Nanoscale Transistors: Device Physics, Modeling, and Simulation

Mark Lundstrom
Electrical and Computer Engineering
Purdue University
West Lafayette, IN 47907

Chapter 3: The Ballistic MOSFET

3.1 Introduction

Silicon MOSFET channel lengths are rapidly approaching 10nm regime where our traditional understanding of transistors must be questioned. In this chapter, we describe an approach to MOSFETs based on concepts that are widely-used in mesoscopic physics [1]. Figure 3.1 shows the lowest conduction subband energy vs. position along the channel of a 10nm MOSFET under a variety of gate and drain biases [2]. The results show that there is an energy barrier between the source and channel and that the height of the barrier is modulated by the gate voltage. The drain current increases as the barrier height in reduced by the increasing gate voltage (Fig. 3.1a for low drain bias and Fig. 3.1b for high drain bias). For an electrostatically well designed MOSFET, the drain bias has a small effect on the height of the source to channel barrier (Figs. 3.1c and 3.1d). Because current flow is controlled by injection across a barrier, a MOSFET is similar to a bipolar transistor, except that in a bipolar transistor the barrier height is directly modulated by the emitter-base voltage while in the MOSFET it is modulated indirectly by the gate voltage [3]. The theory of the nano-scale MOSFETs is based on this simple, physical picture.

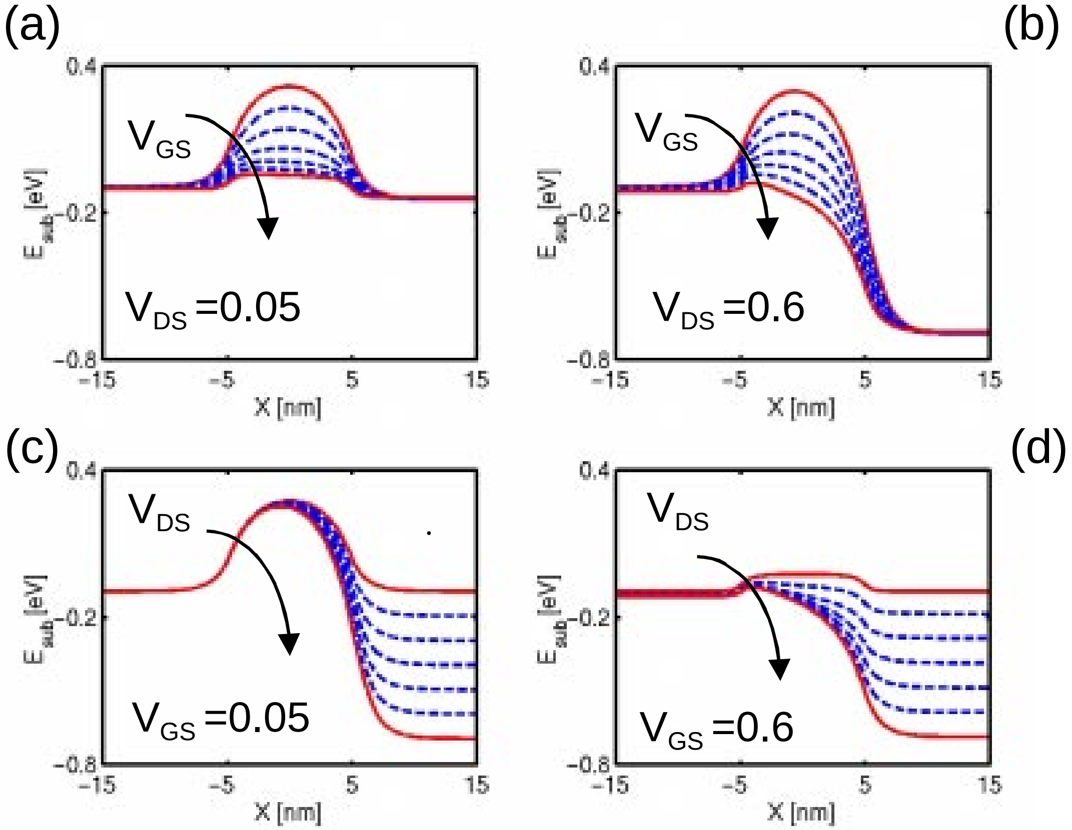

Fig. 1 Energy band diagrams (conduction subband vs. position) for a 10nm channel length, double gate MOSFET under a variety of bias conditions. (From [2])

Before proceeding, we should clarify what is being plotted in Fig. 3.1. Electrical engineers plot "energy band diagrams," the conduction band minimum (or valence band maximum) vs. position. In nanoscale devices, however, quantum confinement effects are strong, so it is more appropriate to plot the electron eigenstates. Figure 3.2 shows the difference for a double gate MOSFET of the type discussed in Chapter 2. Because quantum confinement in the z-direction is strong, the subbands are widely separated and do not couple. We can, therefore, solve a one-dimensional Schrödinger equation in the z-direction at various positions along the channel in the x-direction. Figure 3.2 shows the result. Note that the conduction band minima, $E_C(x, y)$, varies in two dimensions, but the subband profiles vary only in the x-direction. Unless otherwise noted, all energy band diagrams will refer to plots of the subband energy vs. position.

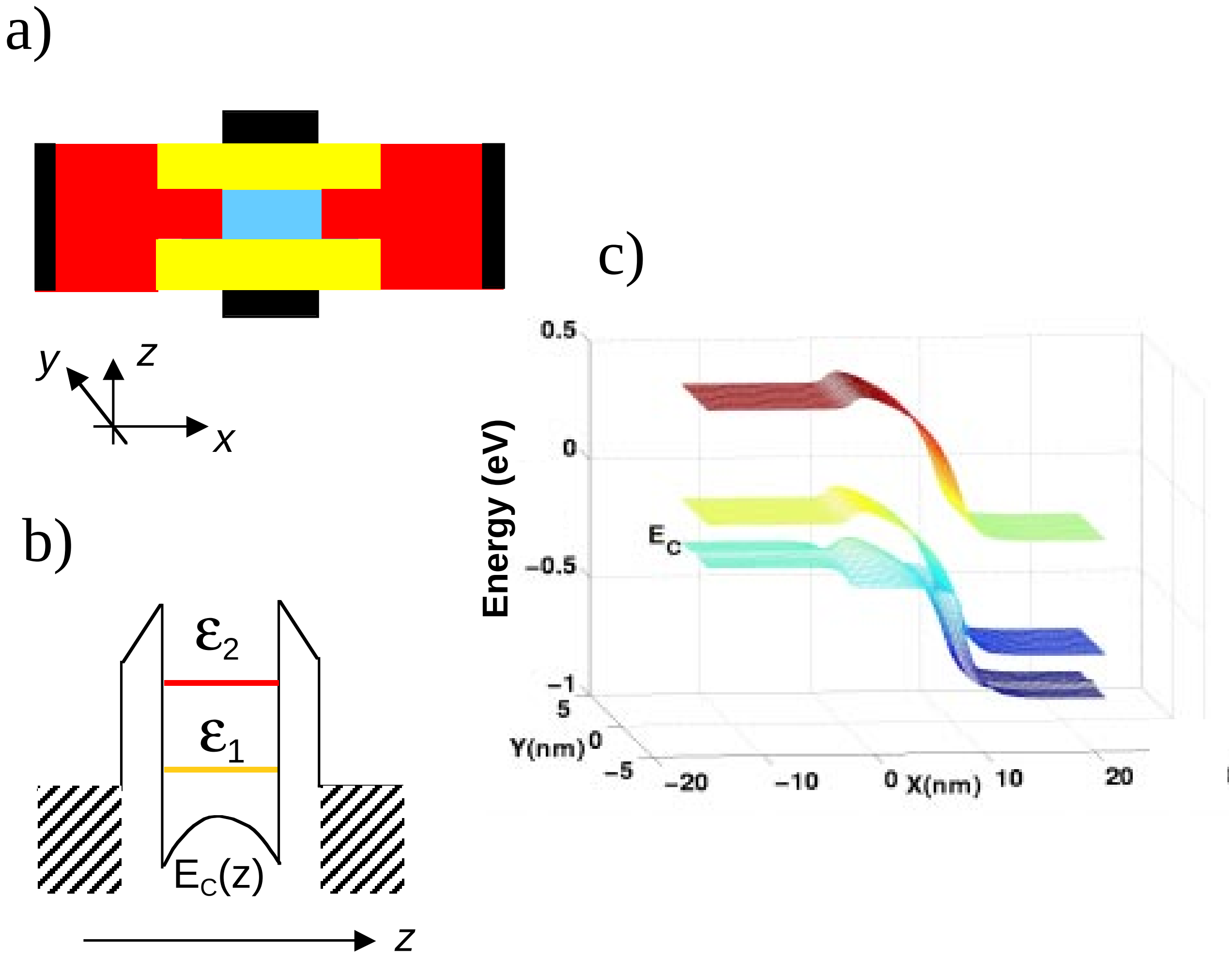

Fig. 3.2 Comparison of the conduction band profile and the subband profile. (a) the double gate device structure. (b) illustration of the quantum confinement in the z-direction, (c) the conduction band and subband profiles in the x-y plane. (Courtesy of Ramesh Venugopal.)

