Journal of Computational Electronics 1: 481–489, 2002
© 2003 Kluwer Academic Publishers. *Manufactured in The Netherlands.*

A Landauer Approach to Nanoscale MOSFETs

M. LUNDSTROM AND J.-H. RHEW
School of Electrical and Computer Engineering, 1285 EE Building, Purdue University, West Lafayette, IN, USA
lundstro@purdue.edu
rhew@ecn.purdue.edu

Abstract. A tutorial introduction to the theory of the nanoscale MOSFET is presented. The essential physics is first reviewed by examining numerical simulations. Next, the analytical theory of the ballistic MOSFET is introduced by presenting a simple derivation based on Boltzmann statistics. The extension to Fermi-Dirac statistics is then discussed. To clarify the connection to well-known results from mesoscopic transport, the $T_L = 0$ K theory is then developed. Finally, qualitative arguments are used to discuss the role of scattering in nano-scale MOSFETs, and the relation of scattering theory to conventional MOSFET theory is examined.

Keywords: MOSFET, transistor, ballistic transport, nanoelectronics

1. Introduction

Silicon MOSFET channel lengths are rapidly decreasing below 100 nm where traditional modeling approaches lose validity. In this paper, we describe an alternative approach to MOSFET models based on concepts that have been widely used in mesoscopic physics (Datta 1997). Figure 1 illustrates the basic physical picture; it shows the lowest conduction subband energy, E_S, vs. position under high gate and drain bias. The height of the source-channel barrier is modulated by the gate voltage. (Note that a plot of the conduction band minimum vs. position may not display a barrier, but it is the lowest quantum confined state as plotted in Fig. 1 that is relevant.) 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 (Johnson 1973). The drain current increases as the barrier height is reduced by the increasing gate voltage. Our objective is to discuss a theory of the nano-MOSFET based on this physical picture. The paper presents a tutorial introduction to recent work is presented.

2. Physics 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 (Ren *et al*. 2000, 2001) (see also, Naveh and Likharev (2000) for an earlier quantum scale study of 10 nm MOSFETs).

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 (Ren *et al*. 2000, 2001). We will, therefore, adopt a classical model in this paper. Scattering complicates the analysis, but some key elements of the essential physics can be established by numerical simulations of ballistic MOSFETs. Figure 2, for example, shows the numerically computed carrier distribution vs. position for a 10 nm channel length ballistic, double gate MOSFET operating under high-bias (Rhew, Ren and Lundstrom 2002). The carrier distribution function is seen to be strongly distorted from its equilibrium

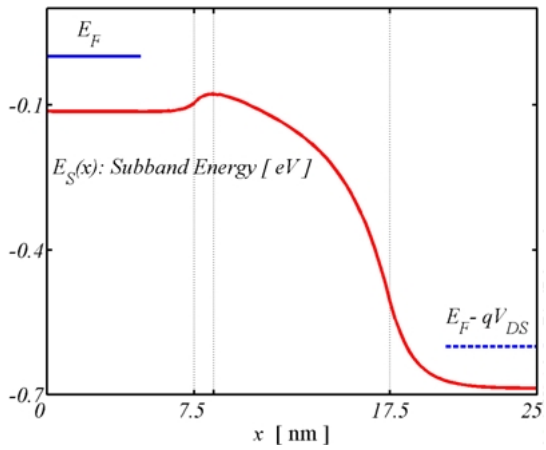

Figure 1. Conduction subband energy vs. position for a MOSFET under high gate and drain bias.

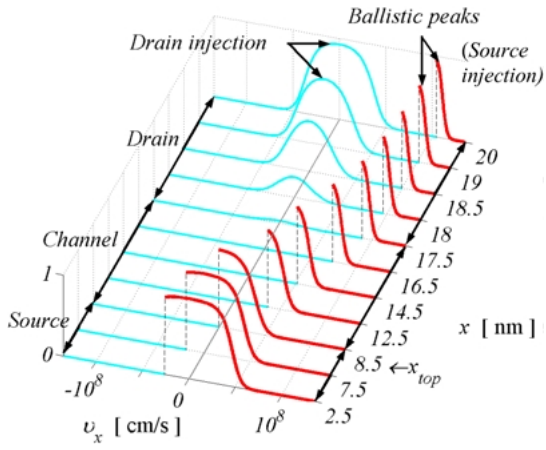

Figure 2. Electron velocity distribution function vs. position for a ballistic n-MOSFET as computed under high gate and drain bias (from Rhew, Ren, and Lundstrom (2002)).

value. Note, for example, the ballistic peak within the channel.

