NCN@Purdue - Intel Summer School: July 14-25, 2008

Physics of Nanoscale Transistors: Lecture 3B:

Theory of the Ballistic MOSFET

Mark Lundstrom

Network for Computational Nanotechnology

Purdue University

West Lafayette, Indiana USA

NCN nanoHUB.org

PURDUE UNIVERSITY

outline

1) **Review**
2) Effective masses
3) Multiple subbands
4) 2D electrostatics
5) Discussion
6) Summary

review

1) In Lecture 3, we generalized

$$I_D = WC_{ox}\left(V_{GS} - V_T\right)\upsilon_T\left(\frac{1 - e^{-qV_{DS}/k_BT}}{1 + e^{-qV_{DS}/k_BT}}\right)$$

to include Fermi-Dirac statistics:

$$I_D = WC_{ox}\left(V_{GS} - V_T\right)\upsilon_T\left[\frac{1 - \mathcal{F}_{1/2}\left(\eta_{F2}\right)/\mathcal{F}_{1/2}\left(\eta_{F1}\right)}{1 + \mathcal{F}_{0}\left(\eta_{F2}\right)/\mathcal{F}_{0}\left(\eta_{F1}\right)}\right]$$

review

2) We also discussed key device parameters for ballistic MOSFETs:

  -ballistic injection velocity
  -ballistic on-current
  -ballistic channel resistance
  -ballistic drain saturation voltage
  -ballistic mobility

what approximations did we make?

1) semiclassical approach (no quantum transport)
2) used a bulk $E(k)$ within the device
3) assumed parabolic energy bands
4) ignored scattering
5) …

questions

1) How do we treat more realistic band structures (e.g. the conduction band of Si?)

2) How to we treat subthreshold conduction and 2D electrostatics?

These questions are addressed in this lecture (Part 2)

outline

1) Review
2) **Effective masses**
3) Multiple subbands
4) 2D electrostatics
5) Discussion
6) Summary

about effective masses

$$N_{2D} = \frac{m^* k_B T}{\pi \hbar^2} \ \#/\text{cm}^2 \qquad\qquad \upsilon_T \equiv \sqrt{\frac{2k_B T}{\pi m^*}} \frac{\mathcal{F}_{1/2}\left(\eta_{F1}\right)}{\mathcal{F}_0\left(\eta_{F1}\right)}$$

For carrier densities, we should use a <u>density-of-states effective mass</u>.

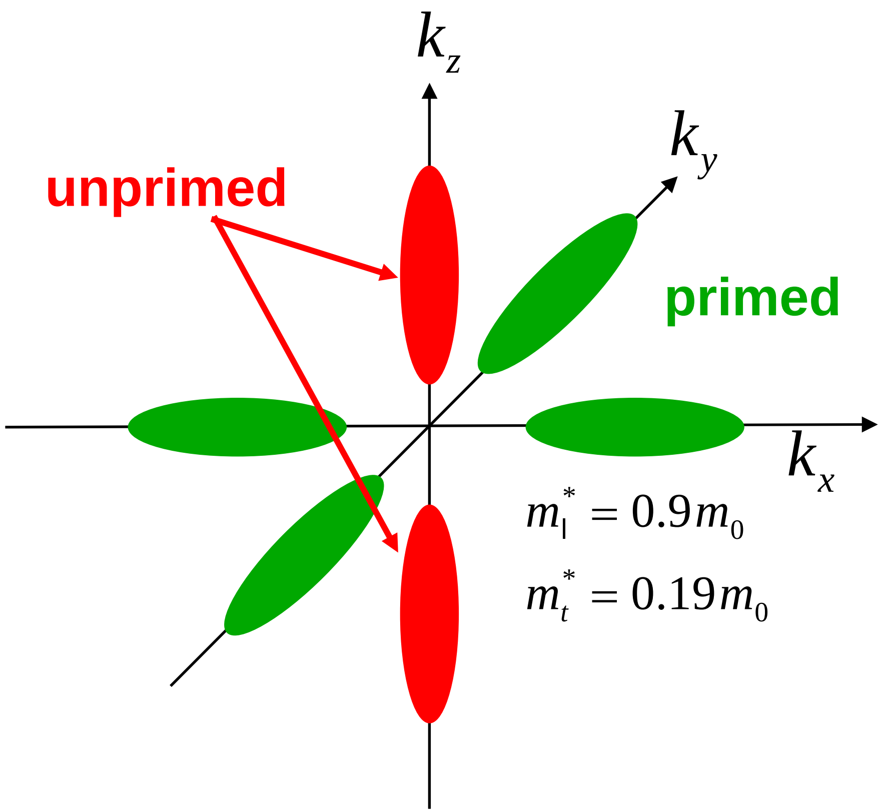
For velocities we should use a <u>conductivity effective mass</u>.

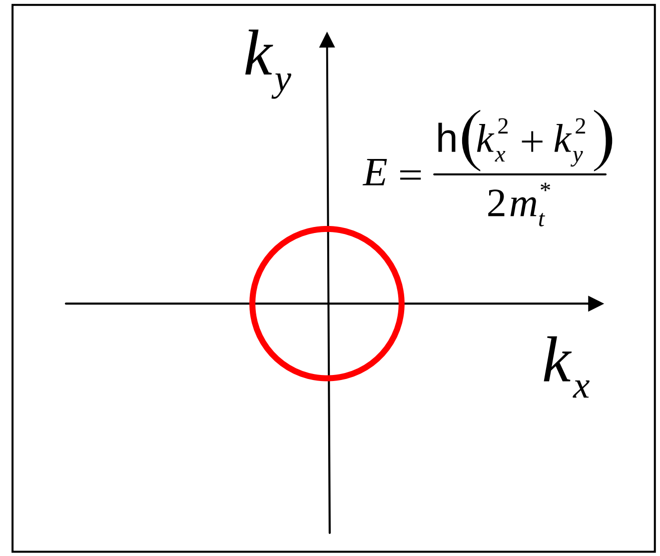
The values of these masses depend on the material and the channel orientation.

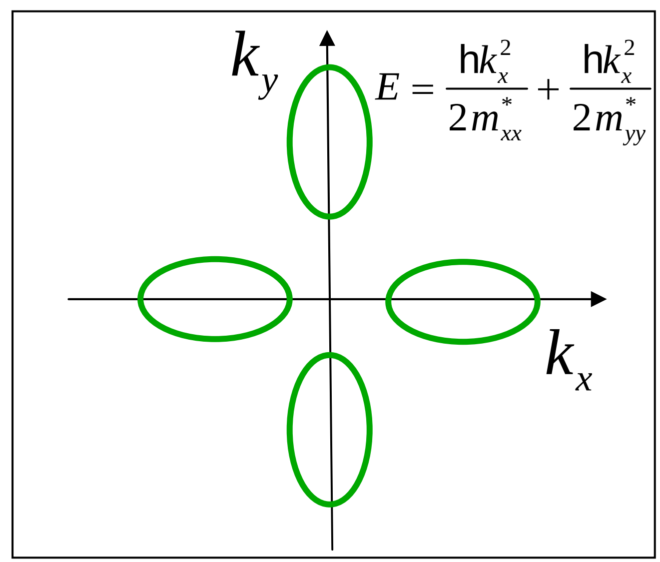
Let’s work this out for the simplest case: (100) Si.

example: (100) Silicon

Si conduction band







confinement masses: $m_{conf}^* = m_{zz}^*$

unprimed: $m_{conf}^* = m_l^*$ primed: $m_{conf}^* = m_t^*$

example: ellipsoidal energy bands

To see how we determine the DOS and conductivity effective masses, let's re-examine our derivation from the beginning - with the drain current:

$$I_D = \frac{2q}{h} \int_{\varepsilon_1(0)}^{\infty} M(E)\left(f_1 - f_2\right) dE$$

So we begin with $M(E)$.

NCN nanoHUB.org

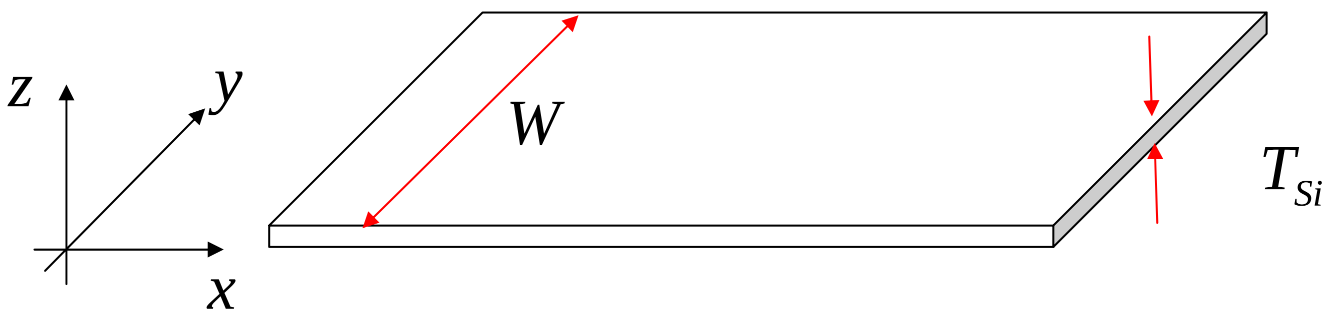
conducting channels: (100) Si [100] transport

$$g_{1D}(E) = g_V \times \frac{1}{\pi\hbar}\sqrt{\frac{m_{yy}^*}{2\left[E - \varepsilon_1(0)\right]}} \quad \text{\#/eV-cm}$$

(Divided by 2 to account for the fact the spin has already been included in the current formula.)

$$M(E) = W\int_{\varepsilon_1(0)}^{E} g_{1D}(E)\,dE = Wg_V\,\frac{\sqrt{2m_{yy}^*\left[E - \varepsilon_1(0)\right]}}{\pi\hbar}$$

M(*E*) = number of transverse modes in the y-direction with cut-off energy less than *E.*



[100] transport direction

current: (100) Silicon [100] transport

To find the drain current, we integrate:

$$I_D = \frac{2q}{h} \int_{\varepsilon_1(0)}^{\infty} M(E)\left(f_1 - f_2\right) dE$$

$$I_D = Wqg_V \left(\frac{m_{yy}^* k_B T}{2\pi\hbar^2} \right) \sqrt{\frac{2k_B T}{\pi m_{yy}^*}} \left[\mathcal{F}_{1/2}\left(\eta_{F1}\right) - \mathcal{F}_{1/2}\left(\eta_{F2}\right) \right]$$

The drain current depends on m_{yy}, which is different for each type of valley.

