Nanoscale Transistors: Device Physics, Modeling, and Simulation

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

Chapter 4: Scattering Theory of the MOSFET

4.1 Introduction

In practice, MOSFETs operate below the ballistic limit because carrier scatter within the device. To determine how far below the ballistic limit a device operates, we can plot eq. (3.45) for the device (knowing its oxide thickness and power supply and threshold voltages) and compare the results to the measured data. Typical data for circa 2000 technology with $L_{eff} \approx 100$ nm is shown in Fig. 4.1. These results show that n-channel MOSFETs deliver roughly one-half of the ballistic on-current while p-channel MOSFET deliver roughly one-third of the ballistic current. Other studies find similar results, [1, 2] although devices that operates much closer to the ballistic limit has also been reported [3]. Our goal in this chapter is to develop an understanding of how scattering affects transistors and to develop a simple model that extends eq. (3.45).

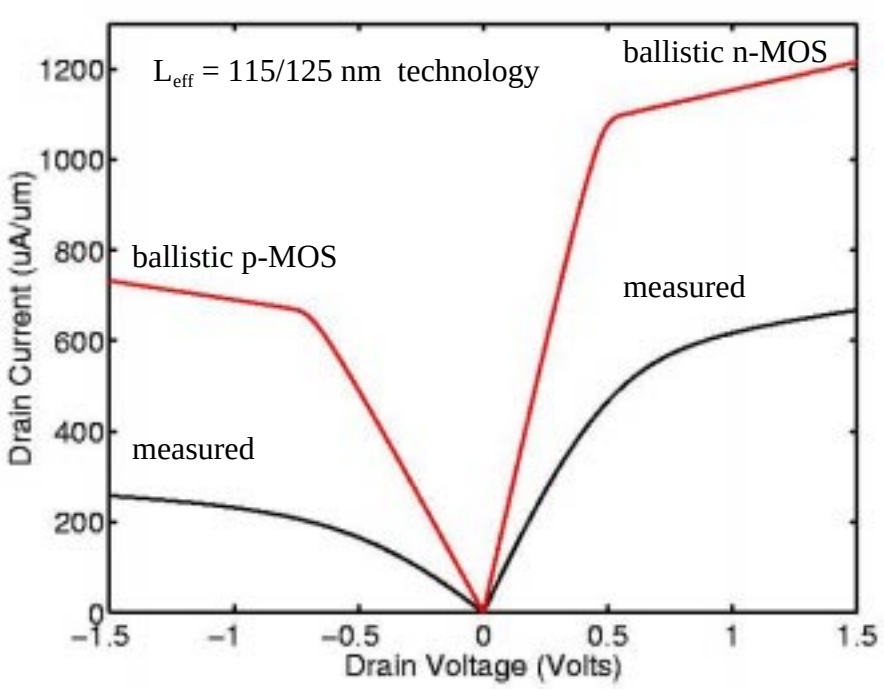

Fig. 4.1 Comparison of measured I-V characteristics of p- and n-channel MOSFETs and the corresponding ballistic limit characteristics. The measured series resistance has been added to the ballistic characteristics. (From D. Rumsey, MSEE Thesis, Purdue University, December 2001.)

The channel transmission coefficient, T, will provide us with a simple way to describe a MOSFET in the presence of scattering. Figure 4.2 defines the transmission coefficient. Under high drain bias in the ballistic MOSFET, only positive k-states at the top of the barrier are occupied, but in the presence of scattering, only a fraction, T_{SD}, of the current injected into the channel transmits across and exits through the drain. A fraction, $(1\text{-}T_{SD})$ backscatters and occupies negative k-states. Similarly, for low drain bias, a fraction, T_{DS}, of the flux injected into the drain transmits to the top of the barrier and adds to the population of the negative k-states. In general, $T_{SD} \neq T_{DS}$, but we will assume that they are equal and see why it leads to an accurate model.