The complex shape of the carrier distribution function would seem to preclude simple, analytical modeling, but Fig. 3 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 (Rhew, Ren and Lundstrom 2002). For $V_{DS} = 0$, the velocity 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 ballistic distribution diminishes. Note, however, that the positive half

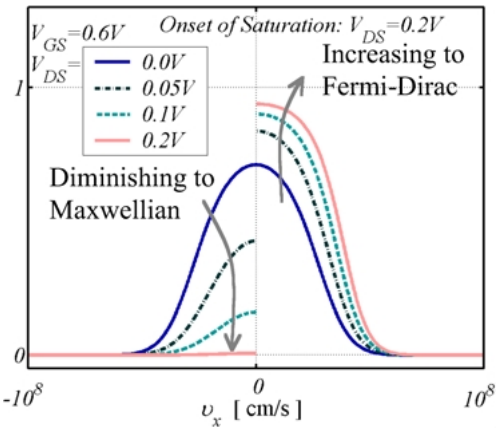

Figure 3. Electron velocity distribution function at the top of the barrier as computed for four different drain bias (from Rhew, Ren and Lundstrom (2002)).

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 (Lundstrom and Ren 2002). 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.

The key concepts illustrated by the numerical simulation are summarized as follows. First, the ballistic 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. Second, 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. Finally, 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. (There has been some discussion about what to call this velocity. The term, thermal velocity, sometimes implies a nondegenerate assumption, but Fermi-Dirac statistics are important. On the other hand, the Fermi velocity implies full degeneracy and is only the velocity at the Fermi level, whereas the appropriate velocity for our purposes if the average velocity of all carriers. Regardless of what we call it, the appropriate velocity for our purposes is the average velocity of a thermal equilibrium, hemi-Fermi-Dirac distribution.)

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

3. Theory of the Ballistic MOSFET ($T_L > 0$ K)

Figure 4 sketches the $E - k_x$ relation at the top of the source-channel barrier. The positive k_x-states are populated by injection from the source according to the source Fermi level, E_F, and the negative k_x-states are populated from the drain according to the drain Fermi level, $E_F - qV_{DS}$. By assuming 2D statistics for the quantum confined carriers 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

$$C_G(V_{GS} - V_T) \approx n_S^+(E_F) + n_S^-(E_F - qV_{DS}). \quad (1)$$

Equation (1) is an equation for the location of the Fermi level. For a given device design, the gate capacitance, C_G, and threshold voltage, V_T, are determined. Equation (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 (Lundstrom and Ren 2002).

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

$$I_D = I^+(E_F) - I^-(E_F - qV_{DS}). \quad (2)$$

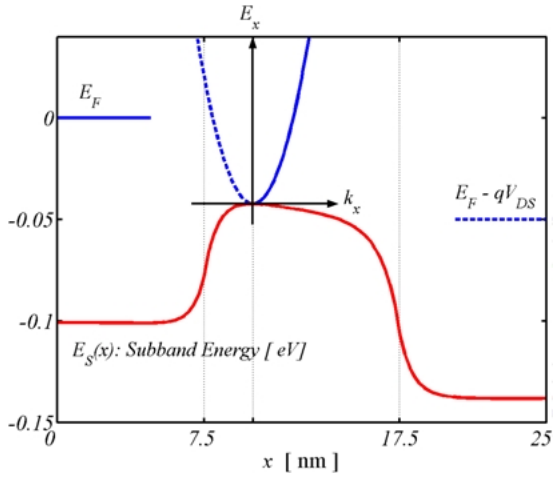

Figure 4. $E - k_x$ relation at the top of the source-channel barrier showing the source and drain Fermi levels, E_F and $E_F - qV_{DS}$.

The approach outlined here is essentially Natori's theory of the ballistic MOSFET (Natori 1994, 2001), which has been further developed by the Purdue group (Datta, Assad and Lundstrom 1998, Assad *et al.* 2000, Lundstrom and Ren 2002).