3.2 Physical View of nano-MOSFETs

MOSFETs are complicated by 2D electrostatics and strong, off-equilibrium transport in the presence of rapidly varying electric fields. These effects have been recently examined by 2D, numerical simulations using the non-equilibrium Green's function approach [4, 5] (see also [6] for an earlier quantum scale study of 10 nm MOSFETs and [7] for a recent review). For extremely short channel lengths, quantum mechanical tunneling from source to drain (through the barrier) degrades device performance, but for channel lengths above 10 nm, MOSFETs behave classically [3]. We will, therefore, adopt a classical model. Scattering, which complicates the analysis, is the subject of Chapter 4, but the essential physics of nanoscale MOSFETs can be established by numerical simulations of ballistic MOSFETs.

Figure 3 shows the numerically computed carrier distribution function vs. position for a 10nm channel length ballistic MOSFET operating under high-bias [8]. The carrier distribution function is seen to be strongly distorted from its equilibrium value. Note, for example, the ballistic peak that develops within the channel.

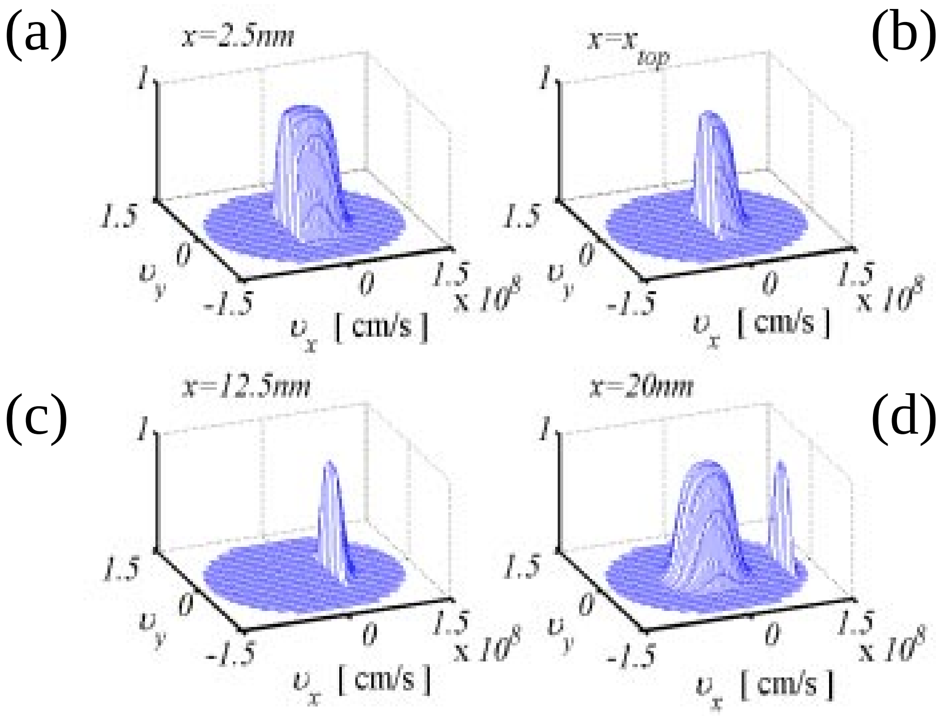

Fig. 3.3. Electron distribution function vs. position for a ballistic n-MOSFET as computed under high gate and drain bias [8]. (a) in the n^+ source, (b) at the top of the barrier, (c) deep within the channel, and (d) within the n^+ drain.

The complex shape of the carrier distribution function would seem to preclude simple, analytical modeling, but Fig. 3.4 shows that things are much simpler at the top of the source-channel barrier. This figure shows the carrier distribution function under high gate bias at four different drain voltages. For $V_{DS} = 0$, the distribution has an equilibrium shape; the positive half of the thermal distribution was injected from the thermal reservoir at the source and the negative half from the drain. As V_{DS} increases, the negative portion of the distribution diminishes. Note, however, that the positive half grows. This occurs because of the MOSFET's electrostatics. In an electrostatically well-designed MOSFET, the gate holds the charge at the top of the barrier approximately constant with drain bias. Above threshold, the value is $\approx C_G(V_{GS} - V_T)$, except for a shift of the threshold voltage, V_T, depending on the magnitude of the 2D short channel effects.

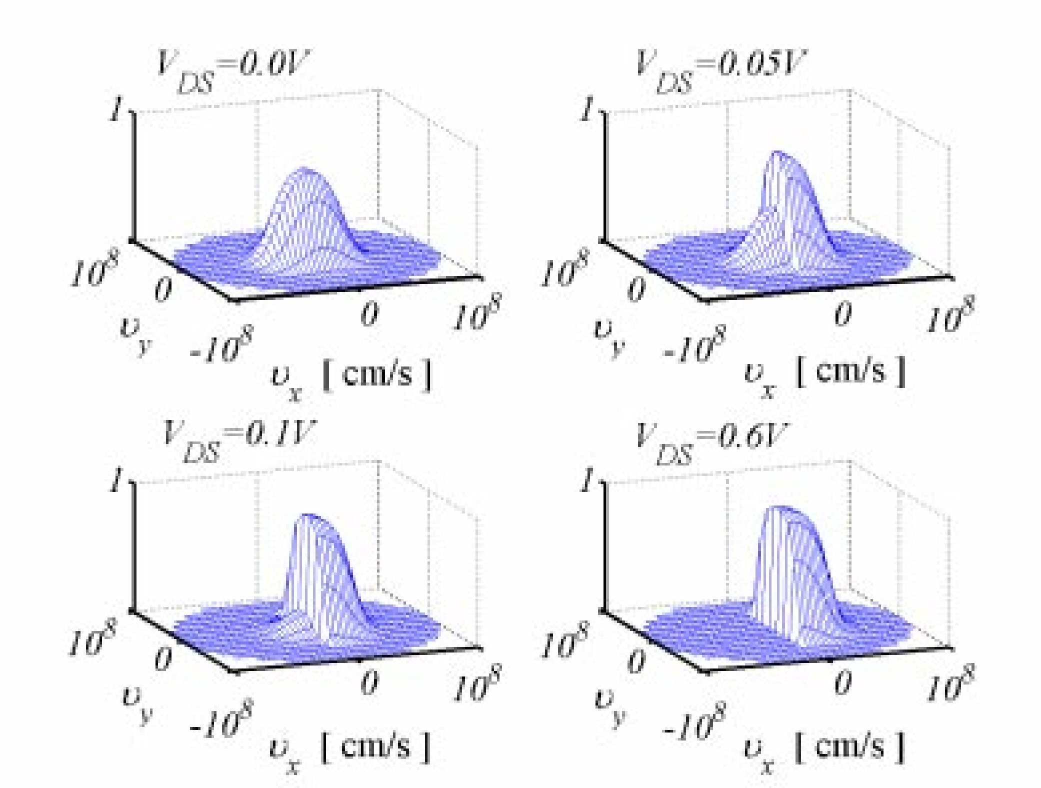

Fig. 3.4. Electron distribution function at the top of the source-channel barrier for a ballistic n-MOSFET as computed under high gate and four different drain biases [8].

For a ballistic MOSFET, the average electron velocity within the channel varies rapidly in space and reacheS very high values. The typical velocity vs. position profiles are displayed in Fig. 3.5 would seem to preclude simple analysis.