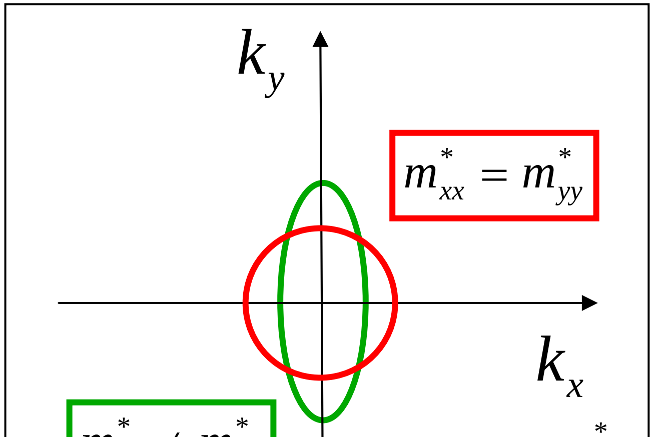
recall: carrier densities in 2D

To find the 2D carrier density, we integrate:

$$n_S^+(0) = \int_{\varepsilon_1(0)}^{\varepsilon_1(top)} \frac{D_{2D}(E)}{2} f_0(E_{F1})\,dE$$

$$n_S^+(0) = \frac{N_{2D}}{2} \mathcal{F}_0(\eta_{F1})$$

$$N_{2D} = \frac{m_D^* k_B T}{\pi\hbar^2} \ \#/\mathrm{cm}^2$$



$$D_{2D}(E) = \frac{m_D^*}{\pi\hbar^2}$$

$$D_{2D}(E) = g_V \frac{\sqrt{m_{xx}^* m_{yy}^*}}{\pi\hbar^2}$$

$$m_D^* = g_V \sqrt{m_{xx}^* m_{yy}^*}$$

NCN nanoHUB.org

example: (100) Silicon [100] transport

Drain current:

$$I_D = Wqg_V \left(\frac{m_{yy}^* k_B T}{2\pi\hbar^2} \right) \sqrt{\frac{2k_B T}{\pi m_{yy}^*}} \left[\mathcal{F}_{1/2}(\eta_{F1}) - \mathcal{F}_{1/2}(\eta_{F2}) \right]$$

Carrier density:

$$n_S(0) = \frac{N_{2D}}{2} \left[\mathcal{F}_0(\eta_{F1}) + \mathcal{F}_0(\eta_{F2}) \right]$$

$$N_{2D} = \frac{m_D^* k_B T}{\pi\hbar^2} \ \#/\mathrm{cm}^2$$

$$\boxed{m_D^* = g_V \sqrt{m_{xx}^* m_{yy}^*}}$$

Velocity:

$$\langle \upsilon \rangle = \frac{I_D}{Wqn_S} = \sqrt{\frac{2k_B T}{\pi m_{xx}^*}} \frac{\left[\mathcal{F}_{1/2}(\eta_{F1}) - \mathcal{F}_{1/2}(\eta_{F2}) \right]}{\left[\mathcal{F}_0(\eta_{F1}) + \mathcal{F}_0(\eta_{F2}) \right]}$$

$$\boxed{m_C^* = m_{xx}^*}$$

NCN nanoHUB.org

example: (100) Silicon [100] transport

the unprimed subbands will be the lowest because: $m^*_{conf} = m^*_l$

$$m^*_{xx} = m^*_{yy} = m^*_t = 0.19 m_0 \qquad m^*_D = 2m^*_t \qquad m^*_C = m^*_t$$

the primed subbands are higher in energy because: $m^*_{conf} = m^*_t$

$$m^*_{xx} \neq m^*_{yy} \qquad m^*_D = 4\sqrt{m^*_t m^*_l} \qquad m^*_C = 4\left[m_t^{-1/2} + m_l^{-1/2}\right]^{-2}$$

NCN nanoHUB.org

more generally (for ellipsoidal energy bands)

For arbitrary crystallographic orientations with different confinement and transport directions, there will be different degeneracy factors, and different effective masses, $m^*_{xx}, m^*_{yy}, m^*_{zz}$

In each case, the appropriate density-of-states and conductivity effective masses can be obtained.

For example, the standard transport direction for (100) Si is [110]. For the unprimed subbands, things don't change, but for the primed bands, the conductivity mass changes to: $m^*_C = 2m^*_t m^*_l \big/ \left(m^*_t + m^*_l\right)$

NCN nanoHUB.org

references

Anisur Rahman, Mark S. Lundstrom, and Avik Ghosh, "Generalized Effective Mass Approach for Cubic Semiconductor n-MOSFETs on Arbitrarily Oriented Substrates," *J. Appl. Phys*., **97**, 053702, March 1, 2005.

Marco De Michielis, David Esseni, and Francesco Driussi, "Analytical Models for the Insight Into the Use of Alternative Channel Materials in Ballistic nano-MOSFETs," *IEEE Trans. Electron Dev*., **54** (1), pp. 115-123, 2007.

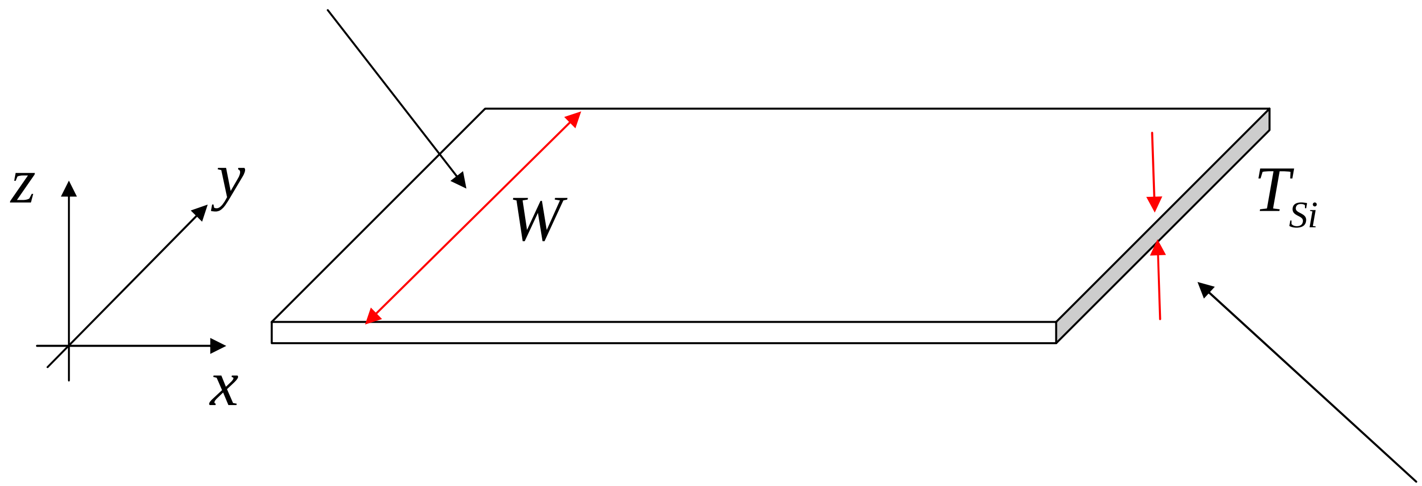
NCN nanoHUB.org

outline

1) Review
2) Effective masses
3) **Multiple subbands**
4) 2D electrostatics
5) Discussion
6) Summary

subbands and modes

For each subband associated with confinement in the z-direction, there is also a set of subbands (modes) associated with confinement in the y-direction.

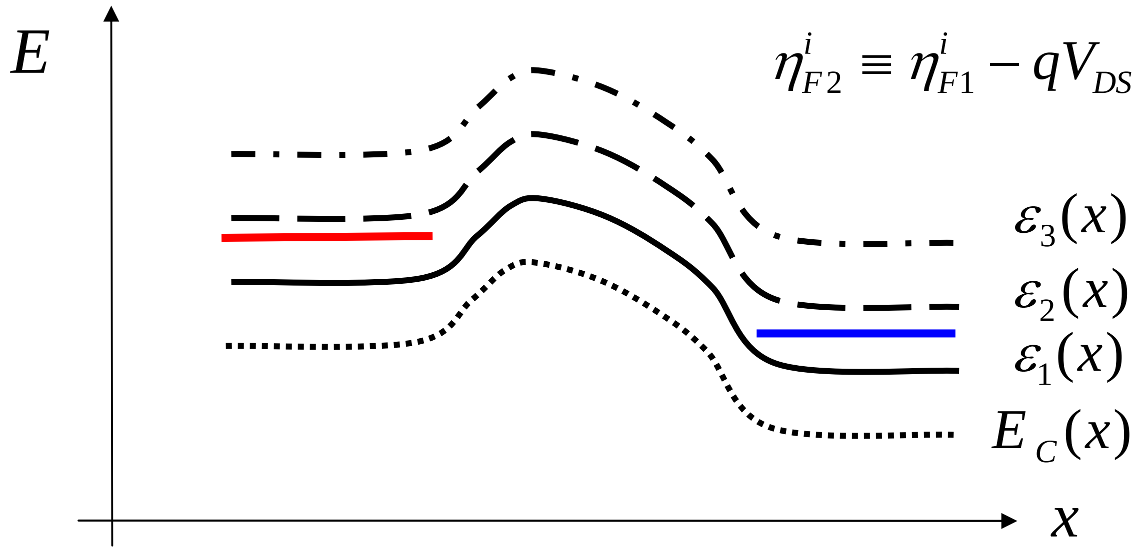


There is a set of subbands associated with confinement in the z-direction.

