge below.

When the gate is at a low potential (negative), mobile majority carriers accumulate at the channel surface and the oxide thickness is the only separation between the balancing opposite charges on each capacitor plate.

As the gate voltage is increased, majority carriers are depleted forming a series capacitance of the oxide cap and depletion cap.

\[ C^{\text{series}} = \frac{1}{\frac{1}{C_1} + \frac{1}{C_2}} \equiv C_1 \parallel C_2 \]

As the gate voltage is increased, the distance to mobile carriers is increased as the depletion width increases – reducing the value of the depletion cap and the the series capacitance.

As the gate voltage is increases even further, mobile minority carriers are attracted to the surface, reducing the separation of the additional opposite charges to just the oxide thickness with the surface controlled by conduction to the source and drain and thus increasing the effective gate capacitance.

In the inversion mode, the charge-neutrality with the top plate is obtained below with a combination of inversion charge and depletion charge.
Surface Coupling for Gating (3)

\[ \Delta \psi_s = \kappa \cdot \Delta V_G + (1 - \kappa) \cdot \Delta V_B \]

- \( \kappa \) represents the gate to surface coupling coefficient; important for depletion mode

\[ \Delta \psi_s = \frac{C_{ox}}{C_{ox} + C_{dep}} \Delta V_G + \frac{C_{dep}}{C_{ox} + C_{dep}} \Delta V_B \]

\[ \frac{C_{ox} + C_{dep}}{C_{ox} + C_{dep}} - \frac{C_{ox}}{C_{ox} + C_{dep}} = 1 - \kappa \]
Capacitive Structure (depletion mode)

- \( V_G = \psi_{surf} + V_{ox} \) (gate voltage is channel surface potential + charge across oxide capacitor)
  - When \( V_G \) is small \( \rightarrow \) depletion mode.
  - May determine voltage across the oxide by realizing \( V_{ox} = f \) (oxide cap., depletion charge defined by \( \psi_{surf} \))
  - The top-plate gate charge is equal and opposite \( (Q'_G = -Q'_B) \), but the bulk depletion charge is associated with a depth and top-plate charge does not have any depth as the charge is associated with mobile carriers.

- \( Q'_C = Q'_B = -\sqrt{2q\varepsilon_S i N_a \psi_{surf}} \) [C/m^2] charge per area looking down and through the bulk
- \( \varepsilon_{Si} = 11.8\varepsilon_o \) permittivity of Si
  (Permittivity is a measure of how an electric field affects, and is affected by, a dielectric medium, and is determined by the ability of a material to polarize in response to the field, and thereby reduce the total electric field inside the material. Thus, permittivity relates to a material’s ability to transmit (or "permit") an electric field. –wikipedia)

- \( \varepsilon_o = 8.854 \times 10^{-12} [\text{F/m}] \) permittivity of free space
- \( N_A \) : acceptor donor density \([1/\text{m}^3]\)
- \( q = 1.602 \times 10^{-19} [\text{C}] \) elementary charge
Source, Gate, and Drain Potentials referenced to BULK potential

For ENHANCEMENT-MODE MOSFET, only very little current passes from the source to drain with only the reversed-biased diodes in the path.

In subthreshold mode, the diffusion of the near-exponential carrier distribution above the barrier is controlled by the terminal voltages. The gate terminal modulates the channel barrier with coefficient $\kappa$ and the bulk controls it with coefficient $(1 - \kappa)$, while the source and drain terminals modulate the respective potentials with coefficient 1.
Towards modeling the MOSFET, we wish to define a “threshold” for a gate voltage about where the device begins to form a significant “inversion layer” allowing additional current to follow from the source to drain.
$V_G \geq V_{TN} \rightarrow \text{Inversion Mode; } V_{TN}: \text{threshold voltage}$

both covered fix charge and mobile charge carriers near the surface

For inversion: $Q'_C = \underbrace{Q'_B}_{\sim \text{fixed max (depletion)}} + \underbrace{Q'_I}_{\text{mobile (inversion)}}$

The term Inversion describes when a p-type material with mobile p-carriers is made to instead have as many n-type mobile carriers, then has some behaviors of an n-type material. The presence of mobile majority carriers is replaced by the opposite presence of mobile minority carriers. When $\psi_{surf} = 2|\phi_F| = 2 \frac{kT}{q} \ln \left( \frac{N_a}{n_i} \right)$, moderate inversion begins.