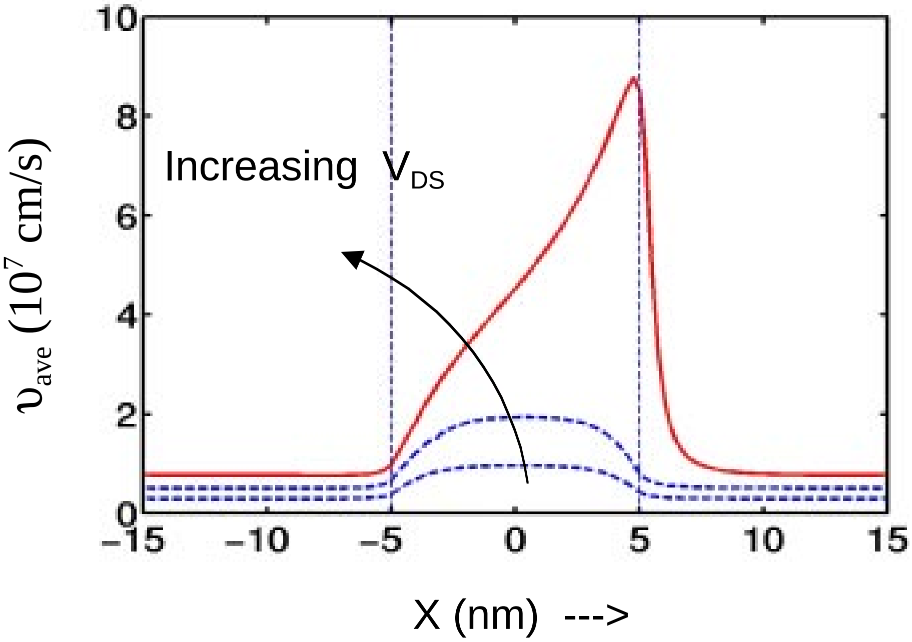

Fig. 3.5. Average electron velocity vs. positon for the ballistic, double gate n-MOSFET under high gate bias and at four different drain biases [8].

Since we have already seen in Fig. 3.4 that the electron density at the top of the barrier is approximately independent of drain bias, it is natural to inquire how the average velocity at the top of the barrier varies with drain bias. As shown in Fig. 3.6, the average velocity rapidly increases with V_{DS}, then saturates. The magnitude of the saturated velocity increases with gate voltage. Velocity saturation is a familiar concept is short channel MOSFET where it occurs in the high field region where carriers are excited to high kinetic energies and scatter frequently [9]. Its occurrence in a ballistic MOSFET where there is no scattering, and the location of the position at the velocity saturates (the top of the barrier where the electic field is zero) is, at first, a surprise.

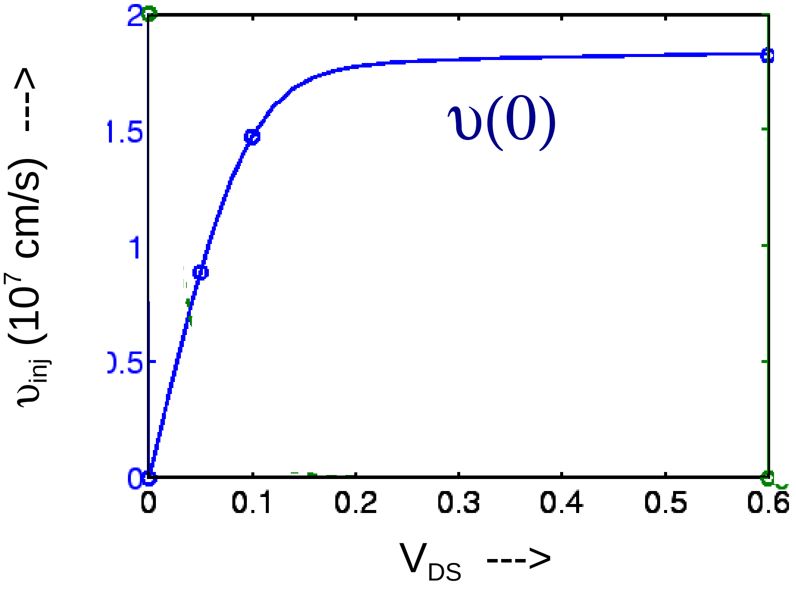

Fig. 3.6. Average electron velocity at the top of the source-channel barrier as a function of drain bias with a high applied gate bias [8].

The key concepts illustrates by the numerical simulation are that:

1) the carrier distribution function at the top of the source-channel barrier consists of two thermal equilibrium halves, one injected from the source and the other from the drain,

2) as the drain bias increases, the negative portion of the distribution diminishes in size, but for an electrostatically well-designed MOSFET, the total carrier density at the top of the barrier is maintained at an approximately constant value, and

3) since the high V_{DS} distribution approaches a hemi-Fermi-Dirac distribution, the average velocity at the top of the barrier saturates at a limiting value which is the average velocity of a thermal equilibrium Fermi-Dirac distribution.

As discussed next, these concepts can be used to develop an analytical theory of the ballistic MOSFET.

3.3 Theory of the Ballistic MOSFET

In this section, we present an overview of the theory of the ballistic MOSFET, without the mathematical details that will be discussed in the following three sections. Figure 3.7 sketches the E-k relation at the top of the source-channel barrier. The positive k-states are populated by injection from the source according to the source Fermi level, E_F, and the negative k-states are populated from the drain according to the drain Fermi level, $E_F - qV_{DS}$. For $V_{DS} = 0$, the positive and negative states are equally occupied, the average velocity is zero, and $I_D = 0$. For small V_{DS}, the drain Fermi level is lowered, fewer negative k-states are occupied, the net velocity is positive, and $I_{DS} \sim V_{DS}$. For large V_{DS}, the negative k-states are empty, the average velocity saturates at a value determined by the gate voltage, which determines the location of the Fermi lwvel within the conduction band. In this case the drain current saturates at the so-called on-current.

By assuming 2D statistics for the quantum confined carrier in the channel, we can relate the carrier populations in the positive and negative halves to their respective Fermi levels. Above threshold, MOS electrostatics then demands that the total carrier density is approximately independent of drain voltage, so we find

$$C'_{ox}\left(V_{GS} - V_T\right) \approx n_S^+\left(E_F\right) + n_S^-(E_F - qV_{DS}). \tag{3.1}$$

Equation (3.1) is an equation for the location of the Fermi level. For a given device design, the gate capacitance, C'_{ox} and threshold voltage, V_T, are determined. Equation (3.1) then determines the location of the Fermi level as a function of gate and drain bias. As the drain bias increases, n_S^- decreases, so E_F increases to maintain charge balance. This occurs physically by the gate electrostatically pushing down the source-channel barrier to let more electrons in from the source. Having determined the Fermi level, the positive and negative fluxes can be evaluated by integrating over the populated states to obtain the drain current as

$$I_D = I^+(E_F) - I^-(E_F - qV_{DS}) \tag{3.2}$$

The approach outlined here is essentially Natori's theory of the ballistic MOSFET [10, 11] which has been further developed recently [12-14].