multiple subbands

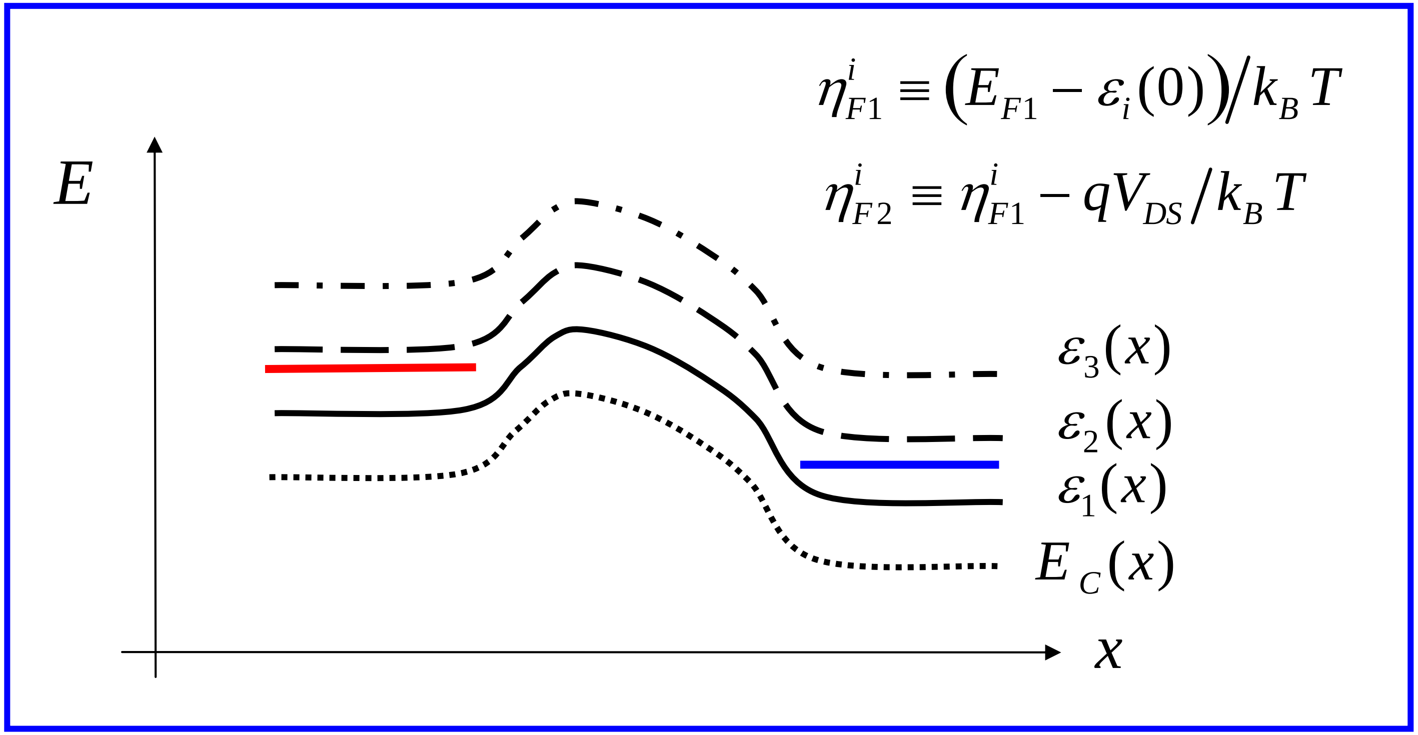
$$\eta_{F1}^{i} \equiv \left(E_{F1} - \varepsilon_i(0)\right)/k_B T$$

$$\eta_{F2}^{i} \equiv \eta_{F1}^{i} - qV_{DS}/k_B T$$



Each subband associated with confinement in the z-direction has many **independent transverse modes** (assuming there is no potential variation in the y-direction, the width of the MOSFET).

multiple subbands



Each subband associated with confinement in the z-direction can be treated as an **independent conduction channel** (with many transverse modes) as long as the potential variation in the x-direction is gentle.

treating multiple subbands

For independent subbands, we can simply add up the contributions to the current and carrier density from each subband.

$$n_S(0) = \sum_i \frac{N_{2D}^i}{2}\left[\mathcal{F}_0\left(\eta_{F1}^i\right) + \mathcal{F}_0\left(\eta_{F2}^i\right)\right]$$

$$I_D = Wq\sum_i \left(\frac{N_{2D}^i}{2}\upsilon_T^i\right)\left[\mathcal{F}_{1/2}\left(\eta_{F1}^i\right) - \mathcal{F}_{1/2}\left(\eta_{F2}^i\right)\right]$$

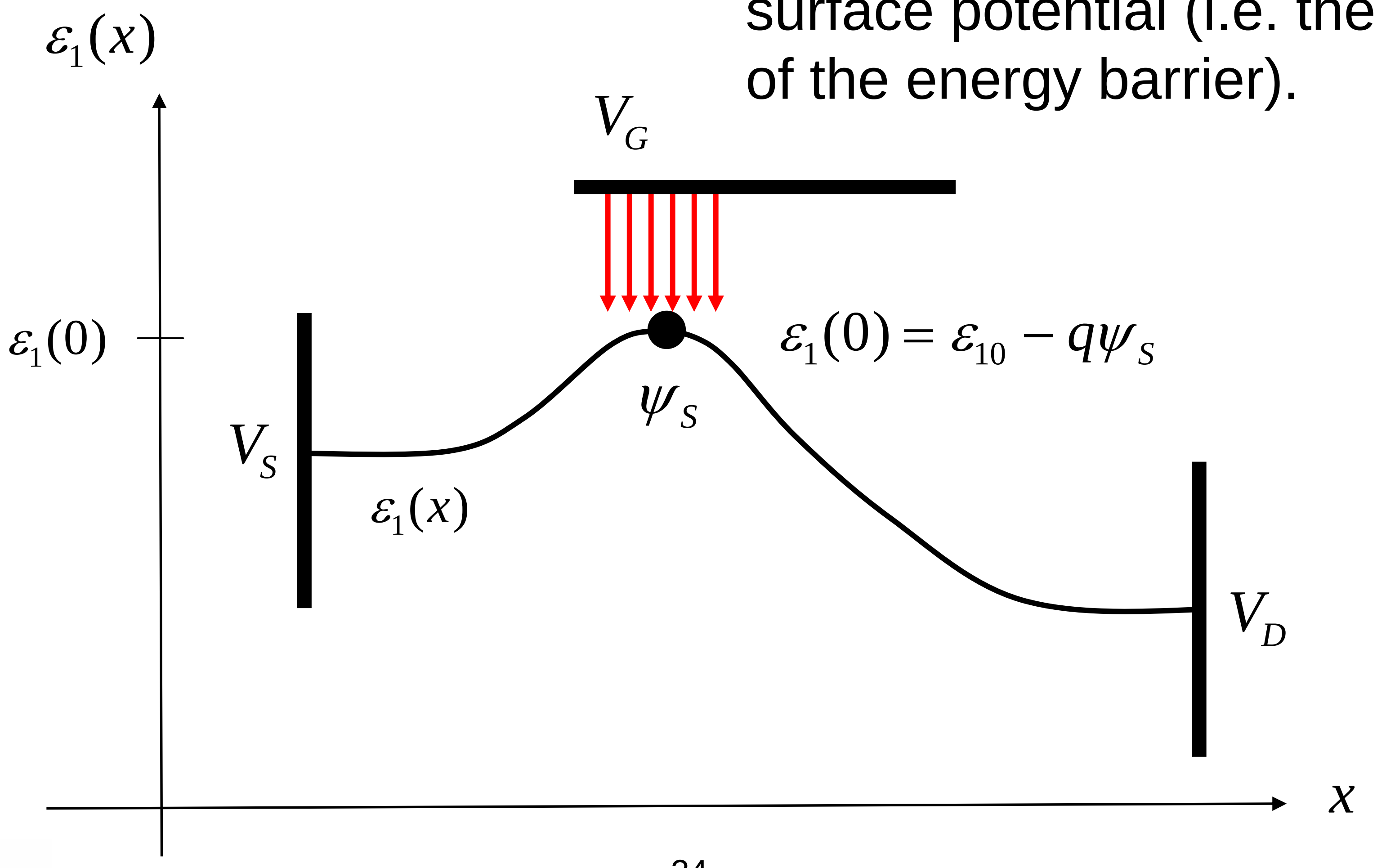
$$\eta_{F1}^i \equiv \left(E_{F1} - \varepsilon_i(0)\right)/k_B T \qquad \eta_{F2}^i \equiv \eta_{F1}^i - qV_{DS}/k_B T$$

outline

1) Review
2) Effective masses
3) Multiple subbands
4) **2D electrostatics**
5) Discussion
6) Summary