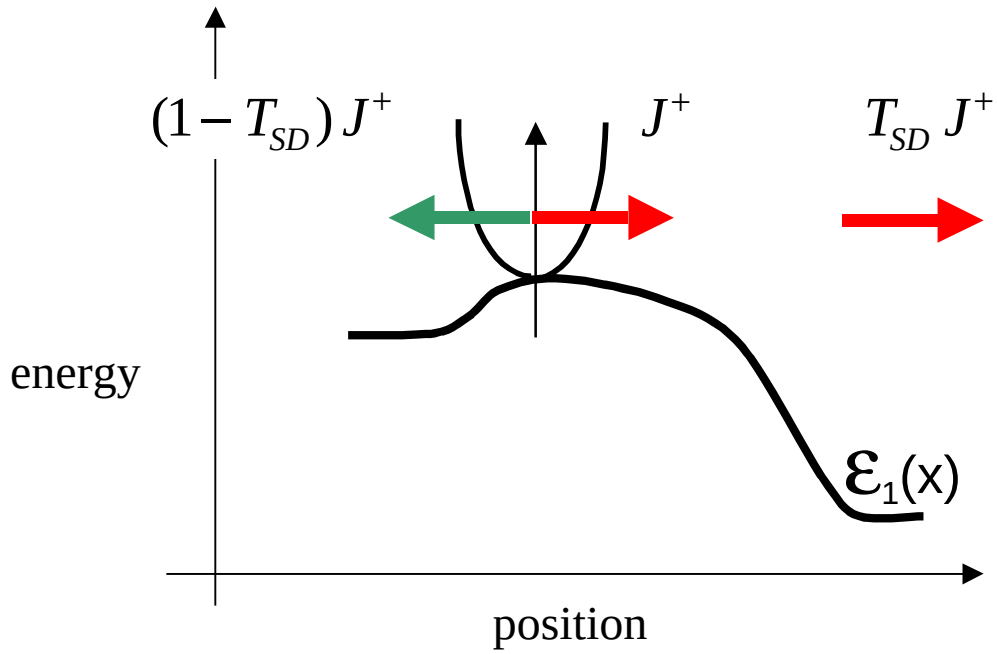

Fig. 4.2 Illustration of how backscattering populates negative k-states at the top of the barrier and reduces the current transmitted to the drain.

4.2 MOSFET physics in the presence of scattering

Before we develop simple models for MOSFETs in the presence of scattering, we will establish the essential physics by examining numerical simulations based on the nonequilibrium Green's function approach [4,5,6,7]. The device simulated is a double gate MOSFET because the simple geometry facilitates the numerical analysis, but the results are expected to apply more generally. For the ballistic MOSFET, the simulations discussed in Sec. 3.2 established 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) 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) 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.

Scattering mixes the positive and negative k-states, so the carrier distribution at the top of the barrier no longer consists of two thermal equal equilibrium halves. Under high drain bias, however, we continue to find that the positive half consists of a near thermal equilibrium population injected from the source. Even in the presence of scattering, however, Fig. 4.3 shows that MOS electrostatic should continues to apply.

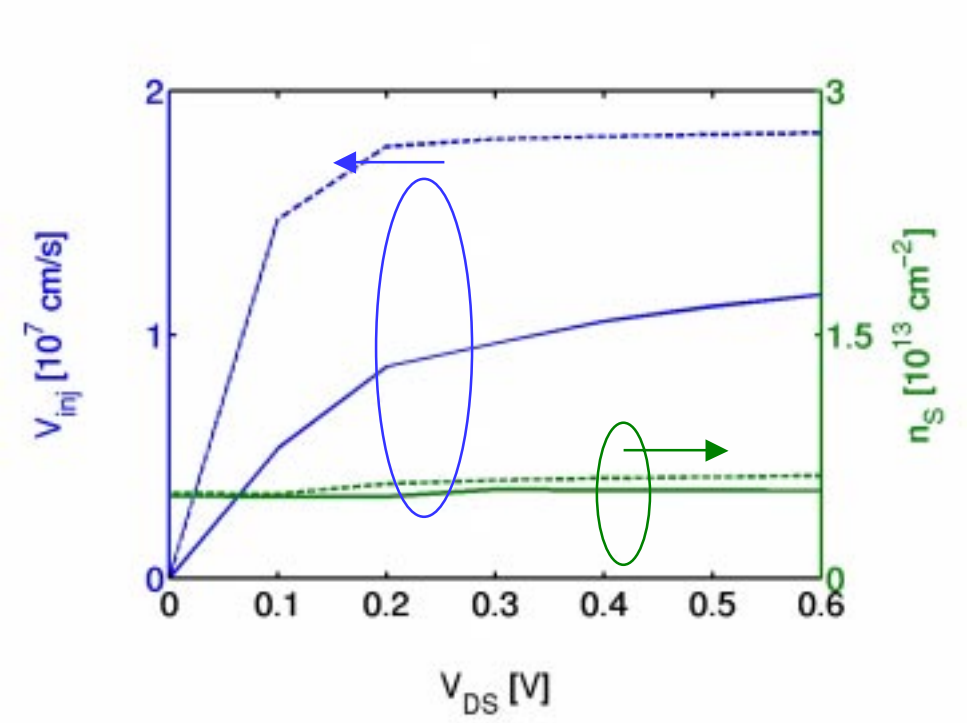

Fig. 4.3 Comparison of the average carrier velocity at the top of the barrier (left) and the inversion layer density (right) when scattering is present (solid lines) and absent (dashed lines). The device is a double gate MOSFET and the results were obtained from a nonequilibrium Green's function simulation [7].