- $\phi_F$: Fermi potential
- $\frac{kT}{q} = V_T$ or $U_T$ or $\phi_t$, referred to by thermal voltage or thermal potential, we will try to use $U_T$ or $\phi_t$ in part to avoid confusion with threshold voltage
- $k = 1.381 \times 10^{-23} \left[ \frac{J}{K} \right] = \text{Boltzmann’s constant}$
- $n_i \approx 1.45 \times 10^{10} \left[ \frac{1}{\text{cm}^3} \right] \circledast 300^\circ \text{K intrinsic carrier concentration}$
- $U_T \approx 26 \text{mV} \circledast 300^\circ \text{K}$
- $N_A \approx 10^{15} \text{cm}^{-3} \rightarrow 2|\phi_F| \approx .58 \text{V}$
We are interested in strong inversion where mobile charge a dominant charge over the bulk charge and grows with increasing gate voltage, \( Q_I = -C'_\text{ox}(V_{GB} - V_{TN0}) \). We can extrapolate the model back to a point \( V_{TN0} \) so that the inversion layer (density of mobile minority carriers) is 0, but it should be noted that at this point the model is not actually valid as the mobile inversion charge is not dominant.

In this model, \( \partial V_G = V_{TN} \) the model predicts \( Q'_I = 0 \rightarrow Q'_C = Q'_B \), (though the model itself is not actually valid at such a gate voltage since the transistor is not strongly inverted at such a voltage)

To estimate the oxide voltage here, use \( Q = CV \) with \( Q \) assumed to be due to bulk depletion charge only

\[
V_{ox} = \left| \frac{Q_B}{C_{ox}} \right| \left( \frac{C}{V/m^2} \right) = \frac{\sqrt{2q\varepsilon_{Si}N_a\phi_0}}{C_{ox}} [V]; \phi_0 \text{ is the value used to approximate the maximum bulk charge reached. The valued used is often estimated as } \phi_0 = 2|\phi_F|, \text{ though a better estimate is } \phi_0 = 2|\phi_F| + \Delta \phi_0 \text{ where } \Delta \phi_0 \text{ is several } \phi_t (6\phi_t).\]
As the gate voltage increases, the surface potential $\psi_s$ will increase. But the bulk charge will no longer increase significantly, mobile charge is attracted to the surface to balance the increasing positive charge on the gate.

$$C'_{ox} = \frac{\varepsilon_{ox}}{t_{ox}} = \frac{3.9 \varepsilon_o}{t_{ox}} \left[ \frac{F}{m^2} \right]$$  oxide capacitance per unit area

The ideal threshold voltage, $V_{TN,ideal} = \phi_0 + \frac{\sqrt{2q\varepsilon_{Si} N_a \phi_0}}{C'_{ox}}$ is the surface potential plus oxide voltage computed from the bulk charge/capacitance (above this voltage, the bulk charge remains relatively constant even as the surface potential changes as additional voltage on the capacitance is created by mobile charge.)
Can’t ignore the work function between the gate and substrate and various trapped charges in the capacitive structure. A voltage $V_{FB}$ is calculated to account for these.

Work function is the energy required to pull electron at the Fermi level away from material into a vacuum, which varies with different materials. The non-ideal insulation with trapped charge affecting the electric field in the oxide capacitance.

$$\Phi_{ms} = (\Phi_G - \Phi_{ox}) + (\Phi_{ox} - \Phi_{SUB}) = (\Phi_G - \Phi_{SUB})$$

$$V_{FB} = (\Phi_G - \Phi_{SUB}) - \frac{Q_f}{C'_{ox}} - \frac{\hat{Q}_{ox}}{C'_{ox}}$$

$V_{FB}$: voltage where there is no charge in the substrate

$$(\Phi_G - \Phi_S) = -\frac{kT}{q} \ln \left( \frac{N_a N_{d,poly}}{n_i^2} \right)$$ (negative value)

for p-poly and p-substrate we can use $$(\Phi_G - \Phi_S) \approx -\frac{kT}{q} \ln \left( \frac{N_{a,poly}}{n_i^2} \right)$$
\( \frac{Q_{ox}'}{C_{ox}'} \) approximates \( \frac{1}{\varepsilon_{ox}} \int_o^{t_{ox}} x \rho(x) \, dx \) where \( x \) is the vertical position in the oxide and \( \rho \) is the concentration \([ \frac{C}{cm^3}]\) of oxide charge at a particular depth.

\begin{align*}
V_{TN}' &= V_{FB} + \phi_0 + \frac{\sqrt{2q\varepsilon_{Si}N_a\phi_0}}{C_{ox}'} = V_{FB} + \phi_0 + \gamma\sqrt{\phi_0} ; \gamma = \frac{\sqrt{2q\varepsilon_{Si}N_a}}{C_{ox}'} \\
\end{align*}

under normal conditions \( V_{TN}' < 0 \), so additional ion implantation is used to adjust the threshold voltage. Assuming a surface implant,

\begin{align*}
V_{TN}^{\dagger} &= V_{FB} + \phi_0 + \gamma\sqrt{\phi_0} \pm \frac{qD_1}{C_{ox}'} ; \text{ where } D_1 \text{ is additional doping per unit area} \\
\end{align*}


\( Q_{I}' \approx -C_{ox}' (V_G - V_{TN}) \)

assume surface doping

for \( V_G \gg V_{TN} \), \( Q_{I}' \approx -C_{ox}' (V_G - V_{TN}) \)

a.k.a. \( V_{eff}, V_{ov}, V_{on} \)

as \( V_G \) increases, more and more mobile charge is made available.