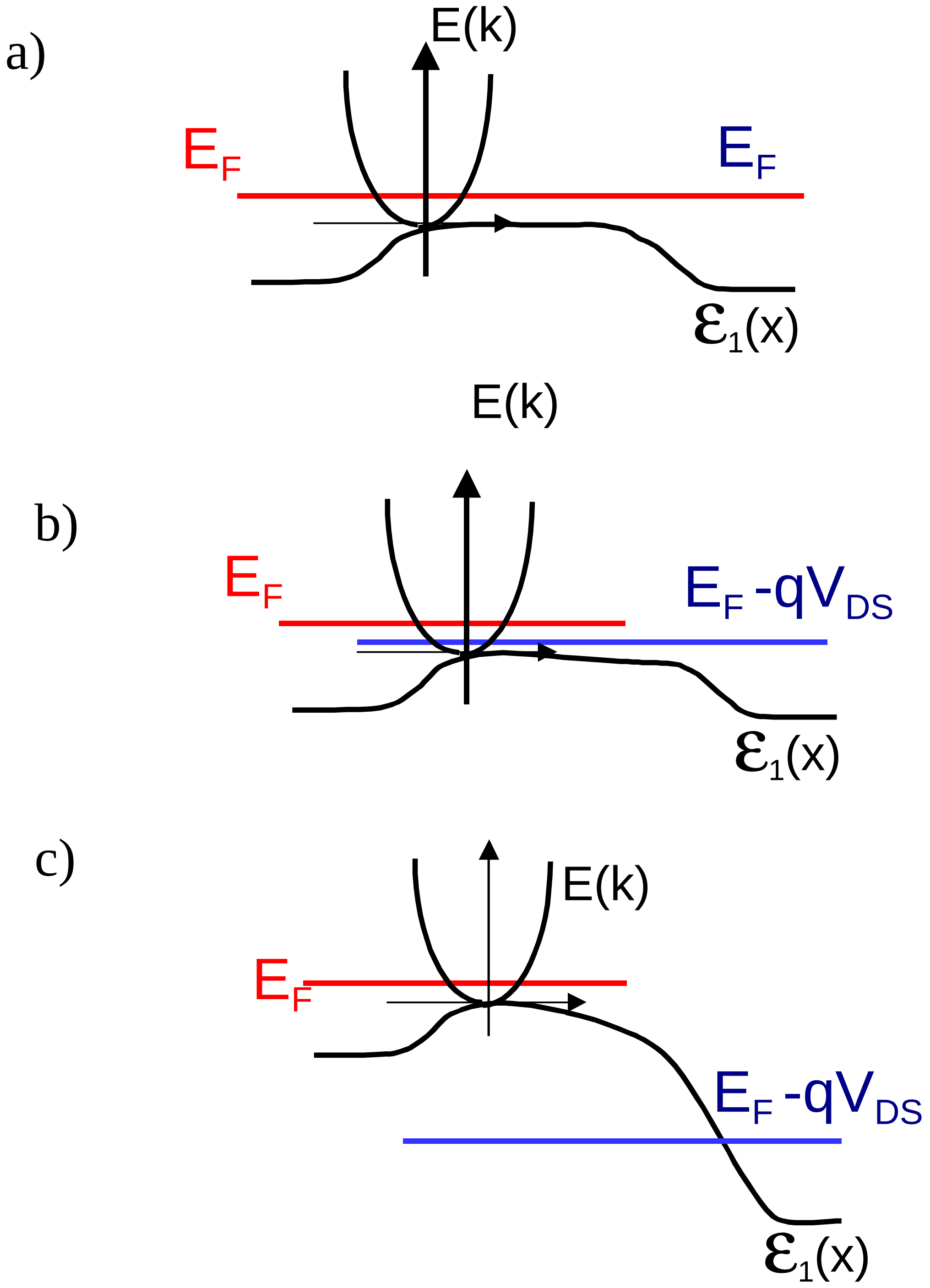

Fig.3.7 E-k relation at the top of the source-channel barrier showing the source and drain Fermi levels, E_F and $E_F - qV_{DS}$. A high gate voltage, which sets the location of the Fermi level in the conduction subband, is assumed. a) $V_{DS} = 0$, b) V_{DS} small, and c) V_{DS} large.

In the following three sections, we will use the approach outlined above, to evaluate the I_{DS}-V_{DS} characteristic of a ballistic MOSFET. The resulting I_D vs. V_{DS} characteristics at room temperature as sketched in Fig. 3.8. Note that the I-V characteristics of a ballistic MOSFET are quite similar to those of a scattering dominated MOSFET.

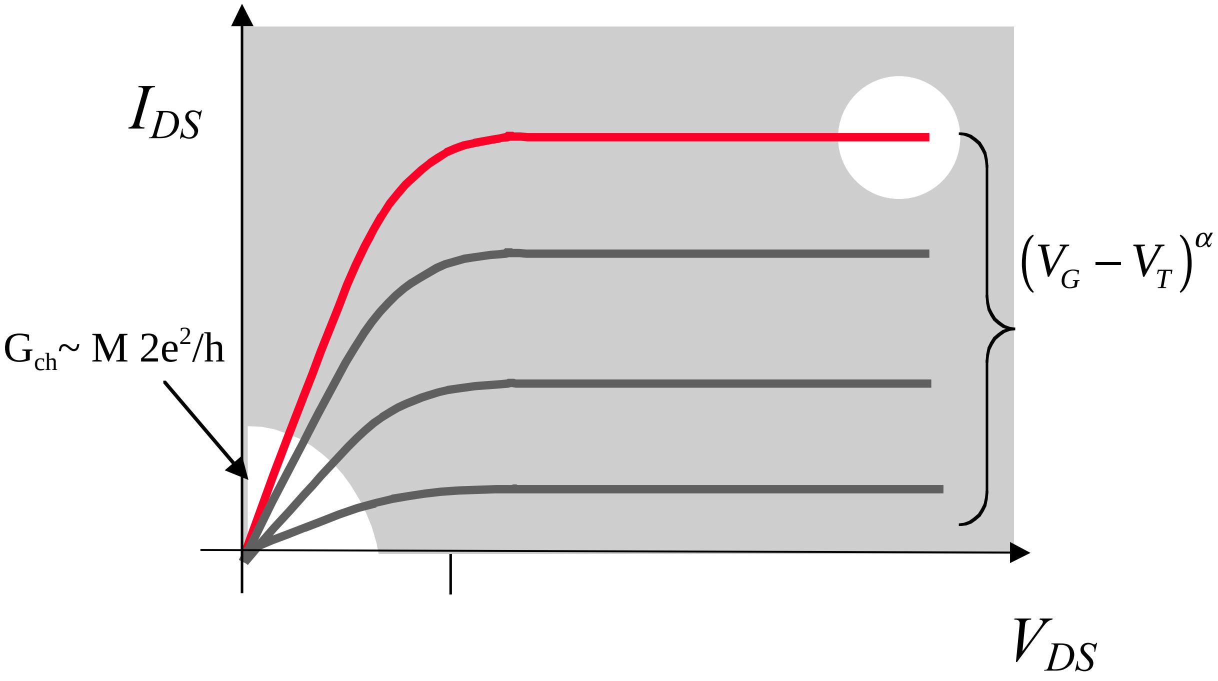

Fig. 3.8 Sketch of the I_{DS} vs. V_{DS} characteristic of a ballistic MOSFET.

The simple expressions for the on-current are illuminating. The on-current is the product of the charge density at the top of the barrier and the average velocity of an equilibrium hemi-Fermi–Dirac distribution. For non-degenerate conditions, we find

$$I_D = W C_G (V_{DD} - V_T)\left[\sqrt{2k_B T / \pi m^*}\right] \tag{3.3a}$$

which shows that the ballistic MOSFET's on-current is proportional to ($V_{GS} - V_T$), just as it is for a velocity saturated MOSFET. As discussed earlier, the channel velocity does saturate in a ballistic MOSFET, but it is the velocity at the top of the barrier that saturates, not the velocity near the drain. In practice, however, nondegenerate statistics do not apply above threshold. In the fully degenerate limit, Eq. (3.3a) becomes

$$I_D = W C_G\left(V_{DD} - V_T\right)\left[\left(8\hbar / 3m^*\right)\sqrt{C_G(V_{GS} - V_T)/q\pi}\right]. \tag{3.3b}$$

MOSFETs typically operate between these two limits, so $I_D \propto \left(V_{GS} - V_T\right)^{\alpha}$ where $1 < \alpha < 1.5$, which is very similar to the result obtained by traditional MOSFET models.

The linear region of the ballistic MOSFET is also interesting. For small V_{DS}, the drain current is

$$I_D = G_{CH} V_{DS}, \tag{3.4}$$

that is, a ballistic MOSFET shows a finite channel conductance. The general expression for G_{CH} involves Fermi-Dirac integrals, but at $T = 0$, it reduces to

$$G_{CH} = M\left(2q^2 / h\right), \tag{3.5}$$

the expected conductance of a ballistic conductor [1]. In Eq. (3.5), M is the number of occupied modes, which is proportional to the width, W, of the MOSFET.