Figure 4.4 provides some insight into how scattering affects a MOSFET. This plot shows the simulated on-current as a function of the number of scatters in the simulation. For this simulation there were 45 finite difference nodes between source and drain. The first scatterer was placed at the last node (at the drain), then with increasing numbers, the scattering progressively moves towards the source. Figure 4.4 shows that when scattering is localized to the drain region, it has little effect on the on-current, but the effect of scattering increases as scattering takes place closer to the source. Although the physical model for scattering is a

simple one (see [7] for a discussion), but the conclusion is confirmed by more rigorous simulations [8]. Simulations, therefore, show that the influence of scattering on the MOSFET's on-current is weighted towards the source.

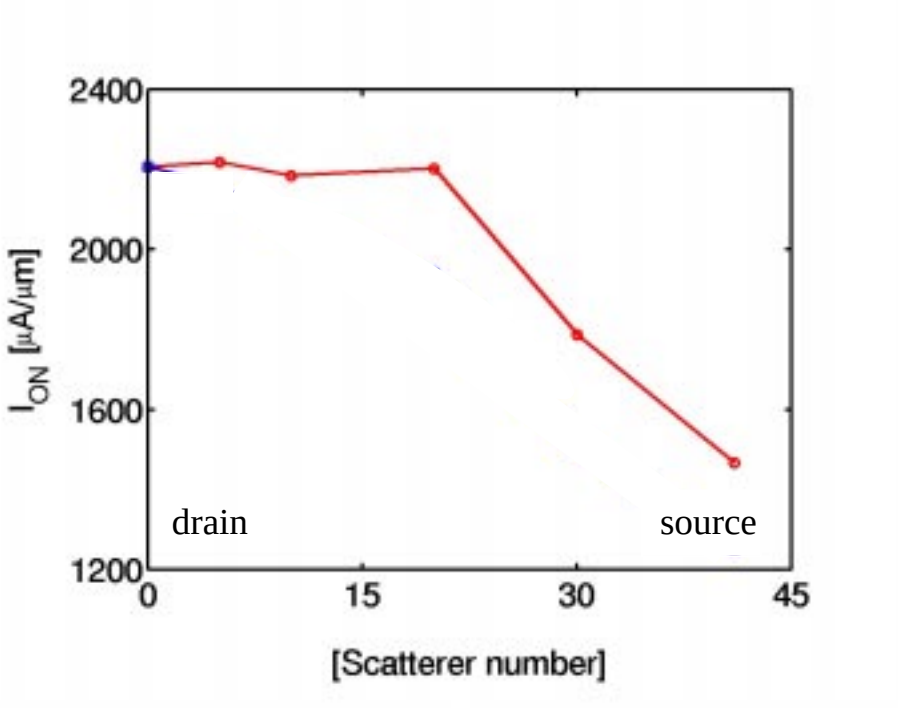

Fig. 4.4 Illustration of how carrier scattering affects the on-current of a 10 nm channel length double gate MOSFET. The channel is divided into 45 slabs, and scatterer 1 is place in the slab at the drain end. With increasing number, they move towards the source. With 45 scatterers present, scattering takes place uniformly from source to drain. (From [7].)

In summary, numerical simulations in the presence of scattering show that:

1) for an electrostatically well-designed MOSFET, the total carrier density at the top of the barrier is maintained at an approximately constant value, as it is for the ballistic MOSFET, and

2) for a fixed gate voltage, the average velocity at the top of the barrier saturates with increasing drain voltage at a limiting value that is below the average velocity of a thermal equilibrium Fermi-Dirac distribution.

3) scattering, which controls the magnitude of the saturated velocity, is most important when in occurs near the source.

We will use observation 1) in developing a scattering theory of the MOSFET in terms of a channel transmission coefficient, *T*, where $T < 1$ in the presence of scattering. We will use this simple theory to establish the limiting source velocity in the presence of scattering and to understand why scattering near the source is more important that near the drain.

4.3 The scattering model

Before we discuss the physics of scattering, we generalize our ballistic theory of the MOSFET, eq. (3.45), to include scattering as described by a channel transmission coefficient, *T*. In addition to assuming a channel transmission coefficient whose calculation will be discussed later, the derivation will require some additional assumptions that are guided by physical reasoning.