3.1. *The Nondegenerate Ballistic MOSFET*

Using the approach outlined above, the $I_{DS} - V_{DS}$ characteristic of a ballistic MOSFET is readily evaluated. The derivation is simplest when nondegenerate carrier statistics are assumed, but above threshold, the carrier gas is degenerate (Assad *et al.* 2000). In this paper, we will present a tutorial derivation assuming Boltzmann statistics, and then quote the more general results.

We begin with the fact that the inversion layer density at the beginning of the channel is composed of a positive-velocity component (populated by injection from the source) and a negative velocity component (populated by injection from the drain),

$$n_S(0) = n_S^+ + n_S^- = n_S^+(1 + n_s^-/n_S^+). \quad (3)$$

At the beginning of the channel, the average carrier velocity *of the positive-velocity carriers* is the average velocity of a thermal equilibrium hemi-Maxwellian distribution,

$$\upsilon(0) = \upsilon_T = \sqrt{2k_BT_L/\pi m^*}, \quad (4)$$

so the ballistic current carried by the positive half of the distribution is

$$I_b^+ = qWn_S^+\upsilon_T \quad (5)$$

with a similar expression for I_b^-. Using Eqs. (5) and (4), we obtain,

$$I_D = W[qn_s(0)]\upsilon_T\left(\frac{1 - n_S^+/n_S^-}{1 + n_S^+/n_S^-}\right). \quad (6)$$

In the nondegenerate limit,

$$n_S^+ = 2\left(\frac{m^*k_BT_L}{2\pi\hbar^2}\right)e^{(E_F - E_S)/k_BT_L}, \quad (7a)$$

and

$$n_S^- = 2\left(\frac{m^* k_B T_L}{2\pi\hbar^2}\right) e^{(E_F - E_S - qV_{DS})/k_B T_L}, \qquad (7b)$$

where E_F is the Fermi level of the source, and $E_F - qV_{DS}$ is the Fermi level of the drain. The bottom of the first subband is at energy, E_S. We assume [100] silicon with only the lowest subband occupied that consists of two ellipsoidal valleys along k_z with $m_x = m_y = m_t$. Thus, the appropriate effective mass to use in Eqs. (4) and (7) is $m^* = m_x = m_t$ and the factor of 2 appears in Eqs. (7a) and (7b) to count the valley degeneracy. Finally, using Eqs. (7a) and (7b) and (1) in Eq. (6), we obtain

$$I_D = WC_{ox}(V_{GS} - V_T)\,\upsilon_T\left(\frac{1 - e^{-qV_{DS}/k_B T_L}}{1 + e^{-qV_{DS}/k_B T_L}}\right), \qquad (8)$$

which is the ballistic I_D vs. V_{DS} and V_{GS} characteristic assuming Boltzmann statistics.

Although Eq. (8) assumes nondegenerate statistics, which do not apply above threshold, the result provides some insight into the performance of ballistic nanotransistors. Under nondegenerate conditions, the drain current saturates when $V_{DS} > V_{DSAT} \approx 2(k_B T_L/q)$. For large drain bias, Eq. (8) reduces to

$$I_D(on) = WC_{ox}\upsilon_T(V_{GS} - V_T). \qquad (9)$$

It is interesting to note that the on-current expression for the ballistic MOSFET is similar to the conventional, velocity saturated model of the short channel MOSFET, except that the velocity is not the high-field saturated velocity but the average velocity of a thermal equilibrium hemi-Maxwellian population. Velocity saturation does occur in the ballistic MOSFET, but the velocity saturates at the top of the barrier (the beginning of the channel) where the electric field is zero—not in the high field region near the drain.

Equation (8) also shows that the ballistic MOSFET has a finite channel conductance. For small drain bias, the exponentials can be expanded to find

$$I_D = \left[WC_G(V_{GS} - V_T)\frac{\upsilon_T}{2(k_B T_L/q)}\right]V_{DS} = G_{CH}V_{DS}. \qquad (10)$$

3.2. *Generalization to Arbitrary Carrier Statistics*

In practice, nondegenerate statistics do not apply above threshold, and the derivation presented above must be generalized. The derivation has been presented before (Natori 1994, Assad *et al*. 2000); only the result for single subband occupancy will be quoted. The derivation proceeds much like that for the nondegenerate case, but the general expressions for the Fermi level dependent carrier densities and average velocities must be used. One complicating factor is the fact that the average velocities of the positive and negative halves of the distribution differ because their Fermi levels differ. Near $V_{DS} = 0$, however, the two velocities are equal, and for high V_{DS}, only the velocity of the positive half matters, so we need not distinguish between the two velocities.