The inversion charge will determine the conduction of a MOSFET in this present model of strong inversion.
Consider a MOSFET structure with a source. The source terminal which affects the potential and the number of carriers available at the source-end of the channel to be formed. By a common convention, we will first reference voltages to the source instead of the bulk terminal. To account for the effect of any source voltage, which affects the potential at the source-end of the channel and the number of available carriers, in strong inversion, the surface potential $\psi_s$ must be increased accordingly for the source-end of the channel. We add a term $V_{SB}$ to the surface potential formerly referenced to the bulk. This means replacing $\phi_0$ with $\phi_0 + V_{SB}$. The new source-referenced threshold voltage expression is

$$V_{TNL}^\dagger = V_{FB} + \phi_0 + V_{SB} + \frac{\sqrt{2q\varepsilon_{Si}N_a(\phi_0 + V_{SB})}}{C_{ox}'} \pm \frac{qD_{1}}{C_{ox}'}$$

The gate voltage with respect to the source must be

$$V_{TNS} = V_{TNL}^\dagger - V_{SB} = V_{FB} + \phi_0 + \frac{\sqrt{2q\varepsilon_{Si}N_a(\phi_0 + V_{SB})}}{C_{ox}'} \pm \frac{qD_{1}}{C_{ox}'}$$

If the bulk and source are at the same potential, $V_{SB} = 0$, we see $V_{TNS}$ is the same as our previous $V_{TN}^\dagger$. 
Very often in an IC circuit, the bulk terminals are connected to a common node, not the sources. It becomes convenient if/when using source-referenced equations to calculate the common baseline component of the threshold voltage when the source is the same potential as the bulk, and then an adjustment can be applied to each transistor based on its source voltage bias point.

First, the baseline threshold is calculated with the source at the bulk potential. That is subtracted from the full equation:

\[
\begin{align*}
\{ & V_{TN0} = V'_{TNS} \bigg|_{V_{SB}=0} = V_{FB} + \phi_0 + \frac{1}{C'_{ox}} \sqrt{2q\varepsilon_{Si}N_a(\phi_0 + 0)} \pm \frac{qD_I}{C'_{ox}} \\
- & V_{TN} = V'_{TNS} \bigg|_{V_{SB}=V_{SB}} = V_{FB} + \phi_0 + \frac{1}{C'_{ox}} \sqrt{2q\varepsilon_{Si}N_a(\phi_0 + V_{SB})} \pm \frac{qD_I}{C'_{ox}} \\
\}
\end{align*}
\]

\[\Delta V_{TN} = \frac{1}{C'_{ox}} \sqrt{2q\varepsilon_{Si}N_a} \left( \sqrt{(\phi_0 + V_{SB})} - \sqrt{(\phi_0)} \right)\]
Source-Referenced Model (3)

\[ V_{TN} = V_{TN0} + \Delta V_{TN} \]

- Common for similar devices with same bulk potential
- Calculated per device for a given source-bulk voltage

Source-Referenced Model Threshold Voltage

\[
V_{TN} = V_{TN0} + \gamma \left( \sqrt{\phi_0 + V_{SB}} - \sqrt{\phi_0} \right);
\]

\[
\gamma = \frac{1}{C'_{ox}} \sqrt{2q\varepsilon_{Si}N_A} \left[ \sqrt{V} \right]
\]

\( \gamma \) is called the body effect constant.

A typical value is 0.75 \( \left[ \sqrt{V} \right] \). Note the square-root dependence:

Initially, as the source voltage is raised the effective threshold voltage increases significantly, but the change is less dramatic for higher source voltages.

<table>
<thead>
<tr>
<th>( V_{SB} )</th>
<th>2.5</th>
<th>5.0</th>
<th>7.5</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>( V_{TN} )</td>
<td>0.73</td>
<td>1.26</td>
<td>1.68</td>
<td>2.04</td>
</tr>
</tbody>
</table>
Summary for source-referenced threshold voltage

\[ V_{TN} = V_{TN0} + \gamma \left( \sqrt{(\phi_0 + V_{SB})} - \sqrt{\phi_0} \right); \]

\[ \gamma = \frac{1}{C'_{ox}} \sqrt{2q\varepsilon_{Si} N_A} \left[ \sqrt{V} \right] \]

\[ \phi_0 \approx 2|\phi_F| + \Delta \phi_0 [V]; \Delta \phi_0 \approx 6\phi_t \]

\( \phi_F \): Fermi potential [V]

\( N_A \): acceptor donor density \( \left[ \frac{1}{m^3} \right] \); \( N_A \approx 10^{15} \text{cm}^{-3} \rightarrow 2|\phi_F| \approx .58 \text{V} \)

\[ \phi_t = \frac{kT}{q} [V] \]

\[ k = 1.381 \times 10^{23} \left[ \frac{J}{K} \right] \]