To evaluate the drain current, we begin with eq. (3.2)

$$I_D = I_D^+ - I_D^- \tag{4.1}$$

and evaluate it at the top of the barrier. For the positive current, we assume a thermal equilibrium flux as in the ballistic case,

$$I^+ = q\left(\frac{N_{2D}}{2}\right)\upsilon_T \mathcal{F}_{1/2}(\eta_F) \tag{4.2}$$

and

$$n_S^+(0) = \left(\frac{N_{2D}}{2}\right)\mathcal{F}_0(\eta_F). \tag{4.3}$$

(Recall that $\upsilon_T = \sqrt{2k_BT/\pi m}$). Our use of thermal equilibrium fluxes is a quasi-equilibrium assumption analogous to the constant quasi-Fermi level assumption of pn junction theory [9]. The negative flux consists of two components, one injected from the drain and another from the fraction of the positive flux that backscatters. In the case of ballistic transport, we saw in Chapter 3 that the contribution from the drain (the only contribution in the ballistic case) is

$$I_b^- = q\left(\frac{N_{2D}}{2}\right)\upsilon_T \mathcal{F}_{1/2}(\eta_F - U_D) \tag{4.4}$$

and

$$n_b^-(0) = \left(\frac{N_{2D}}{2}\right)\mathcal{F}_0(\eta_F - U_D). \tag{4.5}$$

When scattering is present, the contribution to the negative k-states from drain injection will be reduced by a drain-to-source transmission coefficient. Under equilibrium conditions, $T_{SD} = T_{DS}$. A drain bias of only a few k_BT/q suppresses the contribution from the drain completely, so the fact that $T_{SD} \neq T_{DS}$ under high bias is not important. We will, therefore, describe the MOSFET with a single transmission coefficient, T. We can, therefore, write the negative current at the top of the barrier as

$$I^- = (1-T)I^+ + TI_b^-, \tag{4.6}$$

where the first term is the contribution from backscattering of the injected current and the second term is the contribution from the drain. Using eqs. (4.2), (4.4), and (4.6), in eq. (4.1), we find

$$I_D = qTI^+\left[1 - \frac{\mathcal{F}_{1/2}(\eta_F - U_D)}{\mathcal{F}_{1/2}(\eta_F)}\right]. \tag{4.7}$$

To be useful, eq. (4.7) must be related to device parameters, not to internal quantities like I^+. To do so we examine the charge density at the top of the barrier,

$$n_S(0) = n_S^+(0) + n_S^-(0), \tag{4.8}$$

because it can be approximated by $C'_{os}\left(V_G - V_T\right)/q$ above threshold. For the carriers in the positive k-states,

$$n_S^+(0) = \frac{I^+}{q\tilde{\upsilon}_T{}^+} \tag{4.9}$$

where

$$\tilde{\upsilon}_T{}^+ = \upsilon_T \frac{\mathcal{F}_{1/2}(\eta_F)}{\mathcal{F}_0(\eta_F)}. \tag{4.10}$$

Similarly, for the negative ballistic stream,

$$n_b^-(0) = \frac{I_b^-}{q\tilde{\upsilon}_T{}^-} \tag{4.11}$$

where

$$\tilde{\upsilon}_T{}^- = \upsilon_T \frac{\mathcal{F}_{1/2}(\eta_F - U_D)}{\mathcal{F}_0(\eta_F - U_D)}. \tag{4.12}$$

Note that even for ballistic conditions, $\tilde{\upsilon}_T^+ \neq \tilde{\upsilon}_T^-$ because both velocities are Fermi level dependent and the source and drain Fermi levels are different. Under low drain bias, however, $\tilde{\upsilon}_T^+ \approx \tilde{\upsilon}_T^- = \tilde{\upsilon}_T$. Under high drain bias, there is no drain-injected component to worry about, but the distribution of the backscattered carriers is unknown. As a first order approximation, which appears to work well in practice, we will assume at the top of the barrier a common velocity, $\tilde{\upsilon}_T$, equal to eq. (4.10) for both positive and negative streams. With this assumption, eq. (4.6) becomes

$$\frac{I^-}{\tilde{\upsilon}_T} = (1-T)\left(\frac{I^+}{\tilde{\upsilon}_T}\right) + T\left(\frac{I'_b}{\tilde{\upsilon}_T}\right), \tag{4.13}$$

or

$$n_s^-(0) = (1-T)n_S^+(0) + T n_b^- . \tag{4.14}$$

To get the total carrier density, we insert eq. (4.14) in eq. (4.8) to find

$$n_S(0) = (2-T)\, n_S^+(0)\left[1+\left(\frac{T}{2-T}\right)\frac{\mathcal{F}_0(\eta_F - U_D)}{\mathcal{F}_0(\eta_F)}\right]. \tag{4.15}$$

Finally, using eq. (4.7) and (4.15), we find [10]

$$I_D = W C'_{ox}\left(V_G - V_T\right)\left(\frac{T}{2-T}\right)\tilde{\upsilon}_T\left[\frac{1-\dfrac{\mathcal{F}_{1/2}(\eta_F - U_D)}{\mathcal{F}_{1/2}(\eta_F)}}{1+\left(\dfrac{T}{2-T}\right)\dfrac{\mathcal{F}_0(\eta_F - U_D)}{\mathcal{F}_0(\eta_F)}}\right], \tag{4.16}$$

which is the desired result. (We have assumed above threshold operation where $q\, n_S(0) = C'_{ox}\left(V_G - V_T\right)$.)