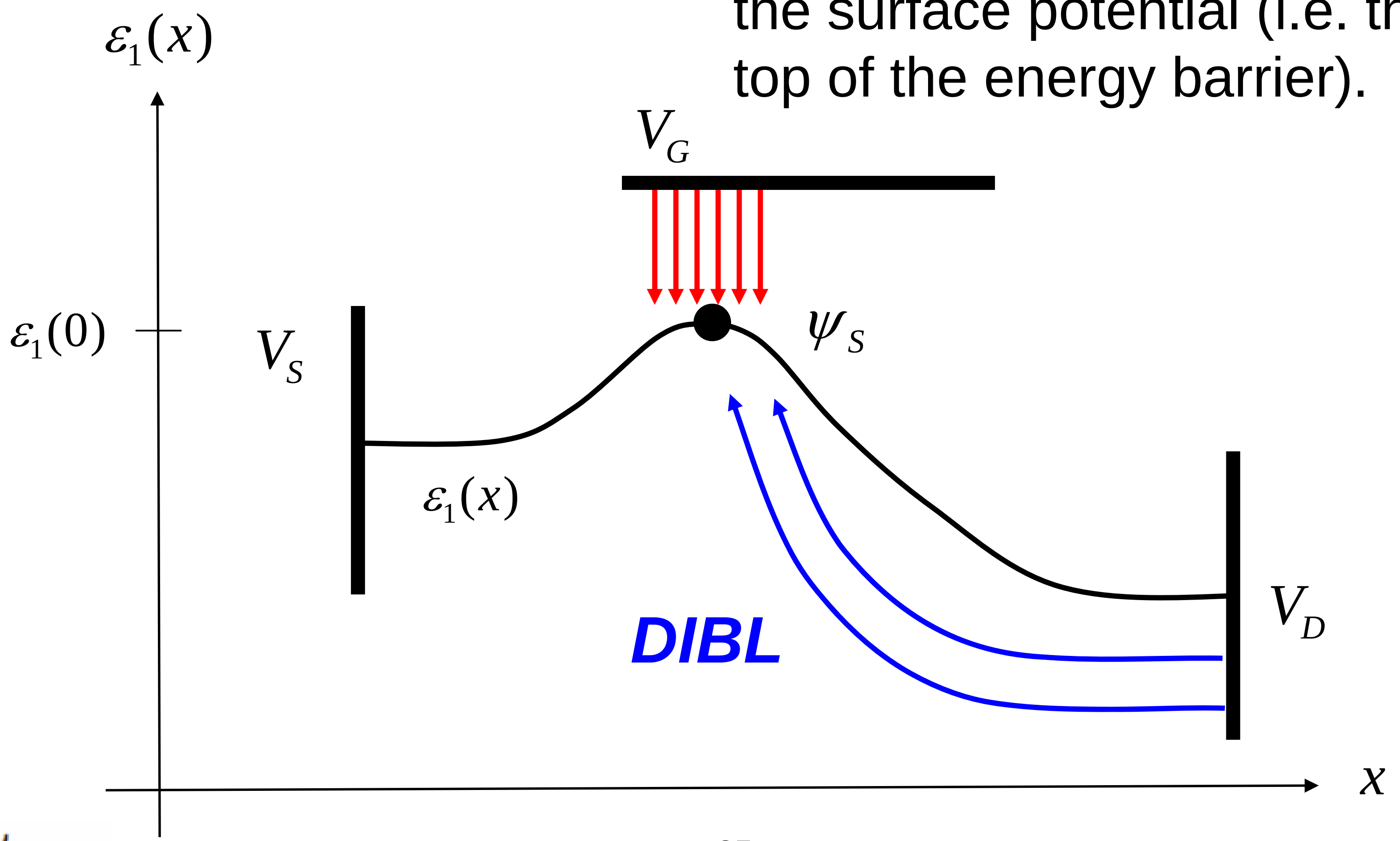
ideal MOS electrostatics

*the **gate voltage*** controls the surface potential (i.e. the top of the energy barrier).

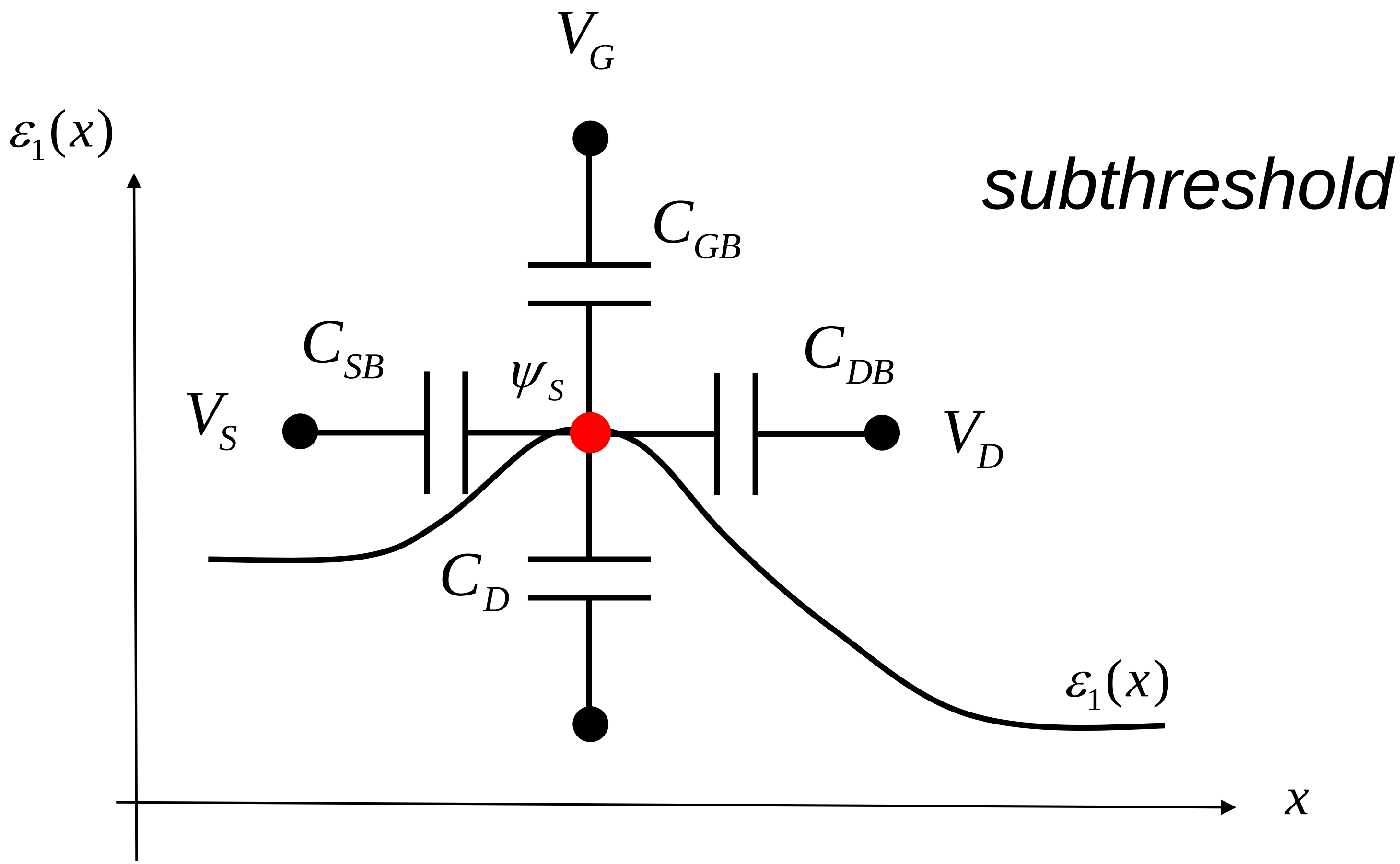


2D MOS electrostatics

the ***drain voltage*** also affects the surface potential (i.e. the top of the energy barrier).



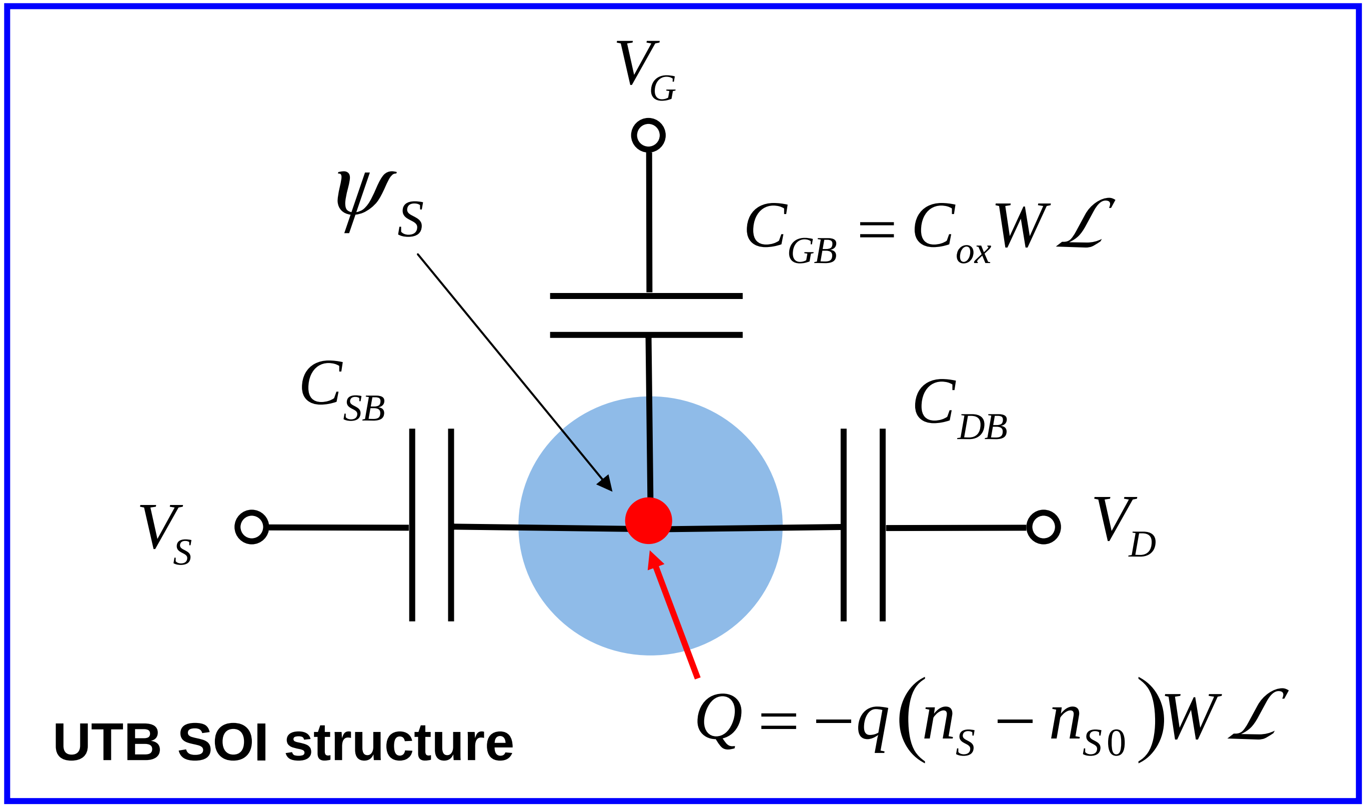
capacitor model for 2D electrostatics



C_D is the depletion capacitance of the semiconductor. If, for example, the MOSFET is an undoped, ultra-thin body structure, then there would be no depletion capacitance.

NCN nanoHUB.org

capacitor model (ii)



$$\psi_S = V_G\left(\frac{C_{GB}}{C_\Sigma}\right) + V_D\left(\frac{C_{DB}}{C_\Sigma}\right) + V_S\left(\frac{C_{SB}}{C_\Sigma}\right) - \frac{q\left[n_S\left(\psi_S\right) - n_{S0}\right] W \mathcal{L}}{C_\Sigma}$$

solving the capacitor model

$$\psi_S = V_G\left(\frac{C_{GB}}{C_\Sigma}\right) + V_D\left(\frac{C_{DB}}{C_\Sigma}\right) + V_S\left(\frac{C_{SB}}{C_\Sigma}\right) - \frac{q\left[n_S\left(\psi_S\right) - n_{S0}\right]W\mathcal{L}}{C_\Sigma} \qquad (1)$$

$$n_S(0) = \frac{N_{2D}}{2}\left[\mathcal{F}_0\left(\eta_{F1}\right) + \mathcal{F}_0\left(\eta_{F2}\right)\right]$$

$$\eta_{F1} \equiv \left[E_{F1} - \varepsilon_1(0)\right]/k_B T$$

$$\eta_{F2} = \eta_{F1} - qV_{DS}/k_B T$$

$$\varepsilon_1(0) = \varepsilon_{10} - q\psi_S$$

(ε_{10} is the value of $\varepsilon_1(0)$ when $\psi_S = 0$
i.e. ‘flatband’ conditions.)