\( T \) is temp in °K

\[ C'_{ox} = \frac{\varepsilon_{ox}}{t_{ox}} \] capacitance per unit area \( \left[ \frac{F}{m^2} \right] \)

\( t_{ox} \) thickness of oxide insulator [m]

\[ \varepsilon_{ox} = 3.9\varepsilon_o \] permittivity of Si02

\[ \varepsilon_{Si} = 11.8\varepsilon_o \] permittivity of Si

\[ \varepsilon_o = 8.854 \times 10^{-12} \left[ \frac{F}{m} \right] \] permittivity of free space

\[ q = 1.602 \times 10^{-19} [\text{C}] \] elementary charge
Drift current

In accordance with the “Gradual Channel Approximation” discussed later, we will ignore the diffusion current and focus on the drift current for now.

Drift current question: For a given applied voltage, what determines the drift current in a resistive material?

Consider flow across a volume through a plane:

Drift current may be thought of as some number of carriers moving at some average speed

1. a number (or density) of mobile charge carriers
2. electric field to push (accelerate) the carriers in a net direction
3. average amount of time a carrier will accelerate before a collision or recombination, which determines the average speed of carriers
Resistance of a Semiconductor

Deriving the effective resistance of a segment of a semiconductor:

\[ R = \frac{L}{\sigma A} \]

- \( \sigma = q\mu_n n \left[ \frac{1}{\Omega \cdot \text{m}} \right] \) note: conductivity is really \( \sigma = q(n\mu_n + p\mu_h) \)
  but with n-type material (or strongly inverted p-type), \( n \gg p \)

- \( \mu_n \): mobility of material based on effective mean carrier lifetime \[ \left[ \frac{m^2}{V \cdot s} \right] \]

- \( n \): carrier density \[ \left[ \frac{1}{\text{m}^3} \right] \]

taking resistance per unit length \( \frac{R}{L} = \frac{1}{\sigma A} \), \( \frac{dR}{dx} = \frac{1}{\sigma A} \)

\( \sigma A \), is the conductivity of a unit length, and is proportional to the cross section area:

\[ \sigma A = \mu_n q(nWH) \]

\( \text{total carrier charge per unit length} \)

Consider flow across a volume through a plane and label the quantity, \( nWH \), in the following plot:
Considering a view of fish in a pond, we may collapse dimensions in the 3D calculation.

We will next consider the surface under the oxide of a transistor and consider flow across a sheet through a line.

A top view of sheet flow and volume flow look same, and in both the flow across a line in the top view can be considered. However, in the case of inversion layer, the carriers we consider are actually at (extremely near) the surface.
The surface potential varies along the channel, due to application of different source and drain potentials.
At any point \( x \) along the surface of the channel, the density of carriers is determined by
\[
Q_I'(x) = -C'_{ox} (V_{GS} - V_{TN} - V(x))
\]
and varies across the channel
\[
\frac{dV}{dx} = -C'_{ox} (V_{GS} - V_{TN} - V(x))
\]
Boundary conditions:
\( V(x = 0) = 0 \) (left-hand side of channel)
\( V(x = L) = V_{DS} \) (right-hand side of channel)
Inversion Layer resistance modeling

Since the inversion layer carrier density varies along the channel, so does the effective resistance of each segment:

For a small segment of a MOSFET channel, we will replace $qnHW$, the total carrier charge per unit length, with $-Q'(x)W$.

Then, $$\frac{dR}{dx} = \frac{1}{-\mu_n WQ'(x)}$$

For any segment, $V = IR$. Or, for infinitely small segments we can write $$\frac{dV}{dx} = I_D \frac{dR}{dx}.$$ $I_D$ does not depend on position $x$ since all segments are in series and therefore must have the same current (assuming no current loss).

Substituting for $\frac{dR}{dx}$ and moving $\frac{1}{dx}$ from the LHS to the RHS we obtain $$dV = I_D \frac{dR}{dx} \, dx = I_D \frac{1}{\mu_n WQ'(x)} \, dx$$
Derivation of Square Law Model

We ended up with a varying effective resistance along the length of the channel. To solve for the current, integration will be used. By substituting for the position-dependent charge density, \( Q_I(x) \), and rearranging further, the following is obtained:

\[
I_D dx = \mu_n W C_{ox} [V_{GS} - V_{TN} - V(x)] dV
\]

Now, integrate position on LHS and voltage on RHS and use the boundary conditions provided earlier.

\[
\int_{x=0}^{L} I_D dx = \int_{V=0}^{V_{DS}} \mu_n W C_{ox} [V_{GS} - V_{TN} - V(x)] dV
\]

\[
\int_{x=0}^{L} I_D dx = \mu_n W C_{ox} \left( \int_{V=0}^{V_{DS}} [V_{GS} - V_{TN}] dV - \int_{V=0}^{V_{DS}} [V(x)] dV \right)
\]

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

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

rearranging terms we have the final square law model for non-saturated operation:

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

note that the derivation employs the gradual channel approximation which assumes that the diffusion current due to carrier gradients \( \frac{dV}{dx} / \frac{dR}{dx} \) is small compared to the drift current from the electric field \(-D \frac{dn}{dx}\). This is not true in subthreshold operation, where diffusion current dominates the drift current.
Non saturation Square Law Model (Above Threshold)

Summary

Variations include the following:

\[ I_{Dn} = k_n' \frac{W}{L} \left( (V_{GS} - V_{TN}) V_{DS} - \frac{1}{2} V_{DS}^2 \right) = k_n' \frac{W}{L} \frac{1}{2} \left( 2 (V_{GS} - V_{TN}) V_{DS} - V_{DS}^2 \right) \]

\[ I_{Dn} = \beta \left( (V_{GS} - V_{TN}) V_{DS} - \frac{1}{2} V_{DS}^2 \right) = \beta \frac{1}{2} \left( 2 (V_{GS} - V_{TN}) V_{DS} - V_{DS}^2 \right) \]
The BSIM model was mainly designed for strong inversion (above threshold) design. The EKV model is a well-behaved bulk-referenced model for weak inversion (subthreshold) and moderate inversion design.

**EKV Model Equation**

\[
I = \frac{K}{2\kappa_{x}} (2\phi_{t})^{2} \left[ \log^{2} \left( 1 + e^{\frac{\kappa(V_{g}-V_{T})-V_{s}}{2\phi_{T}}} \right) - \log^{2} \left( 1 + e^{\frac{\kappa(V_{g}-V_{T})-V_{d}}{2\phi_{T}}} \right) \right];
\]

\[
K = \frac{W}{L} \mu C'_{ox}
\]

: \( I \) is a superposition of forward current and reverse current components
EKV Model (2)

- **Gate Voltage (V):** The x-axis represents the gate voltage in volts (V).
- **IDS (log scale shown):** The y-axis represents the drain-source current (IDS) on a log scale.

**Points:**
- **Depletion:** The region where the current is very low, typically below the detection limit.
- **Weak Inversion:** A region where the current starts to increase slightly as the gate voltage increases.
- **Moderate Inversion:** Region where the current increases more significantly with the gate voltage.
- **Strong Inversion:** Region where the current continues to increase rapidly as the gate voltage increases.

**Additional Gate Charge:**
- Matched mostly by additional depletion charge QB.
- Transition: increase in QB begins to surpass increase in QI at maximum QB.

**Strong Inversion Charge Model:**
- Does not hold at extrapolated intercept that determines VTO.

**Equations:**
- \( \sqrt{I_{DS}} \)
- \( \frac{\kappa}{kT/q} \)
- \( \Delta Q_C = \Delta Q_B \)
- \( \Delta V_g \)

**Symbols:**
- \( V_{TO} \)
- \( C_{ox} \)
- \( Q_B \)

**Interpretation:**
- The graph illustrates the behavior of drain-source current (IDS) as a function of gate voltage, highlighting different regions of operation and charge models associated with it.
Let $x$ be the exponentiated term:
\[
x = \frac{\kappa(V_g - V_T) - V_s}{2\phi_T}
\]

\[
\log^2 \left( 1 + e^{\frac{\kappa(V_g - V_T) - V_s}{2\phi_T}} \right) \approx \begin{cases} 
\left( \frac{\kappa(V_g - V_T) - V_s}{2\phi_T} \right)^2 & \text{when } x \gg 0 \\
\frac{\kappa(V_g - V_T) - V_s}{2\phi_T} e^{\frac{\kappa(V_g - V_T) - V_s}{2\phi_T}} & \text{when } x < 0
\end{cases}
\]

At one extreme, the EKV model behaves like a square law model and on the other it behaves like an exponential function much like a BJT.

$V_{T0}$ is found by extrapolating the $\sqrt{l}$ curve in the strong inversion region to 0A.
\[ i_{DS} = \frac{K}{2\kappa} (2\phi_t)^2 \left[ \left( \frac{\kappa(V_{GB} - V_T) - V_{SB}}{2\phi_T} \right)^2 - \left( \frac{\kappa(V_{GB} - V_T) - V_{DB}}{2\phi_T} \right)^2 \right] \]

When \( V_{DS} \) is large (i.e. \( V_{DB} \ll V_{SB} \)), the forward component (source) dominates the reverse component (drain).

\[ i_{DS} = \frac{K}{2} \left( (V_g - V_T) - \frac{1}{\kappa} Vs \right)^2 = \]
\[ \frac{K}{2} \left( (V_g - [Vs] - V_T) - (\frac{1}{\kappa} - [1]) Vs \right)^2 \]

\[ V_{DSSAT} = \kappa (v_{GB} - V_{T0}) \odot V_{SB} = 0 \text{ (no “body-effect”) } \]

\[ i_{DSSAT} = \frac{\kappa K}{2} (V_{GB} - V_{T0})^2 \]
Source-reference models require an adjustment parameter when the source $V_{SB}$ changes. They cover this change by modifying the threshold voltage parameter according to $V_{SB}$.