The generalization of Eq. (8) to degenerate carrier statistics and single subband occupancy is

$$I_D = WC_G(V_{GS} - V_T)\tilde{\upsilon}_T \times \left(\frac{1 - F_{1/2}(\eta_F - U_{DS})/F_{1/2}(\eta_F)}{1 + F_0(\eta_F - U_{DS})/F_0(\eta_F)}\right) \qquad (11)$$

where $\eta_F = (E_F - \varepsilon)/k_B T_L$ and $U_{DS} = qV_{DS}/k_B T_L$ and $F_{1/2}$ and F_0 are Fermi–Dirac integrals of order 1/2 and 0 and

$$\tilde{\upsilon}_T = \upsilon_T \frac{F_{1/2}(\eta_F)}{F_0(\eta_F)}. \qquad (12)$$

It is interesting to compare Eqs. (9) and (10) for the nondegenerate on-current and channel conductance to the corresponding expressions for general carrier statistics. From Eq. (11), the on-current is

$$I_D = W\,C_G\,(V_{GS} - V_T)\;\tilde{\upsilon}_T. \qquad (13)$$

For a silicon MOSFET with the lowest subband occupied, $\upsilon_T \approx 1.2 \times 10^7$ cm/s, but under typical on-current conditions, $\tilde{\upsilon}_T > 1.5 \times 10^7$ cm/s (Assad *et al*. 2000). Note that because $\tilde{\upsilon}_T$ depends on the location of the Fermi level, and therefore on the gate voltage, the on-current varies as $(V_{GS} - V_T)^\alpha$. We have seen that in the nondegenerate limit, $\alpha = 1$, but we will show in the next section that in the degenerate limit, $\alpha = 1.5$. In practice, therefore, the ballistic MOSFET should have $1 < \alpha < 1.5$.

Finally, by expanding Eq. (11) for small V_{DS}, we obtain the channel conductance as

$$G_{CH} = \left[WC_G(V_{GS} - V_T)\frac{\upsilon_T}{2(k_B T_L/q)}\right]\left(\frac{F_{-1/2}(\eta_F)}{F_0(\eta_F)}\right), \qquad (14)$$

which is identical to Eq. (10) except for a correction factor that accounts for carrier degeneracy.

4. Theory of the Ballistic MOSFET ($T_L = 0$ K)

In this section, we will derive: (1) the linear-region channel conductance G_{CH}, (2) the on-current, and (3) the drain saturation voltage, V_{DSAT}, of a degenerate ballistic MOSFET at $T_L = 0$ K, which are three key parameters of the I-V characteristics. The derivation will relate the ballistic MOSFET to a key result of mesoscopic transport.

4.1. The Linear-Region Channel Conductance G_{CH}

For the subband energy profile under *low* V_{DS} in Fig. 4, the source and the drain contacts inject the currents I_b^+ and I_b^- respectively into the device yielding the net current I_D. First, we will evaluate the current I_b^+ as a function of Fermi-level E_F at the top of the barrier. Note that I_b^- has the same functional form as I_b^+ except that E_F is replaced by $E_F - qV_{DS}$. Then we will derive the channel conductance, G_{CH}.

At $T_L = 0$ K the Fermi-Dirac distribution becomes $f(E) = 1$ if $0 < E < E_F$ and zero elsewhere. Here we implicitly assume that the subband energy at the top of the barrier is set to zero as a reference. The current I_b^+ is obtained from

$$I_b^+ = 2 \times \frac{qW}{A} \sum_{\substack{k_x > 0 \\ k_y}} \upsilon_x f(k_x, k_y) \tag{15}$$

where for parabolic energy bands, $\hbar^2(k_x^2 + k_y^2) = \hbar^2 k^2 = 2m_t E$. In Eq. (15), A is a normalization area, $\upsilon_x = \hbar k_x / m_t$, and m_t is the transverse effective mass of six ellipsoidal valleys of Si. The factor of 2 comes from the two-valley degeneracy of the lowest subband (Assad *et al*. 2000). Converting the sum into an integral with respect to E, we get

$$I_b^+(E_F) = I_b^-(E_F) = 2qW \frac{(2m_t E_F)^{3/2}}{3m_t \pi^2 \hbar^2}. \tag{16}$$