Eqn. (1) is a nonlinear equation for ψ_S. It can be solved by iteration.

solving for I_D

$$\psi_S = V_G\left(\frac{C_{GB}}{C_\Sigma}\right) + V_D\left(\frac{C_{DB}}{C_\Sigma}\right) + V_S\left(\frac{C_{SB}}{C_\Sigma}\right) - \frac{q\left[n_S\left(\psi_S\right) - n_{S0}\right]W\mathcal{L}}{C_\Sigma} \qquad (1)$$

$$\eta_{F1} \equiv \left[E_{F1} - \varepsilon_{10} + q\psi_S\right]/k_B T \qquad \eta_{F2} = \eta_{F1} - qV_{DS}/k_B T$$

$$I_D = Wq\left(\frac{N_{2D}}{2}\upsilon_T\right)\left[\mathcal{F}_{1/2}\left(\eta_{F1}\right) - \mathcal{F}_{1/2}\left(\eta_{F2}\right)\right] \qquad (2)$$

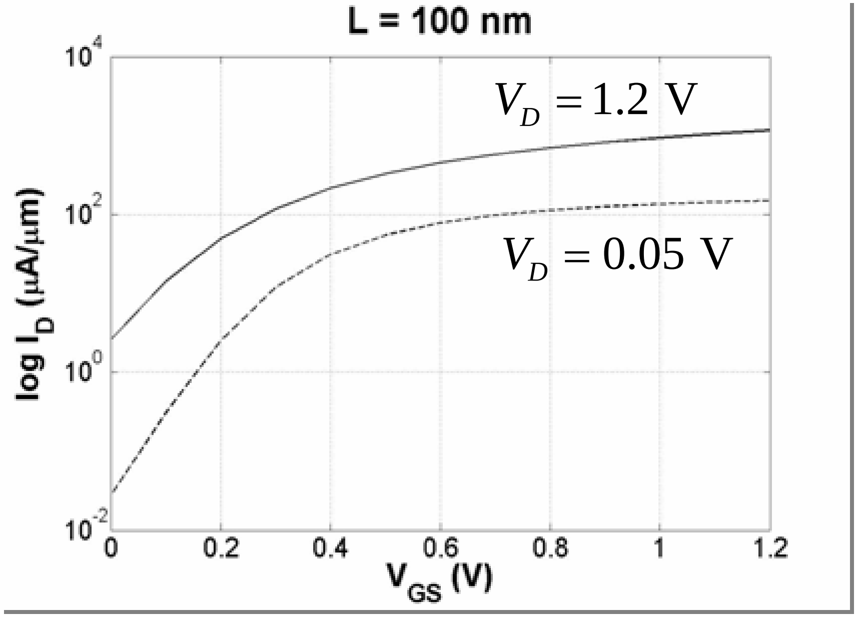
For a given V_G and V_D ($V_S = 0$):

1) Solve (1) for ψ_S

2) Solve (2) for I_D

(~Treats 2D MOS electrostatics, above and below threshold.)

how do we determine the capacitors?



must specify:

C_{GB}

C_{DB}

C_{SB} $\quad \left(\text{or } C_\Sigma = C_{GB} + C_{DB} + C_{SB}\right)$

can do this from measured or simulated data

NCN nanoHUB.org

how do we determine the capacitors?

under high drain bias:

$$I_D \approx Wq\frac{N_{2D}}{2}\upsilon_T \mathcal{F}_{1/2}\left(\eta_{F1}\right)$$

under subthreshold conditions:

$$I_D \approx WqN_{2D}\upsilon_T e^{\eta_{F1}} \sim e^{\left(E_F - \varepsilon_1(0)\right)/k_BT}$$

$$\varepsilon_1(0) = \varepsilon_{10} - q\psi_S$$

$$\boxed{I_D \sim e^{q\psi_S/k_BT}}$$

$$S = 2.3\frac{C_\Sigma}{C_{GB}}\left(k_BT/q\right)$$

$$DIBL = \frac{C_{DB}}{C_{GB}}$$

$$\alpha_G = \frac{C_{GB}}{C_\Sigma} \qquad \alpha_D = \frac{C_{DB}}{C_\Sigma}$$

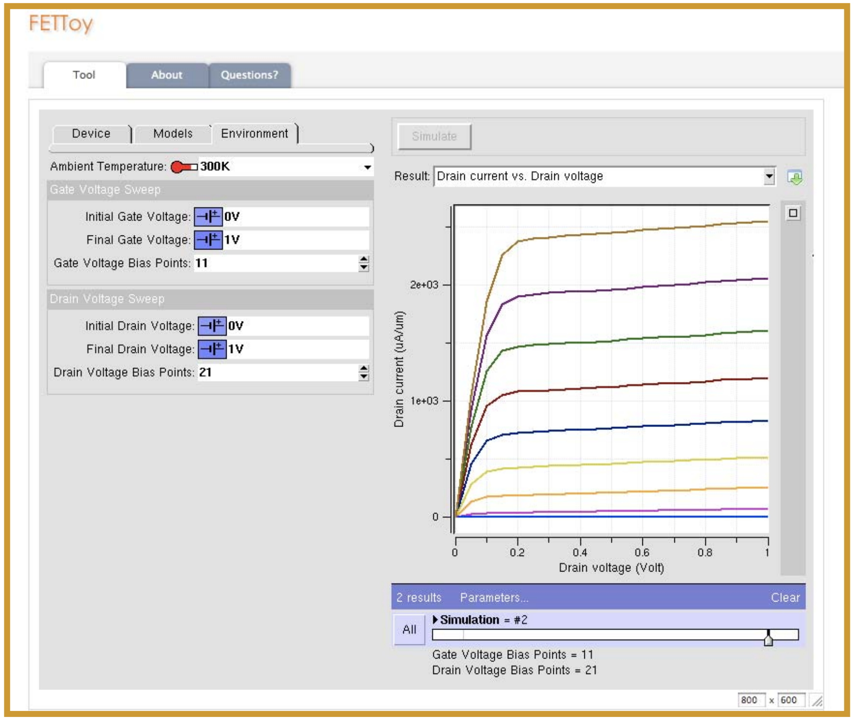
$$S = \frac{2.3\left(k_BT/q\right)}{\alpha_G}$$

$$DIBL = \frac{\alpha_D}{\alpha_G}$$

NCN nanoHUB.org

FETToy

The theory outlined here has been implemented in FETToy, a simulation tool that you can run on nanoHUB.org. You can also download the program to see how the theory is implemented.



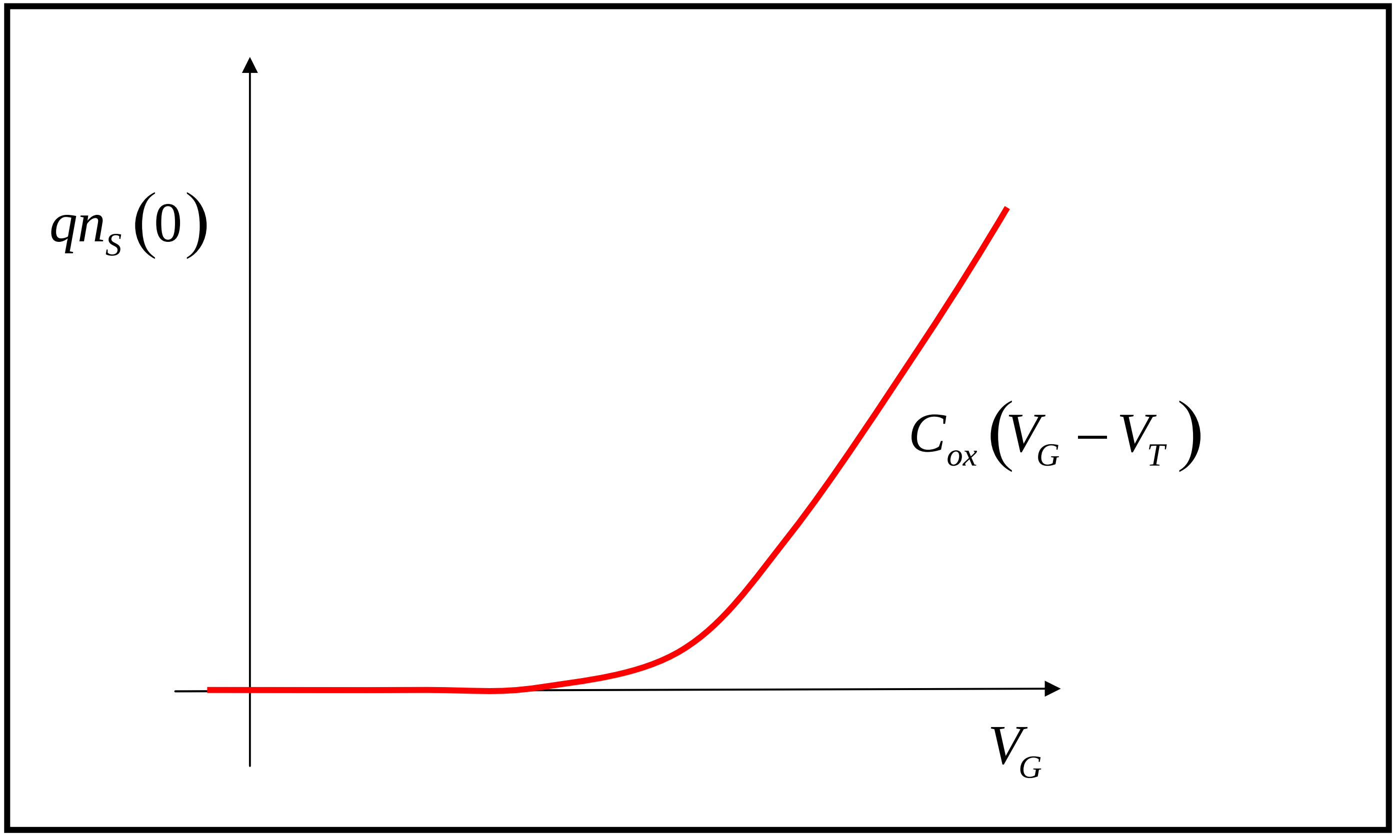
NCN nanoHUB.org

outline

1) Review
2) Effective masses
3) Multiple subbands
4) 2D electrostatics
5) **Discussion**
6) Summary

an aside on MOS electrostatics

in part 1), we assumed ideal, 1D, MOS electrostatics:



an aside on MOS electrostatics (ii)

capacitor model:

$$\psi_S = V_G\left(\frac{C_{GB}}{C_\Sigma}\right) + V_D\left(\frac{C_{DB}}{C_\Sigma}\right) + V_S\left(\frac{C_{SB}}{C_\Sigma}\right) - \frac{q\left[n_S(\psi_S) - n_{S0}\right]W\mathcal{L}}{C_\Sigma}$$