Here a common source-referenced above threshold model:

**Source-Referenced Model**

$$I = \frac{K}{2} \left(V_g - V_s - V_T\right)^2 \text{with } V_T = V_{To} + \gamma \left(\sqrt{\phi_0 + V_{SB}} - \sqrt{\phi_0}\right);$$

$$\phi_0 \approx 2\phi_f + \Delta \phi_o; \Delta \phi_o \approx 6\phi_t \text{ for above threshold fitting but often just } 2\phi_f \text{ is used for } \phi_0$$

Source Referenced Model “Body Effect” (threshold voltage adjustment)

**Source-Referenced Model $V_T$ adjustment termed “Body Effect”**

$$\frac{\partial V_T}{\partial V_s} = \gamma \frac{1}{2} \frac{1}{\sqrt{\phi_0 + V_{SB}}}$$

@ $V_{SB} = 0, \frac{\partial V_T}{\partial V_s} = \gamma \frac{1}{2} \frac{1}{\sqrt{\phi_0}}$
The bulk-referenced EKV model does not require concept of modifying $V_T$ for a “body-effect” when the source changes with respect to the bulk as source referenced models do, this is already modeled through the term $(\frac{1}{\kappa} - [1]) V_s$.

The more stationary $V_T$ in the EKV model is like stationary $V_{T0}$ in the source-referenced model, let’s use $V_{T0}$ for just a moment:

\[ I = \frac{K^2}{2} \left( V_g - V_s - V_{T0} - \left( \frac{1}{\kappa} - 1 \right) V_s \right)^2 \text{(EKV model)} \]

\[ I = \frac{K^2}{2} \left( V_g - V_s - \left( V_{T0} - \left( \frac{1}{\kappa} - 1 \right) V_s \right) \right)^2 \text{(source-referenced model "} V_T \text{")} \]

\[ \frac{\partial}{\partial V_s} V_T'' = \frac{1}{\kappa} - 1 \text{ (change in source-referenced "} V_T \text{" due to } V_{SB}) \]
Comparing this result using our bulk-referenced model with the result from the source-referenced model:

**Relationship of \( \kappa \) and the more familiar \( \gamma \)**

\[
\frac{1}{\kappa} - 1 = \gamma \frac{1}{2} \sqrt{\phi_0}
\]

Alternatively you may see a parameter \( n \):

\[
1 + \frac{\gamma}{2 \sqrt{2\phi_f + V_{SB}}} = \frac{1}{\kappa} = n
\]
Non Saturation Equation

\[ i_{DS} = I_0 e^{\frac{\kappa v_{GB}}{\phi_t}} \left( e^{\frac{-v_{SB}}{\phi_t}} - e^{\frac{-v_{DB}}{\phi_t}} \right); \text{(non-saturation equation)} \]

\[ I_0 = \mu C_{ox} \phi_t^2 \frac{W}{L} \left( \frac{1 - \kappa_{sa}}{\kappa_{sa}} \right) e^{\frac{-\kappa V_{T0}}{\phi_t}} \] [for \( \kappa_a \) see page 101 of Sarpeshkar, Ultra Low Power Bioelectronics]

\[ V_{T0} = V_{FB} + \phi_0 + \gamma \sqrt{\phi_0}; \phi_0 = 2\phi_f \]

may pull \( e^{\frac{-v_{SB}}{\phi_t}} \) term out of parenthesis:

\[ i_{DS} = I_0 e^{\frac{\kappa v_{G}-v_{S}}{\phi_t}} \left( 1 - e^{\frac{-v_{DS}}{\phi_t}} \right) \] (alternate form)

Saturation Equation

when \( \frac{v_{DS}}{\phi_t} \gg 0 \), typically \( V_{ds} = 4\phi_t \approx 100 \text{ mV} \), the forward current dominates and drain voltage changes have little significance:

\[ i_{DS} = I_0 e^{\frac{\kappa v_{G}-v_{S}}{\phi_t}} \left( 1 - e^{\frac{-v_{DS}}{\phi_t}} \right) \approx I_0 e^{\frac{\kappa v_{G}-v_{S}}{\phi_t}} \] (saturation equation)
At some voltage the non-saturated square-law-model equation predicts a peak and then a decrease in current, which does not happen in reality. The peak is identified by setting the derivative to zero:

\[ \frac{\partial I_{Dn}}{\partial V_{DS}} = 0 \rightarrow V_{DS,\text{sat}} = V_{GS} - V_T \]

Through the current equation, \( V_{DS,\text{sat}} \) determines a peak current called the saturation current which is used instead for \( V_{DS} > V_{DS,\text{sat}} \)