Equation (4.16) generalizes the expression for the ballistic MOSFET, eq. (3.45), to the case where scattering in present. Note that when scattering is present, a number of additional, physically motivated assumptions were necessary in order to obtain a simple expression. The key additional assumptions are:

1) that the positive flux injected into the channel continues to be the thermal equilibrium flux injected over the barrier from the source,

2) that carriers injected from the source (drain) that enter the drain (source) are completely absorbed and do not re-enter the device,

3) that the average velocities of the positive and negative streams at the top of the barrier are equal to that of the equilibrium, positive stream.

These assumptions seem to capture the essence of the problem, but note that we have hidden a great deal of complexity in the transmission coefficient, *T*. Comparisons with experiment like that shown in Fig. 4.1 show that T is bias dependent. At low drain bias, $T \approx 0.1$ for a $L_{eff} \approx 100\,\text{nm}$ n-channel MOSFET (≈ 0.05 for a p-MOSFET) while at high drain bias, $T \approx 0.6$. for an n-MOSFET(≈ 0.5 for a p-MOSFET). Computing *T* rigorously is a difficult problem that requires detailed numerical simulations, but the essential physics of scattering in a MOSFET will be discussed in the next two sections.

Before discussing the physics of scattering, it's useful to examine eq. (4.16) under limiting conditions. First of all, it's clear that in the ballistic limit where *T* = 1, we recover the ballistic theory of the MOSFET. For high drain bias, the drain voltage factor approaches unity and we find [11]

$$I_D(on) = W C'_{ox}\left(\frac{T}{2-T}\right)\tilde{\upsilon}_T\left(V_G - V_T\right). \tag{4.17}$$

which can be used to estimate *T* from the measured on-current. Note that the average carrier velocity at the top of the barrier,

$$\langle \upsilon(0) \rangle = \left(\frac{T}{2-T}\right)\tilde{\upsilon}_T, \tag{4.18}$$

is reduced by backscattering. In the linear region, eq. (4.16) reduces to

$$I_D = W C'_{ox}\left(V_G - V_T\right)\frac{T\,\tilde{\upsilon}_T}{2\left(k_B T/q\right)}\left(\frac{\mathcal{F}_{-1/2}(\eta_F)}{F_{1/2}(\eta_F)}\right)V_D. \tag{4.19}$$

Note that the channel conductance is proportional to *T*, as might have been expected, but that the on-current has a more complicated dependence on *T*. This occurs as a result of MOS electrostatics which seeks to maintain the inversion layer density at the top of the barrier constant as the drain bias increases.

4.4 The transmission coefficient under low drain bias

Under low source to drain bias, the channel of a nanoscale MOSFET can be regarded as a region of length, L_{eff}, across which carriers diffuse. (Strictly speaking, the dominance of diffusion over drift implies that the potential drop across the channel is less than k_BT/q.) For these conditions, the channel transmission coefficient is readily evaluated.

Consider a flux of carriers in the positive direction, J^+ and a corresponding flux of carriers in the negative direction, J^-. The positive stream is reduced by scattering out but increased when the negative stream scatters in

$$\frac{dJ^+}{dx} = -\frac{J^+}{\lambda} + \frac{J^-}{\lambda}, \tag{4.20}$$

where λ is the mean-free-path for backscattering (assumed constant) and dx/λ is the probability of scattering is a thickness, dx. A similar equation applies to the negative stream. Using the net flux, $J = \left(J^+ - J^-\right)$, to eliminate J^- in eq. (4.20), we find

$$\frac{dJ^+}{dx} = -\frac{J}{\lambda}, \tag{4.21}$$

which can integrated to find

$$J^+(x) = J^+(0) - J\left(\frac{x}{\lambda}\right). \tag{4.22}$$

Now consider a slab of length. L_{eff}, which represents the channel under low drain bias. As shown in Fig. 4.4, we inject a flux of positive velocity carriers into the slab at x = 0, and a flux $T\,J^+(0)$ emerges from the right. Our goal is to evaluate T. Within the slab, scattering converts part of the injected positive stream to the negative stream, so we have both positive and negative fluxes, At $x = L_{eff}$, however, we have only a positive flux, so $J = J^+(L_{eff}) = T\,J^+(0)$. Using this expression in eq. (4.22), we find

$$T = \frac{J^+(L)}{J^+(0)} = \frac{\lambda}{\lambda + L_{eff}}. \tag{4.23}$$

As expected, the channel transmission coefficient approach zero when $L_{eff} >> \lambda$ and one when $L_{eff} << \lambda$.