1D MOS electrostatics:

$$\psi_S = V_G - \frac{q\left[n_S(\psi_S) - n_{S0}\right]}{C_{ox}} \qquad (1)$$

also, for high drain bias:

$$n_S(0) = \frac{N_{2D}}{2}\mathcal{F}_0\left(\eta_{F1}\right) \qquad (2)$$

Eqns. (1) and (2) can be solved for $\psi_S(V_G)$ and for $n_S(V_G)$.

(assuming strong carrier degeneracy)

an aside on MOS electrostatics (iii)

results:

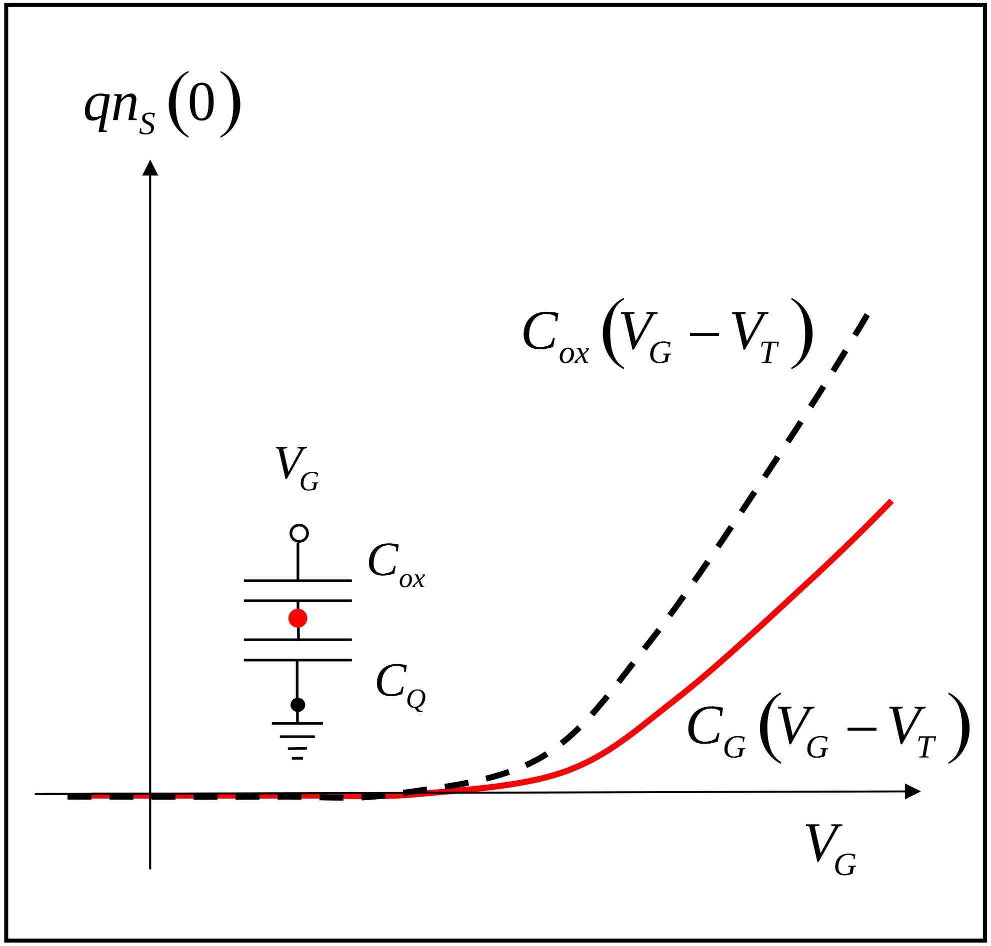
$$\psi_S = \psi_{S0} + \frac{V_G}{1 + C_Q/C_{ox}}$$

$$C_Q = \frac{q^2 m_D^*}{2\pi\hbar^2}$$

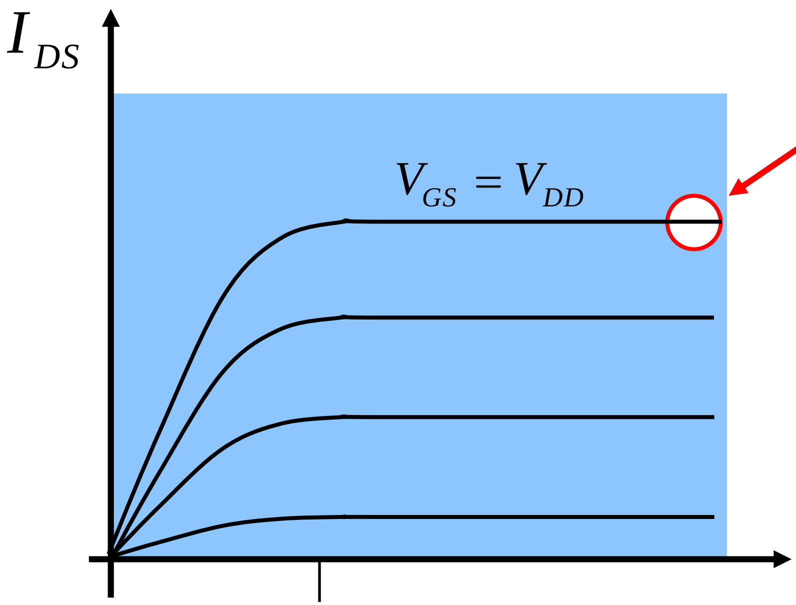
"quantum capacitance"

$$n_S = \frac{C_{ox} C_Q}{C_{ox} + C_Q}\left(V_G - V_T\right)$$

$$\boxed{C_G = \frac{C_{ox} C_Q}{C_{ox} + C_Q}}$$



an aside on MOS electrostatics (iv)



$$I_{ON} = WC_G \upsilon_T \left(V_{GS} - V_T\right)$$

$$C_G = \frac{C_{ox} C_Q}{C_{ox} + C_Q}$$

$$C_Q = \frac{q^2 m_D^*}{2\pi h^2}$$

more generally:

$$C_Q \rightarrow C_S \equiv \frac{\varepsilon_{ox}}{t_{inv}}$$

an aside on MOS electrostatics (v)

$$qn_S(0) = C_G\left(V_G - V_T\right) = \frac{C_{ox}C_Q}{C_{ox}+C_Q}\left(V_G - V_T\right) \qquad C_Q = \frac{q^2 m^*}{2\pi\hbar^2}$$

$$I_D = WQ_I(0)\langle\upsilon(0)\rangle \qquad\qquad \upsilon_T \equiv \sqrt{\frac{2k_BT}{\pi m^*}}\,\frac{\mathcal{F}_{1/2}\left(\eta_{F1}\right)}{\mathcal{F}_0\left(\eta_{F1}\right)}$$

Small effective mass means high injection velocity. :-)

Small effective mass means low inversion layer density. :-(

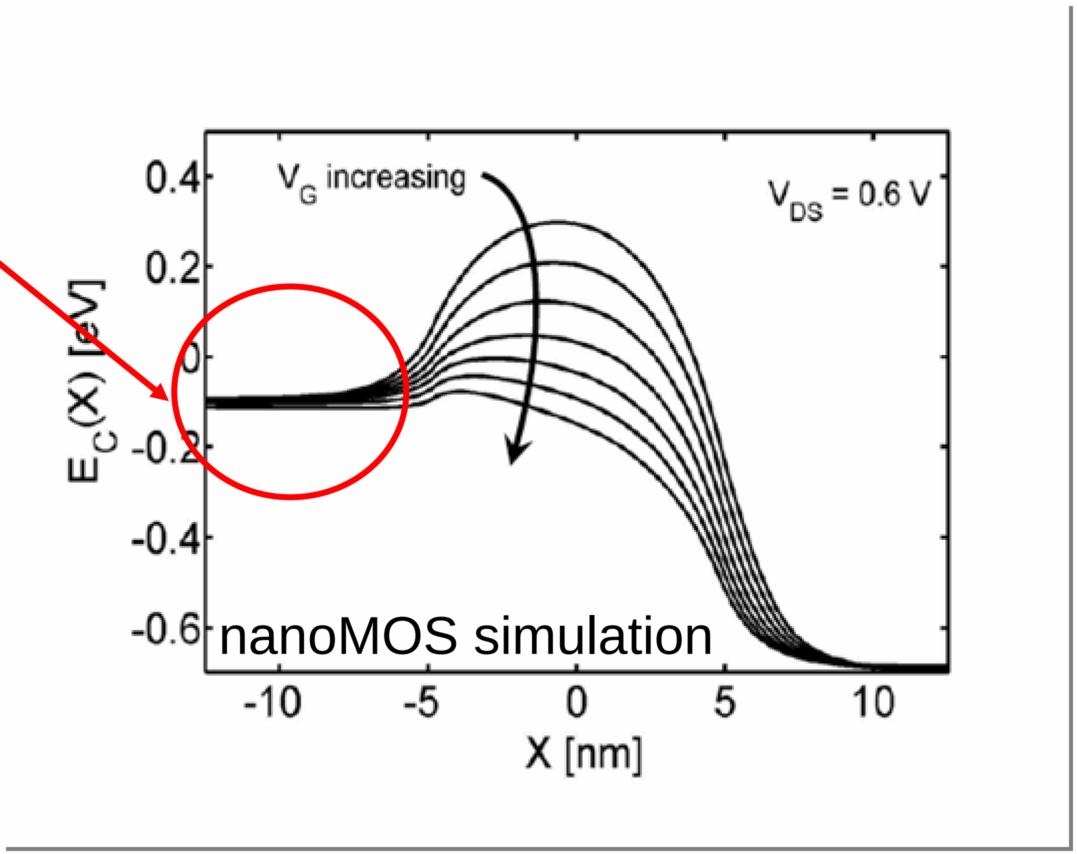
“density-of-states bottleneck”

M. V. Fischetti and S. E. Laux, “Monte Carlo simulation of transport in technologically significant semiconductors of the diamond and zinc-blende structures-Part II: Submicrometer MOSFET’s,” IEEE Trans. Elect. Dev., **38**, 1991.