For small V_{DS}, we manipulate Eq. (2) to obtain

$$\frac{I_D}{qV_{DS}} = \frac{I_b^+(E_F) - I_b^-(E_F - qV_{DS})}{qV_{DS}} \cong \left(\frac{\partial I^+}{\partial E_F}\right). \tag{17}$$

Then, from Eqs. (16) and (17), the channel conductance becomes Landauer's formula (Datta 1997)

$$G_{CH} = q\left(\frac{\partial I^+}{\partial E_F}\right) = M\frac{2q^2}{h}, \tag{18}$$

with the number of transverse modes identified as

$$M = \left(2 \times 2 \times \frac{k_F}{2\pi / W}\right), \tag{19}$$

where the Fermi wavevector is

$$k_F = \sqrt{2m_t E_F}/\hbar. \tag{20}$$

We assume that the transistor width W is large enough so that we can count the number of states in the transverse k_y applying periodic boundary conditions; the spacing between neighboring k_y-states is $2\pi / W$. In order to conduct in x direction, it should be satisfied that $k_x^2 + k_y^2 < k_F^2$, i.e. the maximum of k_y is k_F. Thus, the number of transverse modes participating in conduction is given as Eq. (19) (one factor of 2 for two-valley degeneracy and the other for counting positive and negative k_y states) (Datta 1997).

4.2. The $T_L = 0$ K On-Current

The on-current is achieved at $V_{GS} = V_{DS} = V_{DD}$ where the injection from the drain is totally suppressed, i.e. $n_S^- = 0$, $I^-(E_F - qV_{DS}) = 0$, and $I_D(on) = I^+(E_F)$. Using the relations in (Assad *et al*. 2000), we get

$$n_S = n_S^+ = \frac{m_t}{\pi\hbar^2} E_F = C_G(V_{DD} - V_T)/q, \tag{21}$$

Inserting Eq. (21) into Eq. (16), we obtain

$$I_D(on) = I^+(E_F) = WC_G(V_{DD} - V_T) \times \left[\frac{4\hbar}{3m_t}\sqrt{\frac{2C_G(V_{DD} - V_T)}{q\pi}}\right]. \tag{22}$$

From Eqs. (21) and (22), the on-state carrier injection velocity at the top of barrier is (Assad *et al*. 2000)

$$\upsilon_{inj}(on) = \frac{4\hbar}{3m_t}\sqrt{\frac{2C_G(V_{DD} - V_T)}{q\pi}}. \tag{23}$$

Note that $I_D(on) \propto C_G(V_{DD} - V_T)^{3/2}$, which is similar to the result of conventional MOSFET models.

4.3. The Drain Saturation Voltage V_{DSAT}

At $T_L = 0$ K, the drain current saturates when the drain Fermi level drops below the top of the conduction subband. At $V_{DS} = 0$, both positive and negative k_x-states

at the top of the barrier are occupied and the location of the Fermi level is determined from

$$n_S = 2n_S^+ = C_G(V_{GS} - V_T)/q = \frac{2m_t}{\pi\hbar^2}E_F. \quad (24)$$

For $V_{DS} > V_{DSAT}$, only positive k_x-states are occupied, so the location of the Fermi level at the saturation is determined from

$$n_S = n_S^+ = C_G(V_{GS} - V_T)/q = \frac{m_t}{\pi\hbar^2}E_{FSAT}. \quad (25)$$

Note that $E_{FSAT} = 2E_F$ if we assume an electrostatically well-designed MOSFET in which the inversion layer density at the top of the barrier for given gate bias is approximately constant with drain bias. For this to happen, the source-to-channel barrier lowers to accommodate more carriers coming from the source (Lundstrom and Ren 2002). Thus, the drain saturation voltage is simply

$$V_{DSAT} = E_{FSAT}/q = \frac{C_G}{q^2 \times m_t/\pi\hbar^2}(V_{GS} - V_T). \quad (26)$$

Since $C_G = C_{OX}C_S/(C_{OX} + C_S)$ where the equilibrium semiconductor capacitance is

$$C_S = \left.\frac{\partial(qn_S)}{\partial(E_F/q)}\right|_{V_{DS}=0} = q^2\frac{2m_t}{\pi\hbar^2},$$

Eq. (26) becomes

$$V_{DSAT} = \frac{2}{1 + C_S/C_{OX}}(V_{GS} - V_T) = \frac{2C_G}{C_S}(V_{GS} - V_T) \quad (27)$$

Because $C_S \gg C_G$, the drain saturation voltage is expected to be smaller that the classical value.