3.4 The Ballistic MOSFET-I (T > 0, nondegenerate)

We first derive the I-V characteristic assuming that the electron gas is nondegenerate, because the result is simple and illuminating. We begin with eq. (3.1), which we write as

$$n_S(0) = n_S^+(0) + n_S^-(0), \tag{3.6}$$

where the location, $x = 0$, refers to the charge control point at the top of the barrier. For a nondegenerate gas, the average velocity of a hemi-Maxwellian distribution is

$$\upsilon_T \equiv \sum_{p_x > 0, p_y} \upsilon_x \, f(E) = \sqrt{\frac{2k_B T}{\pi m^*}}\,, \tag{3.7}$$

which is independent of the location of the Fermi level. For a silicon MOSFET, the appropriate effective mass to use depends on the subband. Assuming that only the lowest subband is occupied, the appropriate effective mass is $m^* = m_t = 0.19\, m_0$, which gives a thermal velocity of $\upsilon_T = 1.2 \times 10^7 \text{ cm/s}$.

Because the average carrier velocities are identical for the two streams, the current contributed by carriers with positive velocity is $I^+/W = q\, n_S^+(0)\upsilon_T$ and for the negative current, $I^-/W = q\, n_S^-(0)\upsilon_T$. The net drain current, Eq. (3.2), becomes

$$I_D/W = q\left(n_S^-(0) - n_S^-(0)\right)\upsilon_T . \tag{3.8}$$

Using eqs. (3.6) and (3.8), we find

$$I_D/W = q\, n_S(0)\upsilon_T \frac{\left(1 - n_S^+(0)/n_S^-(0)\right)}{\left(1 + n_S^+(0)/n_S^-(0)\right)}\,. \tag{3.9}$$

Nondegenerate carrier statistic lead to a simple expression for $n_S^+(0)/n_S^-(0)$. Recall that

$$n_S^+(0) = \left(\frac{N_{2D}}{2}\right) e^{(E_F - \varepsilon)/k_B T} \tag{3.10}$$

because the positive k-states are populated according to the source Fermi level. Similarly,

$$n_S^+(0) = \left(\frac{N_{2D}}{2}\right) e^{(E_F - qV_D - \varepsilon)/k_B T} \tag{3.11}$$

because the negative k-states are populated according to the drain Fermi level, $E_F - qV_D$. In these equations,

$$N_{2D} = \left(\frac{m^*}{\pi \hbar^2}\right) k_B T \tag{3.12}$$

is the two-dimensional effective density-of-states, and we have assumed that a single subband, at energy, ε, is occupied. From Eqs. (3.10) and (3.11), we see that $n_S^+ / n_S^- = e^{-qV_D/k_B T}$, so Eq. (3.9) becomes

$$I_D/W = q\, n_S(0) \upsilon_T \frac{\left(1 - e^{-qV_D/k_B T}\right)}{\left(1 + e^{-qV_D/k_B T}\right)}, \tag{3.13}$$

which, using MOS electrostatics for $q\, n_S = C'_{ox}\left(V_G - V_T\right)$, gives the I-V characteristic of the ballistic MOSFET under nodegenerate conditions as

$$I_D = W C'_{ox}\left(V_G - V_T\right) \upsilon_T \frac{\left(1 - e^{-qV_D/k_B T}\right)}{\left(1 + e^{-qV_D/k_B T}\right)}. \tag{3.14}$$

Although the assumption of nondegenerate carrier statistics is a poor one above threshold, Eq. (3.14) is a simple expression that provides some insight into ballistic nanotransistors. A plot of the common source characteristics from Eq. (3.14) gives characteristics like those sketched inn Fig. 3.8. The drain current of a ballistic MOSFET (under nondegenerate conditions) saturates when $V_D > V_{Dsat} \approx 2\left(k_B T / q\right)$. For large drain voltages, we find the on-current as

$$I_D(on) = W C'_{ox} \upsilon_T \left(V_G - V_T\right), \tag{3.15}$$

which has the form of a traditional, velocity-saturated MOSFET model, except that υ_{sat} is replaced by the thermal injection velocity, υ_T. The characteristic exponent for the gate voltage dependence, is α = 1, just as it is in the traditional velocity saturation model. These similarities to the traditional velocity saturation model should have been expected, because we saw that the average carrier velocity does saturate – at the charge control point located at the top of the barrier.

Figure 3.8 and Eq. (3.14) also show that the ballistic MOSFET has a finite channel conductance. For small drain voltages, the exponentials can be expanded to find

$$I_D = \left[W C'_{ox} \left(V_G - V_T \right) \frac{\upsilon_T}{2\left(k_B T / q \right)} \right] V_D = G_{CH} V_D, \tag{3.16}$$

where G_{CH} is the channel conductance. The corresponding result in the degenerate limit is $G = M\left(2q^2 / h\right)$. We have assumed that only one subband is populated, but the number of available modes in the transverse direction is proportional to the width, W, of the MOSFET. Alternatively, we can express the channel conductance as

$$G_{CH} = I_D(\text{on})/(2k_B T / q). \tag{3.17}$$

It is also instructive to compare the ballistic channel conductance to the traditional, drift-diffusion result,

$$I_D = \left[W C'_{ox} \left(V_G - V_T \right) \frac{\mu_{eff}}{L_{eff}} \right] V_D. \tag{3.18}$$

The conductance of a MOSFET cannot be greater than the ballistic limit, so we conclude that the traditional model is valid along as

$$\mu_{eff} \frac{(2kT/q)}{L_{eff}} < \upsilon_T . \tag{3.19}$$

Finally, we note that the origin of the ballistic conductance is easy to appreciate in this nondegenerate picture. Consider Fig. 3.9, which sketches the energy band diagram under high gate bias and low drain bias. At the top of the barrier, we have a positive current, I^+, injected from the source. Because those injected from the drain experience a larger barrier, I^- is smaller by a factor of e^{-qV_D/k_BT}, so the net current is

$$I_D = I^+\left(1 - e^{-qV_D/k_BT}\right) \approx I^+/\left(2k_BT/q\right). \tag{3.20}$$

Under low drain bias, $n_S^+ \approx n_S^- \approx n_S/2$, so

$$I^+ = qW\left(n_S(0)/2\right)\upsilon_T , \tag{3.21}$$

which, when inserted in Eq. (3.20) gives (3.17). So the ballistic conductance is a direct consequence of our thermionic emission model as evaluated under a low drain voltage assumption.

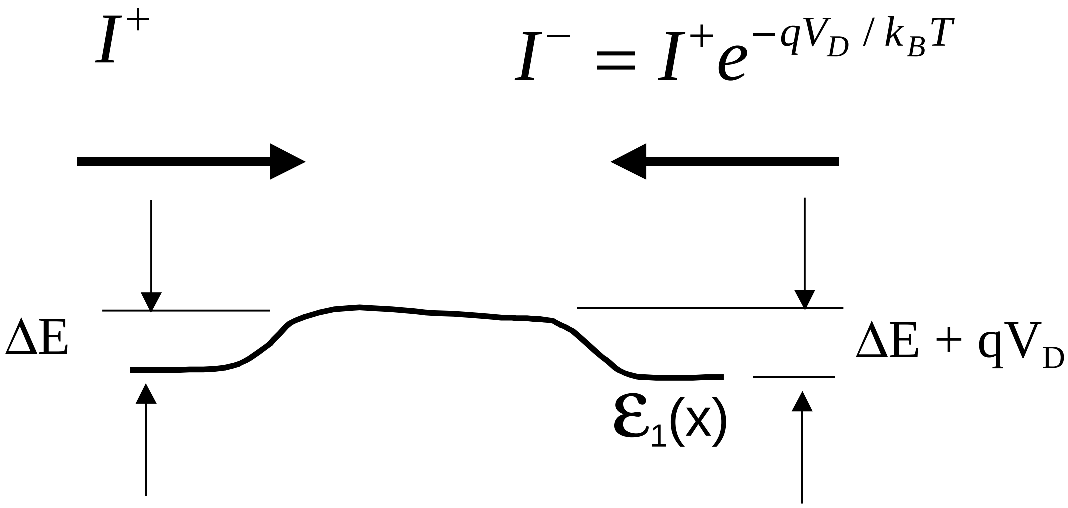

Fig. 3.9 The energy band diagram of a MOSFET under high gate voltage and low drain bias showing the ballistic currents injected from the source, I^+, and from the drain, I^-.