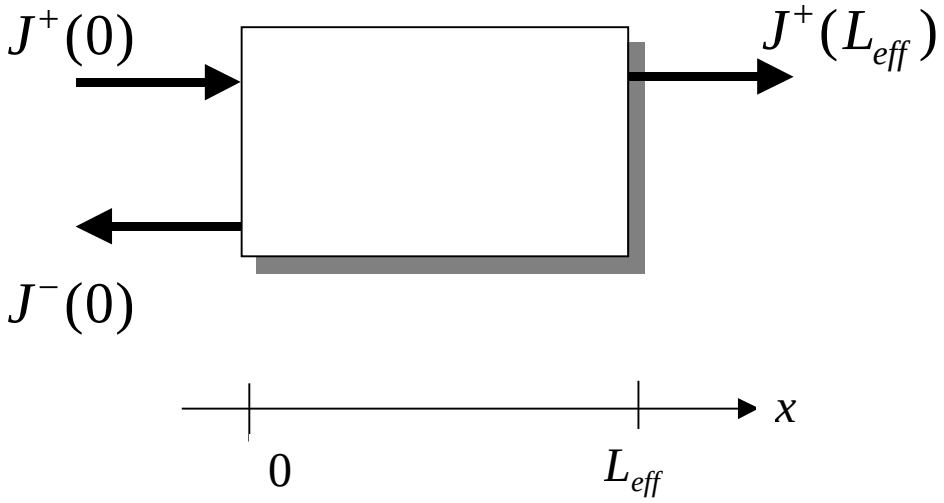

Fig. 4.4 Carrier transport across a field-free semiconductor slab. A flux is injected at $x = 0$ and absorbed at $x = L_{eff}$.

Using eq. (4.23) in the linear region current, eq. (4.19), we find

$$I_D = \left(\frac{W}{L_{eff} + \lambda} \right) \frac{(\upsilon_T \lambda / 2)}{(k_B T / q)} C'_{ox} (V_G - V_T)\ V_D, \tag{4.24}$$

where we have assumed nondegenerate carrier statistics for simplicity. To apply eq. (4.24) in practice, we need to estimate the mean-free-path for backscattering.

To relate the mean-free-path to a familiar macroscopic quantity, we consider the net flux through the slab,

$$J = \left(J^{+} - J^{-}\right) = T\,J^{+}(0). \tag{4.25}$$

The carrier density at $x = 0$ is

$$n_S(0) = \frac{J^{+}(0)}{\upsilon_T} + (1-T)\frac{J^{+}(0)}{\upsilon_T} = (2-T)\frac{J^{+}(0)}{\upsilon_T}, \tag{4.26}$$

which can be solved for $J^{+}(0)$ and used in eq. (4.25) to obtain

$$J = \left(\frac{T}{2-T}\right)\upsilon_T\, n_S(0) = \left(\frac{\lambda\upsilon_T}{2L_{eff} + \lambda}\right)n_S(0). \tag{4.27}$$

Equation (4.27) is valid whether the channel is thin or thick compared to a mean-free-path, but consider the limit when $L_{eff} >> \lambda$ where eq. (4.27) becomes

$$J = \left(\lambda\upsilon_T / 2\right)\left(\frac{n_S(0)}{L_{eff}}\right), \tag{4.28}$$

which is precisely what Fick's Law of Diffusion gives if we define the diffusion coefficient as

$$D \equiv \lambda\upsilon_T / 2 = \left(k_B T / q\right)\mu_{eff}\,. \tag{4.29}$$

The connection to the carrier mobility arises because we have assumed near-equilibrium, nondegenerate conditions for which an Einstein relation applies.

Equation (4.29) provides us with a simple way to estimate the mean-free-path if the mobility is known (subject, however, to the constraint of near-equilibrium, nondegenerate conditions). It also allows us to write eq. (4.24) in a more familiar form

$$I_D = \left(\frac{W}{L_{eff} + \lambda} \right) \mu_{eff} \, C'_{ox} \left(V_G - V_T \right) \, V_D . \tag{4.30}$$

Equation (4.30) is identical to the traditional text book expression [9] except that L_{eff} is replaced by $L_{eff} + \lambda$. As the channel length becomes much smaller than the mean-free-path, the channel conductance approaches the ballistic limit.