5. Carrier Scattering in Nano-MOSFETs

In practice, MOSFETs are observed to deliver an on-current that is roughly 50% of the ballistic limit due to carrier backscattering (Assad *et al.* 1999, Lochtefield and Antoniadis 2001). In the presence of scattering, the on-current can be expressed in terms of a channel transmission coefficient, T (Lundstrom 1997). As for the ballistic MOSFET, the derivation is simplest for nondegenerate carrier statistics, but a more general expression involving Fermi-Dirac integrals can also be derived (Rahman and Lundstrom 2002). In this section, we derive the I-V characteristics assuming Boltzmann statistics then quote the more general result (Rahman and Lundstrom 2002).

5.1. The I-V Characteristic with Scattering

Following Eq. (3) for the ballistic MOSFET, we begin with

$$n_S(0) = n_S^+ + n_S^- = \frac{I_b^+}{WqU_T} + (1 - T)\frac{I_b^+}{WqU_T} + T\frac{I_b^-}{WqU_T} \quad (28)$$

where the first term is the injected flux (assumed to be the same as in the ballistic case), the second term is the fraction of the injected ballistic flux that back-scatters from the channel, and the third term is fraction of the drain-injected ballistic flux that transmits to the top of the barrier. In Eq. (28), we have introduced a channel transmission coefficient, T, where $0 < T < 1$. We have also assumed that all fluxes have the same average velocity, U_T, and that T is the same from source to drain as from drain to source. These assumptions are not as easy to justify as in the ballistic case, but we note that near equilibrium $T_{SD} \approx T_{DS}$ and under high drain bias, $T_{DS} \approx 0$, so there is little need to distinguish between them. By recognizing that $I_b^-/I_b^+ = e^{-qV_{DS}/k_BT}$, we can express Eq. (28) as

$$n_S(0) = \frac{I_b^+}{WqU_T}[(2 - T) + Te^{-qV_{DS}/k_BT_L}]. \quad (29)$$

To evaluate the drain current in the presence of scattering, we write

$$I_D = I^+ - I^- = I_b^+ - [(1 - T)I_b^+ + TI_b^-], \quad (30)$$

which can be expressed as

$$I_D = TI_b^+[1 - e^{-qV_{DS}/k_BT_L}]. \quad (31)$$

Finally, by combining Eqs. (29) and (31) we obtain the I-V characteristic in the presence of scattering as

$$I_D = WC_{ox}(V_{GS} - V_T)U_T\left(\frac{T}{2 - T}\right) \times \left(\frac{1 - e^{-qV_{DS}/k_BT_L}}{1 + \left(\frac{T}{2-T}\right)e^{-qV_{DS}/k_BT_L}}\right) \quad (32)$$

which generalizes Eq. (8) to include the effects of scattering. The extension to Fermi-Dirac statistics may be obtained by replacing exponentials with Fermi-Dirac integrals as we did in going from Eq. (8) to Eq. (11) (see also Rahman and Lundstrom (2002)).

To obtain the on-current in the presence of scattering, we evaluate Eq. (32) for large V_{DS} to find

$$I_D = WC_G(V_{GS} - V_T)\left[\frac{T}{2-T}\right]\upsilon_T. \qquad (33)$$

More generally, υ_T should be replaced by $\tilde{\upsilon}_T$. According to Eq. (33),

$$\langle \upsilon(0) \rangle = \upsilon_T\left(\frac{T}{2-T}\right). \qquad (34)$$

whenever scattering is present ($0 < T < 1$), the average velocity at the beginning of the channel is less than the thermal equilibrium injection velocity. For typical MOSFETs, $T \approx 0.6$ under high drain and gate bias, so the on-current is 40–50% of the ballistic limit.

By expanding Eq. (32) for small V_{DS}, we find the channel conductance as

$$G_{CH} = \left[WC_G(V_{GS} - V_T)\frac{\upsilon_T}{2(k_B T_L/q)}\right]T, \qquad (35)$$

which generalizes Eq. (10) to the non-ballistic case. Comparisons with experiment indicate that $T \approx 0.1$ for small V_{DS}, which is much smaller than the high V_{DS} value.