3.5 The Ballistic MOSFET II (T = 0, degenerate)

Figure 3.10 sketches the I_D vs. V_D characteristic of an ideal ballistic MOSFET at T = 0. (Not shown in this figure is a sharp transition vs. gate voltage as the device switches from off to on.) Because the transition from the linear to saturated region is difficult to treat simply, we will evaluate only three quantities: 1) the channel conductance, G_{CH}, 2) the on-current, I_D (on), and 3) the drain saturation voltage, V_{Dsat}. (This section will conclude by estimating the width of the liner to saturation region transition.)

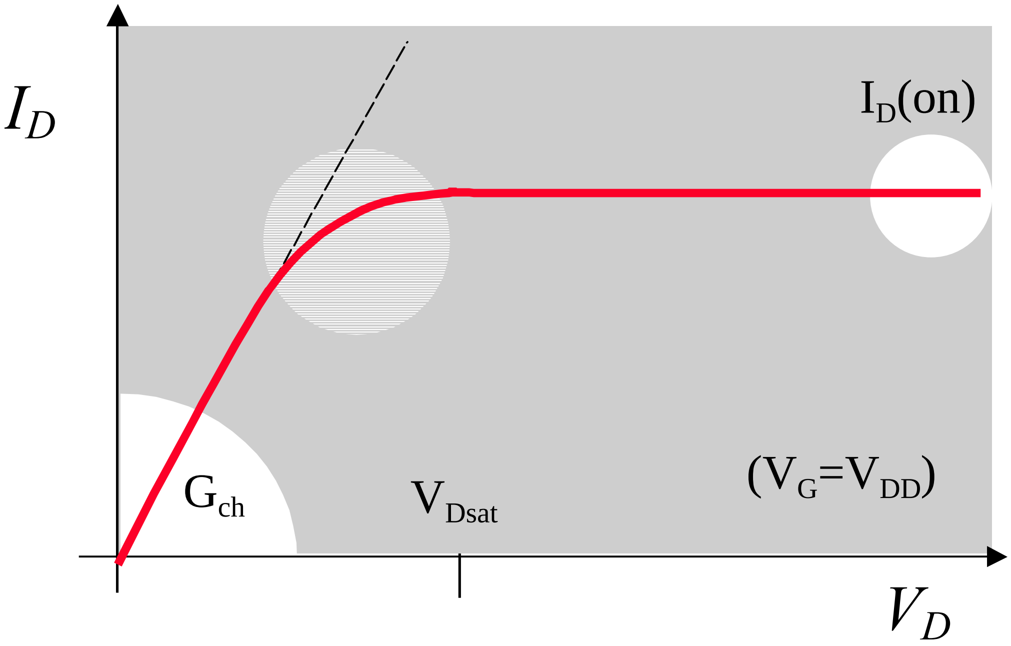

Fig. 3.10 Sketch of the I_D vs. V_D characteristic at $T = 0$.

3.5.1 The T = 0K channel conductance

To evaluate the channel conductance, we begin with Eq. (3.2) and assume that $V_D << V_{Dsat}$. In this region, I^+ and I^- are nearly equal so

$$I^- = I^+ - \left(\frac{\partial I^+}{\partial E_F}\right) qV_D. \tag{3.22}$$

Using Eq. (3.22) in (3.2) produces

$$I_D = qW\left(\frac{\partial I^+}{\partial E_F}\right)V_D. \tag{3.23}$$

For a degenerate carrier gas,

$$I^{+} = qW\frac{\hbar k_F^3}{3\pi^2 m} = qW\frac{\left(2mE_F\right)^{3/2}}{3m\pi^2\hbar^2} \tag{3.24}$$

which can be used to evaluate eq. (3.23) and find

$$I_D = \left[\left(\frac{W k_F}{\pi}\right)\frac{2q^2}{h}\right]V_D. \tag{3.25}$$

Our derivation assumes that the transistors width, W, in the z-direction is large so that we can integrate over the transverse modes. As illustrated in Fig. 3.11, we impose periodic boundary conditions so that the allowed k-states in the z-direction are spaced by $2\pi/W$. For a given maximum wavevector in the z-direction, k_F, the number of transverse k-states are occupied is

$$M = 2\times\frac{k_F}{2\pi/W} \tag{3.26}$$

so the channel conductance of a ballistic MOSFET at T = 0, from eq. (3.25), assumes a familiar form

$$G_{CH} = M\frac{2q^2}{h}. \tag{3.27}$$

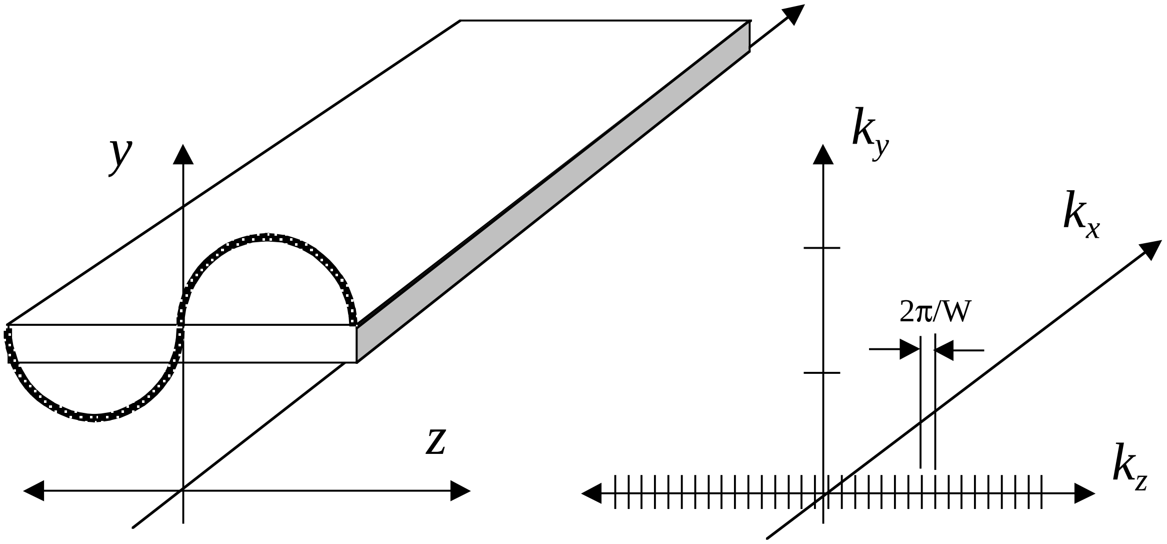

Fig. 3.11 Illustration of how period boundary conditions are applied in the transverse (z) direction (i.e. the MOSFET width) and the resulting quantization of k-states in the z-direction

3.5.2 The T = 0K on-current

The transistor's on current consists of I^+ alone, so from Eq. (3.24)

$$I_D(\text{on}) = qW\frac{\hbar k_F^3}{3\pi^2 m}, \tag{3.28}$$

which needs to be related to the gate voltage. Recall that for a two-dimensional carrier gas,

$$n_S = \frac{n_S^+}{2} = \frac{k_F^2}{4\pi} = \frac{C'_{ox}\left(V_G - V_T\right)}{q} \tag{3.29}$$

(the extra factor of two in the denominator occurs because only positive k-states are occupied under on-state conditions). Using Eq. (3.29) in (3.28), we find the on-current as

$$I_D(on) = WC'_{ox}\left(V_G - V_T\right)\left[\left(8\hbar/3m\right)\sqrt{C'_{ox}\left(V_G - V_T\right)/q\pi}\right], \tag{3.30}$$

which shows that the characteristic exponent that describes the gate voltage dependence of the on-current is $\alpha = 3/2$. In the previous section, we saw that $\alpha = 1$ for nondegenerate conditions, so ballistic MOSFET will display a characteristic exponent of $1 \le \alpha \le 3/2$. This is the same range of exponents observed for sub-micron MOSFETs that presumably operate far below the ballistic limit, so the gate voltage dependence doesn't provide a clear signature for identifying a ballistic MOSFET.