When parasitic series resistance is eliminated, the measured linear region current is observed to be about 0.1 for 100nm generation n-MOSFETs. From eq. (4.19), this ratio is T, so for a $L_{eff} \approx$ 100nm channel length, we find that $\lambda \approx$ 10 nm. According to eq. (4.29), this corresponds to a mobility of approximately 250 cm^2/V-s, which is roughly the expected electron mobility for a 100nm MOSFET. Within the accuracy of our simplifying assumptions, we conclude that this simple theory roughly accounts for scattering under low drain voltages.

4.5 The transmission coefficient under high drain bias

It is considerably more difficult to develop a simple expression for the channel transmission coefficient under high drain bias. For such conditions, carriers are injected into a short, high-field region and carrier transport is far from equilibrium. Well-known off-equilibrium transport effects such as velocity overshoot occur, and simple analytical treatments are hard to come by [12]. Nevertheless, we seek a simple analytical expression for T that is consistent with experiment and provides physical insight.

Figure 4.5 is a sketch of the conduction subband profile vs. position along the channel of a small MOSFET under high drain and gate bias. The channel consists of two portions, a

low field region near the source, which is under strong gate control, and a high field region near the drain. The transition between the low and high field regions is often very sharp. Transport in the channel of a MOSFET is similar to transport in a bipolar transistor. Carriers diffuse across the low-field portion of the channel (which is analogous to the base) after which they are collected by the high-field portion of the channel, which is analogous to the collector. Following this reasoning, we postulate that the transmission coefficient under high drain bias is given by

$$T = \frac{\lambda}{\lambda + \ell}, \tag{4.31}$$

which is analogous to eq. (4.23). In eq. (4.31), ℓ is a critical length, some fraction of the channel length, that has to be specified. Since diffusion dominates drift when the potential drop across the region is less than k_BT/q, we are tempted to identify ℓ as the distance over which the first k_BT/q of potential drops. If we did so, our result would be analogous to the well-known Bethe condition for thermionic emission in a metal-semiconductor diode [13]. Equation (4.31) can, in fact, be mathematically derived from McKelvey's flux theory (see [12, 14, 15]), but the flux equations are based on a near equilibrium assumption which makes the derivation little better than the plausibility argument we have given. The physical reasoning that leads to eq. (4.31) is, however, consistent with numerical experiments by Monte Carlo simulation. These simulations show that when carriers are injected into a high field region, if they penetrate more than a short distance into the high-field region, then even if they do backscatter, they are unlikely to emerge from the high-field region [16, 11]. Our task, therefore, is to identify the distance, ℓ, beyond which if a carrier scatters it will be unlikely to return to the source.

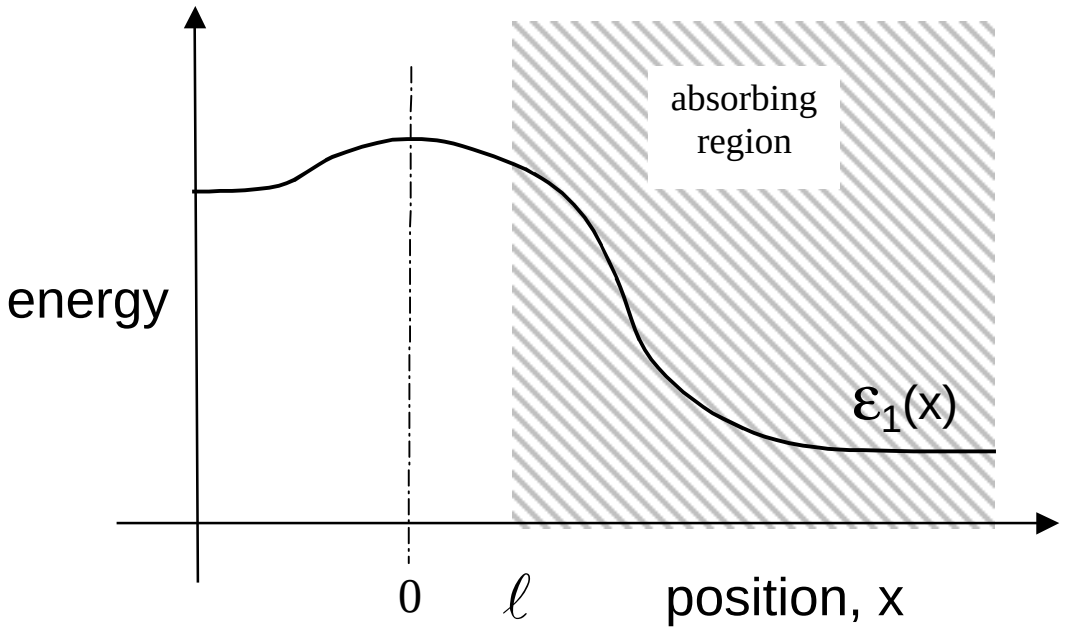

Fig. 4.5 Illustration of the conduction subband profile vs. position under high drain bias. For a well-designed MOSFET, the channel is typically divided into a low field region near the source and a high field region near the drain.