5.2. *The Channel Transmission Coefficient*
沟道传输系数

Understanding a nanoscale MOSFET boils down to understanding how the channel transmission coefficient depends on device structure, temperature, and bias. Under low V_{DS}, the channel transmission coefficient is just the transmission coefficient for a field-free slab of length, L (Datta 1997)

$$T = \frac{\lambda}{\lambda + L}, \qquad (36)$$

where λ is the mean-free-path for backscattering. Typically, $T < 0.1$ for low V_{DS}, but for high V_{DS} it increases to $\approx$0.6.

Computing T under high drain bias is complicated by the strong off-equilibrium transport effects that occur. It has, however, been argued that Eq. (36) can be applied to high bias condition if the channel length, L, is replaced by an effective length, ℓ, so that Eq. (36) becomes (Lundstrom 1997)

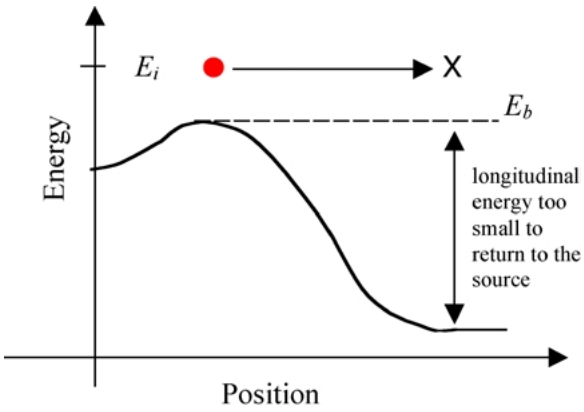

Figure 5. Illustration of an electron injected into the channel with energy, E_i, over the barrier at E_b. It backscatters at location, X, but only those with longitudinal energy above the dashed line can return to the source.

$$T = \frac{\lambda}{\lambda + \ell}. \qquad (37)$$

Because ℓ is typically a small fraction of the channel length, scattering near the source is what controls the on-current.

A rigorous derivation of the high drain bias T is complicated by the strong off-equilibrium transport effects that occur in short channel devices. A simple argument, however, explains why the on-current is most sensitive to backscattering near the source. Consider Fig. 5, which shows an electron with energy, E_i, and kinetic energy, $E_0 = E_i - E_b$, injected into the channel. It is important to note that Fig. 5 is a plot of the longitudinal energy, but an electron may backscatter at an angle (i.e. into a transverse mode). Only a fraction of the carrier that backscatter have sufficient longitudinal kinetic energy to return to the source, the others will be reflected by the channel potential and exit the drain. The fraction that backscatter with sufficient longitudinal energy to return to the source is a decreasing function of distance into the channel (Lundstrom 1997), so scattering near the source is what controls the on-current.

5.3. *Relation to Conventional MOS Theory*

We have outlined a nontraditional theory of the MOSFET, but it bears a surprising similarity to

traditional expressions. To relate the two approaches, recall that the near-equilibrium diffusion coefficient is related to the near-equilibrium mean-free-path for backscattering by (Shockley 1962)

$$D = \frac{\lambda \upsilon_T}{2} \tag{38}$$

and that the diffusion coefficient can be related to the mobility by the Einstein relation. We can, therefore, express Eq. (35) in more familiar terms as

$$G_{CH} = \frac{W}{L + \lambda} \mu_{eff} C_G (V_{GS} - V_T) \tag{39}$$

which is identical to the conventional expression except that the channel length is replaced by the channel length plus one mean-free-path. As the channel length shrinks, Eq. (39) remains finite and approaches the channel conductance of a ballistic MOSFET, Eq. (10).