It is also useful towrite the on-current as

$$I_D(\text{on}) = W\big(qn_S(0)\big)\langle \upsilon(0) \rangle \tag{3.31}$$

which can be re-expressed as

$$I_D(\text{on}) = WC'_{ox}\big(V_G - V_T\big)\left(\frac{4\upsilon_F}{3\pi}\right). \tag{3.32}$$

From the corresponding expression for the nondegenerate case, eq. (3.15), we see that for degenerate conditions,

$$\upsilon_T \to \frac{4\upsilon_F}{3\pi}, \tag{3.33}$$

which is both greater than υ_T and gate voltage dependent,

$$\langle \upsilon(0) \rangle = \frac{4\upsilon_F}{3\pi} = \frac{8\hbar}{3m}\sqrt{\frac{C'_{ox}\big(V_G - V_T\big)}{q\pi}}\,. \tag{3.34}$$

The 3/2 power characteristic of the gate voltage dependence of the on-current arises because the carrier density varies as $(V_G - V_T)$ and the carrier velocity as the square root of the same quantity.

We can also relate the T = 0 on-current to the channel conductance by using Eq. (3.25) and (3.28) to write

$$G_{CH} = \frac{I_D(on)}{\left(2E_F / 3q\right)}. \tag{3.35}$$

From the corresponding expression for the nondegenerate case, eq. (3.17), we see that for degenerate conditions,

$$\frac{2k_BT}{q} \rightarrow \frac{2E_F}{3q}.$$

3.5.3 The T = 0K drain saturation voltage

At T = 0K, the drain current saturates when the drain Fermi level drops below the conduction band. Figure. 3.12 shows the location of the Fermi level under equilibrium conditions (V_D = 0) and for $V_D = V_{Dsa.}$ For V_D = 0, both positive and negative k-states are occupied and the location of the Fermi level is determined from

$$n_S(0) = \frac{C'_{ox}\left(V_G - V_T\right)}{q} = g_{2D}E_F, \tag{3.36}$$

where g_{2D} is the two dimensional density of states. For drain voltages beyond V_{Dsat}, only positive k-states are occupied, so the location of the Fermi level is determined from

$$n_S(0) = \frac{C'_{ox}\left(V_G - V_T\right)}{q} = \frac{g_{2D}}{2}E'_F. \tag{3.37}$$

The result is that E'_F, the location of the source Fermi level under high drain bias, is approximately twice its value in equilibrium. (We are assuming an electrostatically well designed MOSFET in which the inversion layer density at the top of the barrier is approximately constant with drain bias.) The mechanism for pushing the Fermi level higher into the subband for higher drain voltages is an electrostatic one. The source to channel barrier lowers to allow more carriers in from the source.

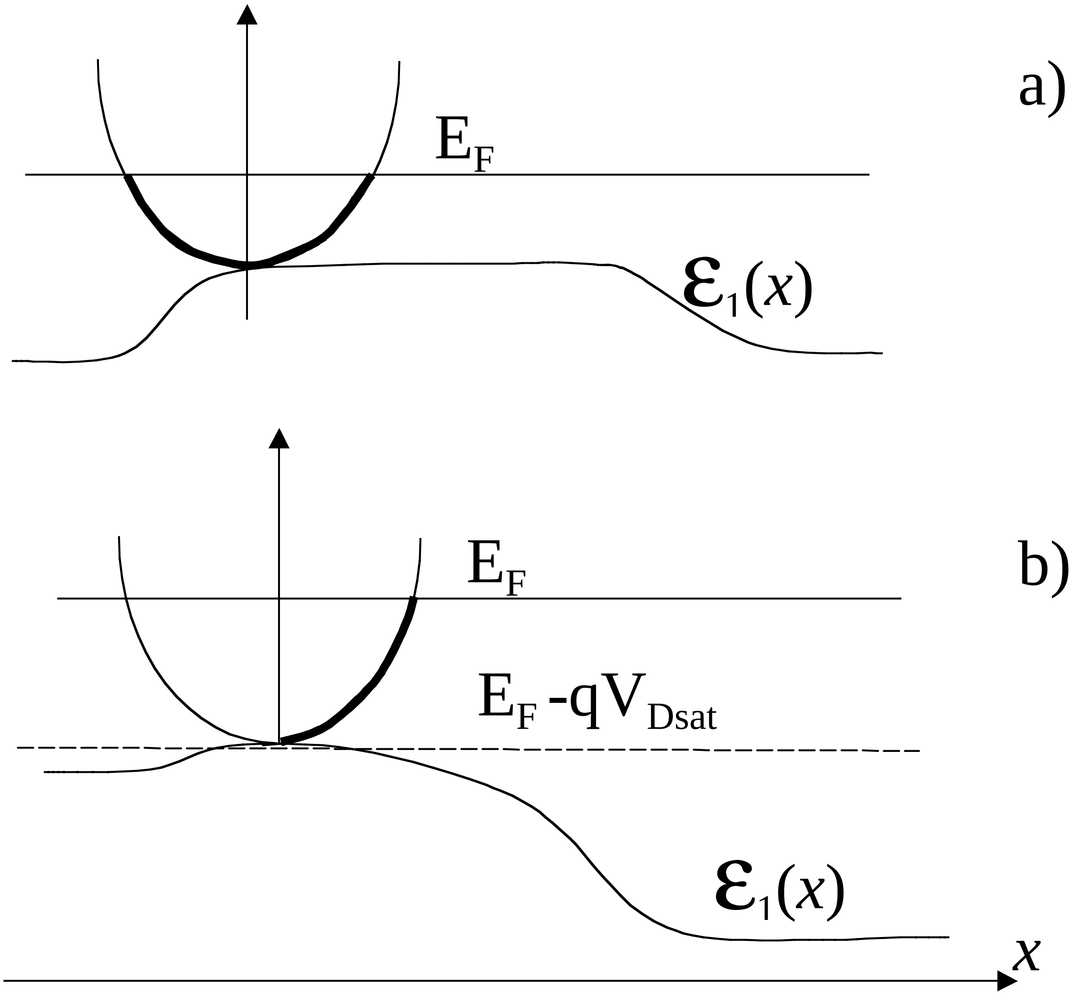

Fig. 3.12 Illustration of how the Fermi level location, E_F, and the source-channel barrier height differ at (a) $V_D = 0$ (a) and (b) $V_D = V_{Dsat}$.

Using these ideas, we see that drain current saturation occurs when $V_D = E_F'$, so

$$V_{Dsat} = \left(\frac{C_{ox}'}{q^2 (g_{2D}/2)} \right) (V_G - V_T). \tag{3.38}$$

For a long channel MOSFET, the drain saturation voltage is (V_G – V_T). The extra factor in Eq. (3.38) is typically much less than one, so for a ballistic MOSFET at T = 0K, the drain saturation voltage is much smaller than for a long channel MOSFET. Finally, since the drain saturation voltage is simple E_F/q, the eq. (3.35) can be used to show

$$V_{Dsat} = 1.5 \times \left(\frac{I_D(\text{on})}{G_{CH}} \right), \tag{3.39}$$

which gives some indication of how the I-V characteristic rounds off as saturation is approached.

3.6 The Ballistic MOSFET I-V - III

Having established the important features of ballistic MOSFETs, we are ready to derive a general expression for arbitrary levels of carrier degeneracy. We begin with eq. (3.1) and recall that for two-dimensional carriers

$$n_S^+ = \frac{N_{2D}}{2} \log\left(1 + e^{(E_F - \varepsilon_1)/k_B T}\right) = \frac{N_{2D}}{2} \mathcal{F}_0(\eta_F) \tag{3.40}$$

where $\mathcal{F}_0$ is the fermi-Dirac integral of order 0 [15] and $\eta_F = (E_F - \varepsilon_1)/k_B T$. A similar expression with η_F replaced by (η_F – U_D) [where $U_D = V_D/(k_B T/q)$] applies to electrons in the negative k-states. From eq. (3.1), we find

$$n_S^+(0) = \frac{n_S(0)}{\left(1 + n_S^- / n_S^+\right)}, \tag{3.41}$$

where

$$\frac{n_S^+(0)}{n_S^-(0)} = \frac{\mathcal{F}_0(\eta_F)}{\mathcal{F}_0(\eta_F - U_D)}. \tag{3.42}$$