P. M. Solomon, and S.E. Laux, “The ballistic FET: Design, capacitance and speed limit,” in *IEDM Tech. Dig*., Dec. 2001.

source electrostatics

Note that the conduction band in the source drops slightly as V_{GS} increases. Why?



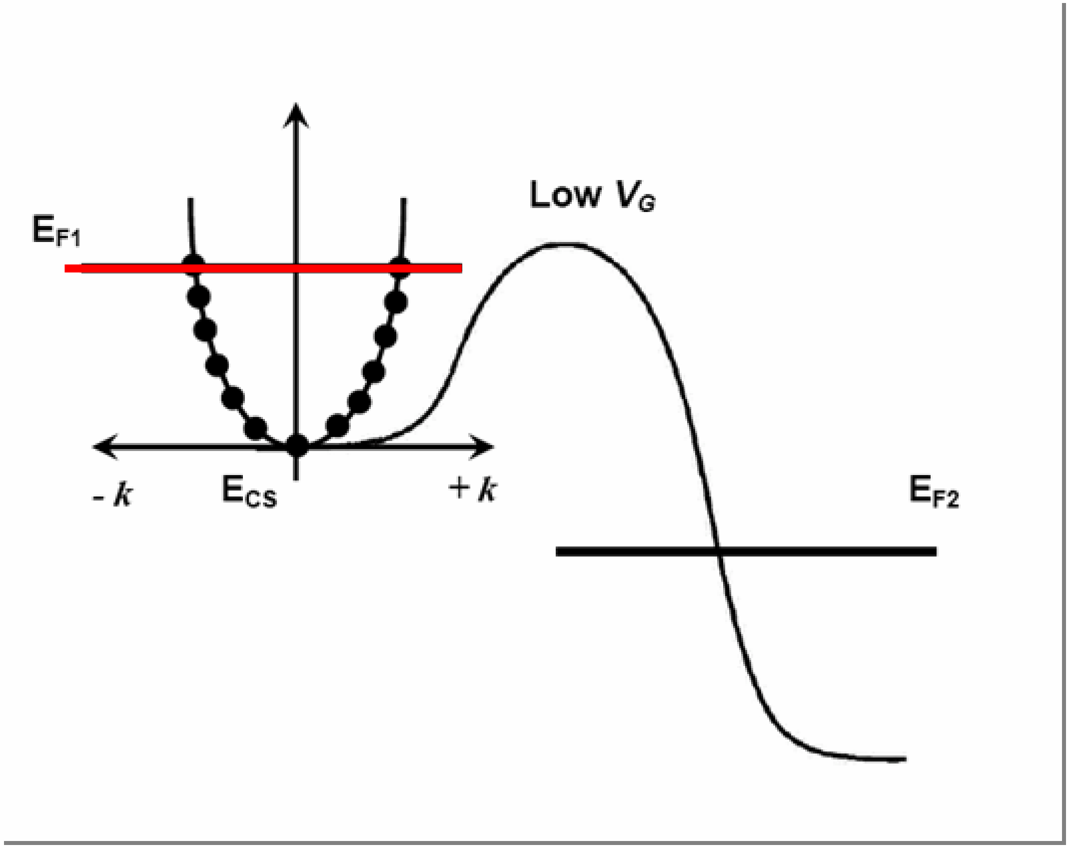
The reason has to do with the electrostatics of the source under ballistic conditions.

A. Rahman, et al., *IEEE TED,* **50**, 1853, 2003.

source electrostatics (ii)

Electrons are injected from the contact into the ballistic source. Most reflect off of the barrier, so +k and -k states are occupied.

The total electron density in the source is equal to the total dopant density.

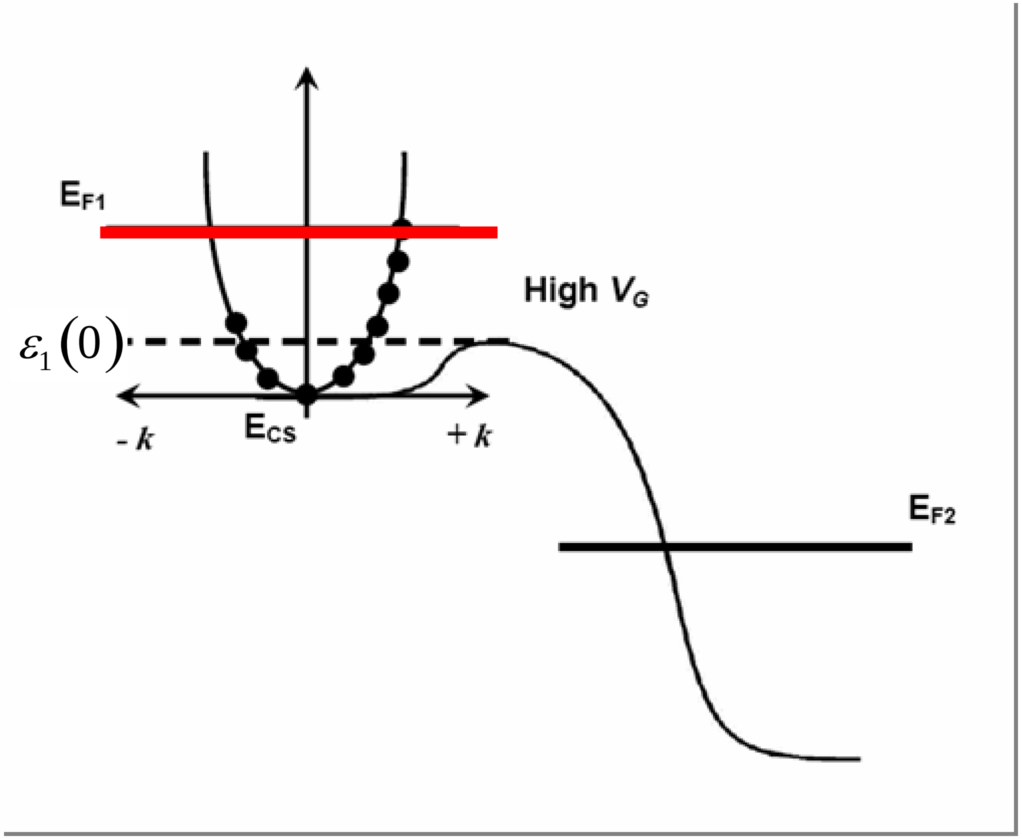


A. Rahman, et al., *IEEE TED,* **50**, 1853, 2003.

source electrostatics (iii)

Under high gate bias, less reflection occurs and fewer -k states are occupied.

This cannot be the correct band diagram, because the total electron charge is now less than the total dopant charge.

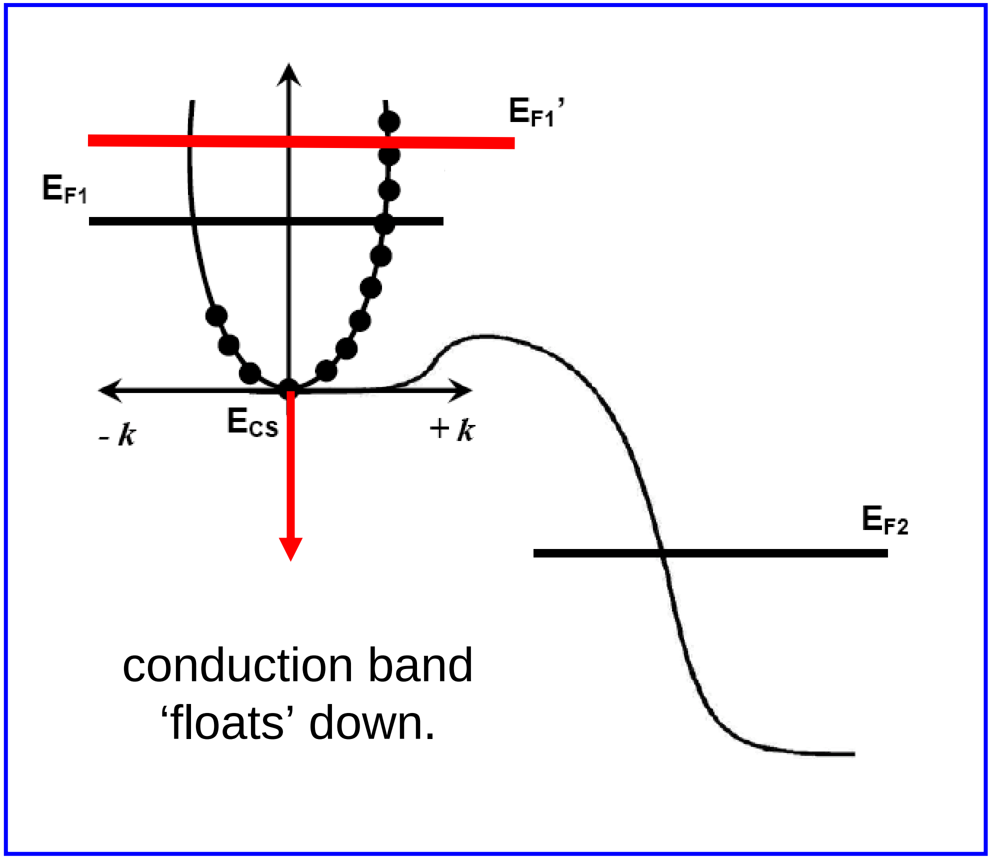


A. Rahman, et al., *IEEE TED*, **50**, 1853, 2003.

source electrostatics (iv)

To satisfy the Poisson equation, the conduction band in the source moves down to let more +v electrons in.