### Saturation Current

\[ I_{Dn,\text{sat}} = \frac{k'}{2} \frac{W}{L} (V_{GS} - V_{Tn})^2 = \frac{\beta}{2} (V_{GS} - V_{Tn})^2 \]
In summary thus far, (ignoring the yet to be discussed channel length modulation) we have

\[ I_{Dn} = \begin{cases} k_n \frac{W}{L} \frac{1}{2} \left( 2 (V_{GS} - V_{TN}) V_{DS} - V_{DS}^2 \right), & \text{if } V_{DS} \leq V_{DS,sat} \\ \frac{k'}{2} \frac{W}{L} (V_{GS} - V_{Tn})^2 = I_{Dn,sat} & \text{if } V_{DS} \geq V_{DS,sat} \end{cases} \]
A better approximation when $V_{DS} > V_{DS,sat}$ takes into account channel length modulation. As the drain voltage increases, a depletion region grows around the drain, pulling carriers out of the inversion layer, and the inversion layer is reduced to a shorter length.

Assume $L$ is a function of $V_D$:

$$L_{M,(V_D)} = L - L_R$$

(reduction in length due to drain voltage)

Reexamine current with and without modulation:

- No modulation: $I_{sat} = \frac{k'L}{2} \frac{W}{L} (V_{GS} - V_{Tn})^2$
- With modulation: $I_{sat,M} = \frac{k'L}{2} \frac{W}{L_{M,(V_D)}} (V_{GS} - V_{Tn})^2$

A drain-dependant saturation-mode current:

$$I_{Dn,sat}(V_D) = I_{Dn,sat}\left(\frac{L}{L_{M,(V_D)}}\right) = I_{Dn,sat}\left(\frac{L}{L-L_R}\right)$$
let $L_R$ be defined by a parameter $r$: 
$$L_R = r \times (V_D - V_{D,sat})$$

Examine the rate of change of current:
$$\frac{dl_{Dn,\text{sat}}(V_D)}{dV_D} = \frac{dl_{Dn,\text{sat}}(V_D)}{dL_{M,\text{sat}}(V_D) \times dV_D} = \left( -\frac{k'}{2} \frac{W}{L_{M,\text{sat}}(V_D)} (V_{GS} - V_{Tn})^2 \right) \times (-r)$$

$$\frac{dl_{Dn,\text{sat}}(V_D)}{dV_D} = I_{Dn,\text{sat}}(V_D) \times \frac{r}{L_{M,\text{sat}}(V_D)}$$

We will capture the proportional (in the sense of $\frac{\Delta L}{L}$) rate of reduction of $L$ w.r.t. drain voltage by a new parameter $\lambda$:
$$\lambda = \frac{r}{L} \text{ (depends on initial length)}$$
when $\lambda$ is smaller, $\frac{\Delta L}{L}$ tends to be smaller

Rate of change of current:
$$\frac{dl_{Dn,\text{sat}}(V_D)}{dV_D} = I_{Dn,\text{sat}}(V_D) \times \lambda \text{ ; (sensitive to saturation current and length)}$$
A drain-dependant saturation-mode current:

- \( I_{Dn,\text{sat}}(V_D) = I_{Dn,\text{sat}}\left(1 + \lambda (V_{DS} - V_{DSSAT})\right) \)
- Continuous with non-sat square-law model, meets at \((V_{DSSAT}, I_{Dn,\text{sat}})\)

Other variations:
These are continuous with non-sat models only if the non-sat models also account for channel length modulation (the first derivative is continuous as well)

\[
I_D = \frac{k'}{2} \frac{W}{L} (V_{GS} - V_{Tn})^2 (1 + \lambda V_{DS})
\]

or use \( e^{V_{DS}/V_A} \approx (1 + V_{DS}/V_A) \) for small \( V_{DS} \)
Graphically, we can identify a point where the $I_{DS}$ curves intersect the $V_{DS}$ axis and refer to it as $V_A$, the Early voltage. Then, $\lambda \approx \frac{1}{V_A}$

For high $V_{DS}$ voltages, even this model does not apply. Longer devices have lesser proportional change in the length and have a lower $\lambda$

For hand calculations there are two common approaches:

1. Provided a given $\lambda$ for a given $L$, scale $\lambda$ inversely proportional to the actual $L$
2. Assume a fixed $\lambda$ regardless of $L$
In saturation, the right-most point of the inversion depiction always represents the same potential, as this point moves to the left it represents a shortening channel (smaller resistance).

In the view below, the right-side capacitor segment is so depleted of any significant amount of carriers inversion and only sees carriers swiftly moving by. Recall segments have the same charge flow.
Rather than use the boundary of the two equations to be \( V_{DS} = V_{DS,\text{sat}} = V_{GS} - V_{TN} \), we can use \( V_{GB} - V_{DB} = V_{TN} \). As the drain is increased from 0 V, if the drain comes within a threshold voltage of the gate, there is no longer inversion at the drain end of the channel and we must use the saturation equation. Not, this is in contrast to use of a definition of \( V_{DS_{\text{Sat}}} \).
Saturation Occurs when $V_{GB} - V_{DB} > V_{TN}$

\[ I_{Dn, sat} = \frac{k'}{2} \frac{W}{L} (V_{GS} - V_{Tn})^2 = \frac{\beta}{2} (V_{GS} - V_{Tn})^2 \]

\[ I_{Dn} = I_{Dn, sat} (1 + \lambda (V_{DS} - V_{DSSAT})); \lambda = \frac{1}{V_A} \]
As the reverse bias voltage of drain increases, the effective length of the channel region shortened. $\frac{\Delta L}{\Delta V_D}$ is negative.