To evaluate the current, recall that the average carrier velocity of the positive stream depends on the location of the Fermi level according to

$$\langle \upsilon^+ \rangle = \upsilon_T \frac{\mathcal{F}_{1/2}(\eta_F)}{\mathcal{F}_0(\eta_F)} \tag{3.43a}$$

and similarly for the negative stream,

$$\langle \upsilon^- \rangle = \upsilon_T \frac{\mathcal{F}_{1/2}(\eta_F - U_D)}{\mathcal{F}_0(\eta_F - U_D)}. \tag{3.43b}$$

Instead of eq. (3.2) for the net drain current, we find

$$I_D = I^+ - I^- = qW\left[n_S^+(0)\langle \upsilon^+ \rangle - n_S^-(0)\langle \upsilon^- \rangle\right]. \tag{3.44}$$

Finally, using eqs. (3.41) – (3.44), we obtain

$$I_D = qWn_S\,\tilde{\upsilon}_T\left[\frac{1 - \mathcal{F}_{1/2}(\eta_F - U_D)/\mathcal{F}_{1/2}(\eta_F)}{1 + \mathcal{F}_0(\eta_F - U_D)/\mathcal{F}_0(\eta_F)}\right], \tag{3.45}$$

where

$$\tilde{\upsilon}_T = \sqrt{\frac{2k_BT}{\pi m}}\,\frac{\mathcal{F}_{1/2}(\eta_F)}{\mathcal{F}_O(\eta_F)}. \tag{3.46}$$

Under nondegenerate conditions, Fermi-Dirac integrals of all orders reduce to exponentials, e^{η}, so eq. (3.45) reduces to eq. (3.13). For strong carrier degeneracy, the results of Sec. 3.5 are approached.

A sharp drain saturation voltage does not exist, but it is useful to examine the results for the on-current and channel conductance and compare the to previous results. For high drain bias, the factor in square brackets reduces to unity and we find the on-current as

$$I_D(on) = W C'_{ox} \tilde{\upsilon}_T \left(V_G - V_T\right), \tag{3.47}$$

which is identical to the nondegenerate expression, ea. (3.15), except for the implicit gate bias dependence in $\tilde{\upsilon}_T$. Similarly, by expanding eq. (3.45) for small drain voltages, we find the channel conductance and can write it as

$$G_D = \left[W C'_{ox}\left(V_G - V_T\right)\frac{\upsilon_T}{2\left(k_B T/q\right)}\right]\left(\frac{\mathcal{F}_{-1/2}(\eta_F)}{F_0(\eta_F)}\right), \tag{3.48}$$

which is identical to eq. (3.16) except for the correction factor for carrier degeneracy. Finally, we have seen that there is a close relation between the ballistic on-current and the ballistic channel resistance. In the general case, eqs. (3.47) and (3.48) give

$$G_{CH} = \frac{I_D(\mathrm{on})}{(2k_B T/q)}\left(\frac{\mathcal{F}_{-1/2}(\eta_F)}{\mathcal{F}_{1/2}(\eta_F)}\right), \tag{3.49}$$

which is identical to eq. (3.17) except for the appearance of a degeneracy correction factor.

3.7 Discussion

The simple treatment outlined in this chapter is complicated in practice when multiple subbands are occupied and by two-dimensional electrostatics. These issues are the subject of several recent papers [2, 4, 5, 6, 7, 8, 13]. Our simple treatment provides insights that are useful in interpreting experimental data or comprehensive simulations, and it often provides adequate estimates of the ballistic current that can be used to assess the performance of real devices. Given the appropriate device parameters, one can readily compute the ballistic characteristics from eq. (3.45). For quick estimates, it’s often useful to know how key paramters vary with the inverrsion layer density, a quantity that can often be estimated reasonably well for a device. For example, Fig. 3.13 is a plot of the ballistic injection velocity (left axis) and the ballistic channel conductance as a function of n_S. Given an estimate of the inversion layer density at the beginning of the channel, one can readily use this plot to estimate the ballistic on-current and the ballistic channel resistance, which can be compared against actual devices.

Fig. 3.13 Ballistic injection velocity (left) and the ballistic channel conductance (right axis) vs. the inversion layer density at the beginning of the channel, $n_S(0)$. One-subband occupation at room temperature are assumed and the results were obtained from eqs. (3.46) and (3.48).

The source-channel barrier played a key role in the theory we have presented, but. it is possible to compute the ballistic I-V characteristic without reference to the barrier at all. The key quantity is $n_S(V_{GS})$, which can be computed for an MOS capacitor in equilibrium, then shifted in voltage to account for short-channel effects. Modulation of the source-channel barrier height is what provides inversion charge to balance the gate charge. Our assumption that +k-states at the top of the barrier are populated according to the source Fermi level removes the need to treat transport over the barrier explicitly (much like the “Law of the Junction” for a bipolar transistor). In practice, transport across the barrier will be less than perfect, and the effect will appear as a series resistance in the device.

3.8 Summary

The theory presented in this chapter describes the ballistic MOSFET under ideal conditions where two-dimensional electrostatics do not produce strong short channel effects. In practice, devices are further complicated by carrier scattering. The rigorous treatment of scattering is difficult, but we shall see in the next chapter, that even in the presence of scasttering, the key features discussed in this chapter retain their validity.

Chapter 3 References

[1] S. Datta, Electronic Transport in Mesoscopic Systems, Cambridge University Press, Cambridge, UK, 1997.

[2] M.S. Lundstrom and Z. Ren, "Essential Physics of Carrier Transport in Nanoscale MOSFETs," to appear in IEEE Trans. Electron Dev., Jan. 2002.

[3] E.O. Johnson, "The insulated-gate field-effect transistor - a bipolar transistor in disguise," RCA Review, 34, pp. 80-94, 1973.

[4] Z. Ren, R. Venugopal, S. Datta,, M.S. Lundstrom, D. Jovanovic, and J.G. Fossum, "The ballistic nanotransistor: A simulation study," IEDM Tech. Digest, pp. 715-718, Dec. 10-13, 2000.

[5] Z. Ren, R. Venugopal, S. Datta, and M.S. Lundstrom, "Examination of design and manufacturing issues in a 10 nm Double Gate MOSFET using Nonequilibrium Green's Function Simulation," IEDM Tech. Digest, Washington, D.C., Dec. 3-5, 2001.

[6] Y. Naveh and K.K. Likharev, "Modeling of 10-nm-scale ballistic MOSFET's," IEEE Electron Dev. Lett., 21, pp. 242-244, 2000.

[7] Konstantin Likharev, "Sub-20-nm electron devices," in *Advanced Semiconductor and Organic Nano-techniques*, Pt. 1, ed. by H. Morkoc, Academic Press, 2002.

[8] J.-H. Rhew, Zhibin Ren, and Mark Lundstrom, "Numerical study of a ballistic MOSFET," submitted for publication, 2001.

[9] Charles G. Sodini, Ping-Keung Ko, and J.L. Moll, "The Effect of High Fields on MOS Device and Circuit Performance," *IEEE Trans. Electron Dev.*, **31**, p. 1386, 1984.

[10] K. Natori, "Ballistic metal-oxide-semiconductor field effect transistor, J. Appl. Phys., 76, pp. 4879-4890, 1994.

[11] K. Natori, "Scaling limit of the MOS transistor – A Ballistic MOSFET, IEICE Trans. Electron., E84-C, pp. 1029-1036, 2001.

[12] S. Datta, F. Assad, and M.S. Lundstrom, "The Si MOSFET from a transmission viewpoint," Superlattices and Microstructures, 23, pp. 771-780, 1998.

[13] F. Assad, Z. Ren, D. Vasileska, S. Datta, and M.S. Lundstrom, "On the performance limits for Si MOSFET's: A theoretical study," IEEE Trans. Electron Dev., 47, pp. 232-240, 2000.

[14] F. Assad, Z. Ren, S. Datta, M.S. Lundstrom, and P. Bendix, “Performance limits of Si MOSFET’s,” IEDM Tech. Digest, pp. 547-549, Dec. 1999.

[15] J.S. Blakemore, *Semiconductor statistics*, New York, Macmillian, 1962

[16] P. Van Halen and D. L. Pulfrey, “Acurate short series approximations to Fermi-Dirac integrals of order,-1/2, 1/2, 1, 3/2, 2, 5/2, 3 and 7/2,” J. Appl. Phys., **57,** pp. 5271-5274, 1985