Similarly, under high drain bias, Eq. (33) can be expressed in terms of the mobility. Note that we have argued that scattering in a low-field region near the beginning of the channel controls the on-current, and in this region the carriers have not been heated significantly. The result is that to first order, the near-equilibrium mean-free-path (which can be related to the near-equilibrium mobility) also controls the on-current (Lundstrom 2001). By expressing Eq. (33) in terms of the near-equilibrium mobility, we find

$$I_D = \frac{W C_G \upsilon_T m (V_{GS} - V_T) V_{DS}}{1 + m V_{DS}}, \tag{40}$$

where

$$m = \frac{\mu_{eff}}{2 L \upsilon_T} \tag{41}$$

and we have assumed that $\ell = L(k_B T_L / q V_{DS})$. Equation (41) is identical to a conventional short channel MOSFET model with the exception that υ_{sat} is replaced by υ_T (Veeraraghavan and Fossum 1988).[1]

The surprising connection between the scattering theory of the MOSFET and traditional MOS theory may be traced back to the close connection between the flux method and drift-diffusion (Shockley 1962). The fact that υ_T for silicon happens to be close to the high field saturated velocity, υ_{sat}, helps explain why traditional MOSFET models continue to provide a useful first-order description of nanoscale MOSFETs.

6. Summary

In this paper, we presented a tutorial introduction to the theory of the nanoscale MOSFET. The essential physics was first reviewed by examining numerical simulations. Next, the analytical theory of the ballistic MOSFET was introduced using a simple derivation based on Boltzmann statistics. The extension to Fermi-Dirac statistics was also discussed. To clarify the connection to well-known results from mesoscopic transport, the $T_L = 0$ K theory was developed. Finally, qualitative arguments were used to discuss the role of scattering in nanoscale MOSFETs, and to relate scattering theory to conventional MOSFET theory is examined.

Note that although the source-channel barrier played a key role in our discussions, it is possible to compute the ballistic I-V characteristic without reference to this barrier. 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.

Acknowledgment

This work was supported by the Semiconductor Research Corporation.

Note

1. This connection was first pointed out by Lixin Ge and Kuenwoo Kim at the University of Florida, March, 1999.

References

Assad F., Ren Z., Datta S., Lundstrom M.S., and Bendix P. 1999. Performance limits of Si MOSFET's. IEDM Tech. Digest, 547–549.

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

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

Datta S., Assad F., and Lundstrom M.S. 1998. The Si MOSFET from a transmission viewpoint. Superlattices and Microstructures 23: 771–780.

Johnson E.O. 1973. The insulated-gate field-effect transistor—a bipolar transistor in disguise. RCA Review 34: 80–94.

Lochtefield A. and Antoniadis D. 2001. On experimental determination of carrier velocity in deeply scaled NMOS: How close to the thermal limit? IEEE Electron Dev. Lett. 22: 95–97.

Lundstrom M.S. 1997. Elementary scattering theory of the MOSFET. IEEE Electron Dev. Lett. 18: 361–363.

Lundstrom M. 2001. On the mobility versus drain current relation for a nanoscale MOSFET. IEEE Electron Dev. Lett. 22: 293–295.

Lundstrom M.S. and Ren Z. 2002. Essential physics of carrier transport in nanoscale MOSFETs. IEEE Trans. Electron Dev. 49: 133–141.

Natori K. 1994. Ballistic metal-oxide-semiconductor field effect transistor. J. Appl. Phys. 76: 4879–4890.

Natori K. 2001. Scaling limit of the MOS transistor—A Ballistic MOSFET, IEICE Trans. Electron. E84-C: 1029–1036.

Naveh Y. and Likharev K.K. 2000. Modeling of 10-nm-scale ballistic MOSFET's. IEEE Electron Dev. Lett. 21: 242–244.

Rahman A. and Lundstrom M. 2002. A compact scattering model for the nanoscale double-gate MOSFET. In: IEEE Trans. Electron Dev. 49: 481–489.

Ren Z., Venugopal R., Datta S., Lundstrom M.S., Jovanovic D., and Fossum J.G. 2000. The ballistic nanotransistor: A simulation study. IEDM Tech. Digest 715–718.

Ren Z., Venugopal R., Datta S., and Lundstrom M.S. 2001. 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.

Rhew J.-H., Ren Z., and Lundstrom M. 2002. A numerical study of a ballistic transport in a nanoscale MOSFET 46: 1899–1906.

Shockley W. 1962. Diffusion and drift of minority carriers in semiconductors for comparable capture and scattering mean-free-paths. Phys. Rev. 125: 1570–1576.

Veeraraghavan S. and Fossum J.G. 1988. A physical short channel model for the thin-film SOI MOSFET applicable to device and circuit CAD. IEEE Trans. Electron Dev. 35: 1866–1875.