The proportional effect on the drain current is inversely proportional to the length of a channel (for longer channels the effect on the current is small): $\frac{\Delta L}{L}$ is smaller for larger L.
Above Threshold:

- \( I_D = k'_n \frac{W}{L} \frac{1}{2} (2(V_{GS} - V_{TN}) V_{DS} - V_{DS}^2) = k'_n \frac{W}{L} \frac{1}{2} ((2(V_{GS} - V_{TN}) - V_{DS}) V_{DS}) \)
- for \( V_{DS} \ll 2(V_{GS} - V_{TN}), \)
  \[ I_D \approx k'_n \frac{W}{L} \frac{1}{2} (2(V_{GS} - V_{TN}) V_{DS}) = \beta (2(V_{GS} - V_{TN}) V_{DS}) \]

Looking at Ohm’s law as a template, in the triode/linear region of operation the device behaves like gate-voltage-controlled resistor between the source and drain

- \( R_{DS,\text{triode}} = \frac{V_{DS}}{I_D} = \frac{1}{k'_n \frac{W}{L} (V_{GS} - V_{TN})} = \frac{1}{\beta (V_{GS} - V_{TN})} \)
Temperature Effects

- \( \mu(T) = \mu(T_r) \left( \frac{T}{T_r} \right)^{-k_3} \) [Tsividis, Operation and Modeling of the MOS Transistor, pg 224]

- “\( V(T) = V_T(T_r) = k_4(T - T_r) \) where \( k_4 \) is usually between 0.5 mV/K and 3 mV/K, with larger values in this rage corresponding to heavier doped substrates, thicker oxides, and smaller values of \( V_{SB} \)” [Tsividis, Operation and Modeling of the MOS Transistor, pg 225]
$L_{\text{eff}} = L_{\text{drawn}} - 2LD$; LD is the amount of manufacturing lateral diffusion that reduced the channel length at each end. There may also be a manufacturing adjustment parameter given for the width accounting for field oxide encroachment.

$W_{\text{eff}} = W_{\text{drawn}} - 2W_{\text{FOE}}$
Drawn Vs Effective W and L (2)

Top View Drawing

Select Mask (n-select or p-select)

diffusion opening

polysilicon

Side Cross Section

Cross Section

Select Mask

$W_{\text{eff}} = W_{\text{Drawn}} - 2W_{\text{FDE}}$

Top View Drawing

Trasistor as Drawn

Direction of Current Flow

Channel area as drawn

Birds Peak
Additional Gate to bulk Capacitance provided per-unit-length units (rather per-unit width)

Reduction of Drawn Channel

Gate-Source/Gate-Drain Overlap provides overlap capacitance $C_{\text{ox}}$, provided through a per-unit-width parameter

*note, actual encroachment into the channel and into the junctions may be different
**Low-Frequency Small-Signal Models (using unilateral model for Early Effect)**

![Circuit Diagram]

<table>
<thead>
<tr>
<th></th>
<th>$r_\pi$</th>
<th>$g_m$</th>
<th>$g_s$</th>
<th>$r_0$</th>
<th>$g_m r_0$</th>
<th>$g_m r_\pi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>sub-VT mos (sat)</td>
<td>$\infty$</td>
<td>$\frac{\kappa i}{\phi_t} = \kappa g_s$</td>
<td>$\frac{i}{\phi_t}$</td>
<td>$\frac{V_A}{I}$</td>
<td>$\frac{V_A}{\phi_t}$</td>
<td>$\infty$</td>
</tr>
<tr>
<td>above-VT mos ($\kappa \approx 1$)(sat)</td>
<td>$\infty$</td>
<td>$\sqrt{2KI} = \frac{2I}{V_g - V_s - V_T} = \frac{2I}{V_{on}}$</td>
<td>$g_m$</td>
<td>$\frac{V_A}{I}$</td>
<td>$\frac{2V_A}{V_g - V_s - V_T} = \frac{2V_A}{V_{on}}$</td>
<td>$\infty$</td>
</tr>
<tr>
<td>BJT (fwd. active)</td>
<td>$\frac{\beta\phi_t}{I_C}$</td>
<td>$\frac{i_C}{\phi_t}$</td>
<td>$g_m$</td>
<td>$\frac{V_A}{I_C}$</td>
<td>$\frac{V_A}{\phi_t}$</td>
<td>$\beta$</td>
</tr>
</tbody>
</table>