Charge balance in the source has been restored, but the top of the barrier is now lower, so ***I_D* is higher.**



A. Rahman, et al., *IEEE TED*, **50**, 1853, 2003.

source electrostatics (v)

For a discussion of how to treat these “floating source” effects in our “top of the barrier model,” see:

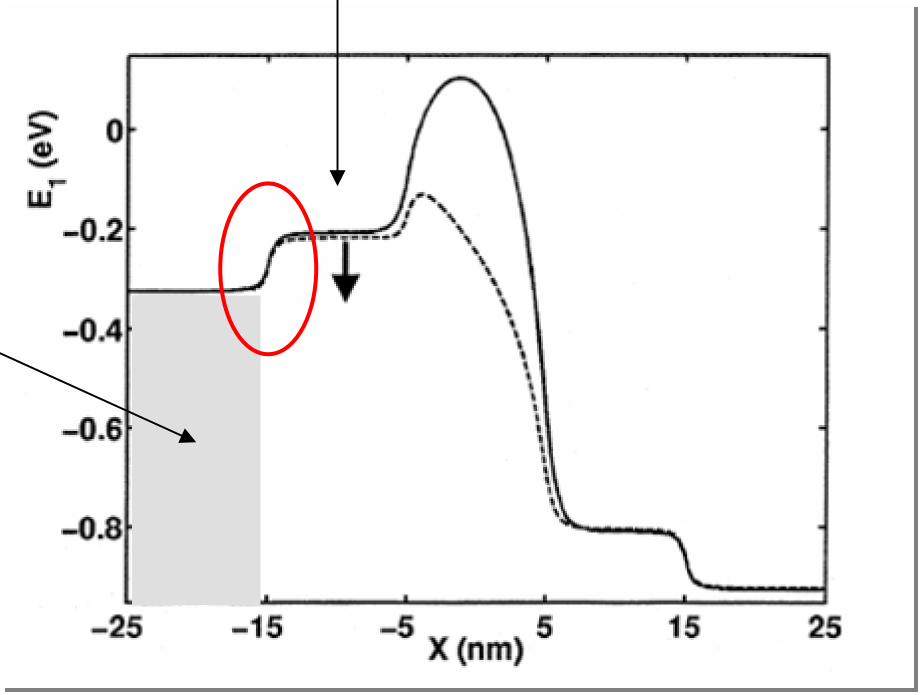
Anisur Rahman, Jing Guo, Supriyo Datta, and Mark Lundstrom, “Theory of Ballistic Nanotransistors,” *IEEE Trans. Electron. Dev*., Nanoelectronics, **50**, pp. 1853-1864, 2003.

NCN nanoHUB.org

numerical simulation

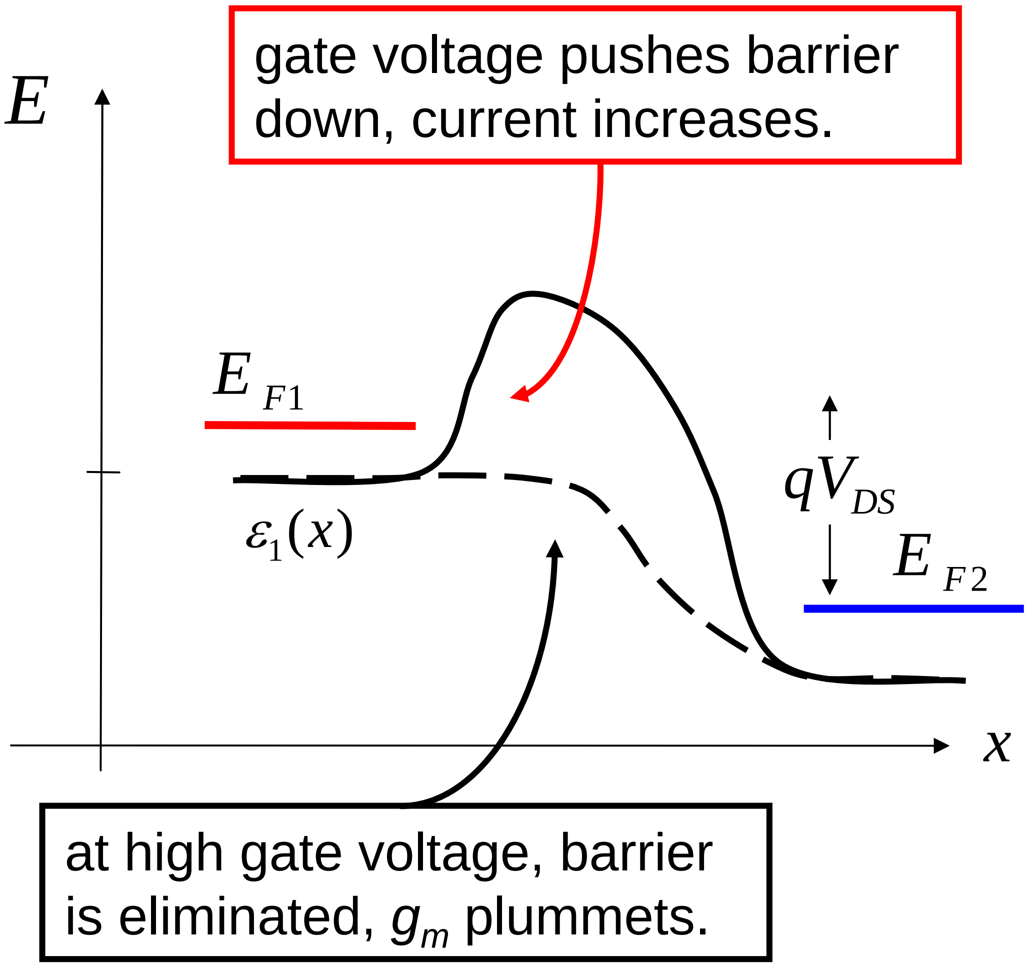
source

n^{++}source
contact



Zhibin Ren, Ramesh Venugopal, Sebastien Goasguen, Supriyo Datta, and Mark S. Lundstrom, "nanoMOS 2.5: A Two-Dimensional Simulator for Quantum Transport in Double-Gate MOSFETs," *IEEE Trans. Elec. Dev.*, **50**, pp. 1914-1925, 2003.

“source exhaustion”



Another view: The charge in the channel cannot be greater than the charge in the source.

$$Q_I(\max) = qN_{DS}x_j$$

This effect can be important when the source is not heavily doped, for example, in a III-V FET.

source exhaustion comments

Source exhaustion, as I have described is it is purely electrostatic effect that is present in a ballistic or drift-diffusion model.

Another effect, “source starvation,” which has to do with the injection of carriers from the 3D contact to the 2D channel, may also be important.

See:

M. Fischetti, T. O’ Reagan, S. Narayanan, C. Sachs, S. Jin, J. Kim, and Y. Zhang,“Theoretical study of some physical aspects of electronic transport in nMOSFETs at the 10-nm gate-length,” *IEEE Trans. Elect. Dev*., **54**, 2007.

NCN nanoHUB.org

outline

1) Review
2) Effective masses
3) Multiple subbands
4) 2D electrostatics
5) Discussion
6) **Summary**

summary

We have generalized

$$I_D = WC_{ox}\left(V_{GS} - V_T\right)\upsilon_T \left[\frac{1 - \mathcal{F}_{1/2}\left(\eta_{F2}\right)/\mathcal{F}_{1/2}\left(\eta_{F1}\right)}{1 + \mathcal{F}_0\left(\eta_{F2}\right)/\mathcal{F}_{10}\left(\eta_{F1}\right)}\right]$$

to include 2D and subthreshold electrostatics as well as the effect of the quantum capacitance above threshold:

$$\psi_S = V_G\left(\frac{C_G}{C_\Sigma}\right) + V_D\left(\frac{C_D}{C_\Sigma}\right) + V_S\left(\frac{C_S}{C_\Sigma}\right) - \frac{qn_S\left(\psi_S\right)}{C_\Sigma}$$

$$I_D = Wq\left(\frac{N_{2D}}{2}\upsilon_T\right)\left[\mathcal{F}_{1/2}\left(\eta_{F1}\right) - \mathcal{F}_{1/2}\left(\eta_{F2}\right)\right]$$

and we discussed some of the implications.

suggested exercise: subthreshold conduction

$$\psi_S = V_G\left(\frac{C_{GB}}{C_\Sigma}\right) + V_D\left(\frac{C_{DB}}{C_\Sigma}\right) + V_S\left(\frac{C_{SB}}{C_\Sigma}\right) - \frac{q\left[n_S(\psi_S) - n_{S0}\right]W\mathcal{L}}{C_\Sigma}$$

$$\eta_{F1} \equiv \left[E_{F1} - \varepsilon_{10} + q\psi_S\right]/k_B T \qquad \eta_{F2} = \eta_{F1} - qV_{DS}/k_B T$$

$$I_D = Wq\left(\frac{N_{2D}}{2}\upsilon_T\right)\left[\mathcal{F}_{1/2}(\eta_{F1}) - \mathcal{F}_{1/2}(\eta_{F2})\right]$$

Exercise: Simplify for 1D electrostatics and subthreshold conduction and derive the subthreshold I-V characteristics of a ballistic MOSFET.

suggested exercise: bulk MOSFET

$$\psi_S = V_G\left(\frac{C_{GB}}{C_\Sigma}\right) + V_D\left(\frac{C_{DB}}{C_\Sigma}\right) + V_S\left(\frac{C_{SB}}{C_\Sigma}\right) - \frac{q\left[n_S(\psi_S) - n_{S0}\right]W\mathcal{L}}{C_\Sigma}$$

$$\eta_{F1} \equiv \left[E_{F1} - \varepsilon_{10} + q\psi_S\right]/k_B T \qquad \eta_{F2} = \eta_{F1} - qV_{DS}/k_B T$$

$$I_D = Wq\left(\frac{N_{2D}}{2}\upsilon_T\right)\left[\mathcal{F}_{1/2}(\eta_{F1}) - \mathcal{F}_{1/2}(\eta_{F2})\right]$$

Exercise: Repeat the derivation and develop a top of the barrier model for a bulk MOSFET.