Figure 4.6a illustrates a scattering event for an electron with energy E_i (kinetic energy $(E_i - E_b)$ that is injected in the channel and scatters first at location $x = x_1$. The electron enters the channel with some momentum, $\mathbf{p_0}$, is the x-y plane, as shown in Fig. 4.6b. The longitudinal electric field in the channel accelerates the electron so that its x-directed momentum increases, At x_1, where it scatters first, its momentum is $\mathbf{p}_1$, where

$$\frac{p_1^2}{2m} = E_i + qV(x_1). \tag{4.32}$$

We assume that the electron scatters isotropically and elastically to momentum $\mathbf{p_1'}$. After scattering, an electron can surmount the potential barrier, $qV(x_1)$, and exit through the source only if its velocity is in the negative x-direction and if $\frac{1}{2}m\upsilon_x^2 > qV(x_1)$. The shaded region

in Fig. 4.6b is the only fraction that contributes to channel backscattering as we have defined it, and one can readily show that to surmount the barrier

$$\theta < \cos^{-1}\left(\frac{qV(x_1)}{qV(x_1)+(E_i-E_b)}\right). \tag{4.33}$$

The probability that an electron which scatters first at $x = x_1$ returns to the source is

$$P(x_1) = \frac{2\theta}{2\pi} = \frac{1}{\pi}\cos^{-1}\sqrt{\frac{qV(x_1)}{qV(x_1)+(E_i-E_b)}} \tag{4.34}$$

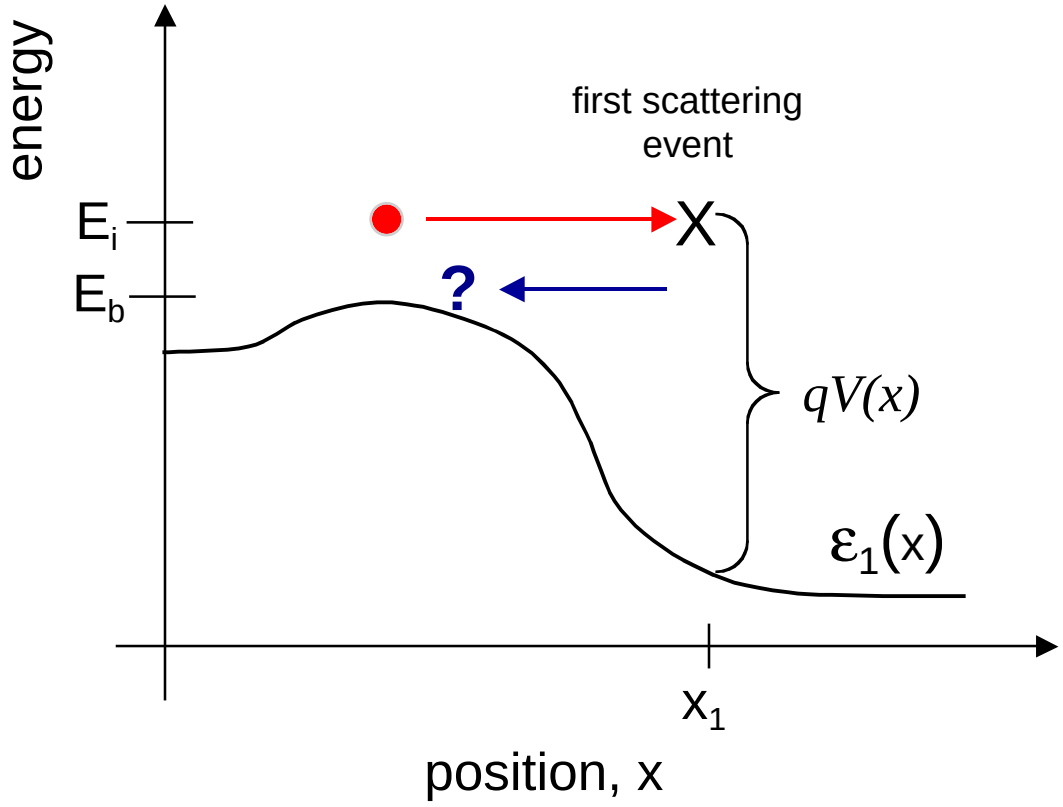

Fig. 4.6a Illustration of an electron injected into the channel with an energy, Ei, that undergoes its first scattering event at $x = x_1$. We seek to compute the probability that it returns to the source. It is important to remember that the electron is also free to move in the y-direction.

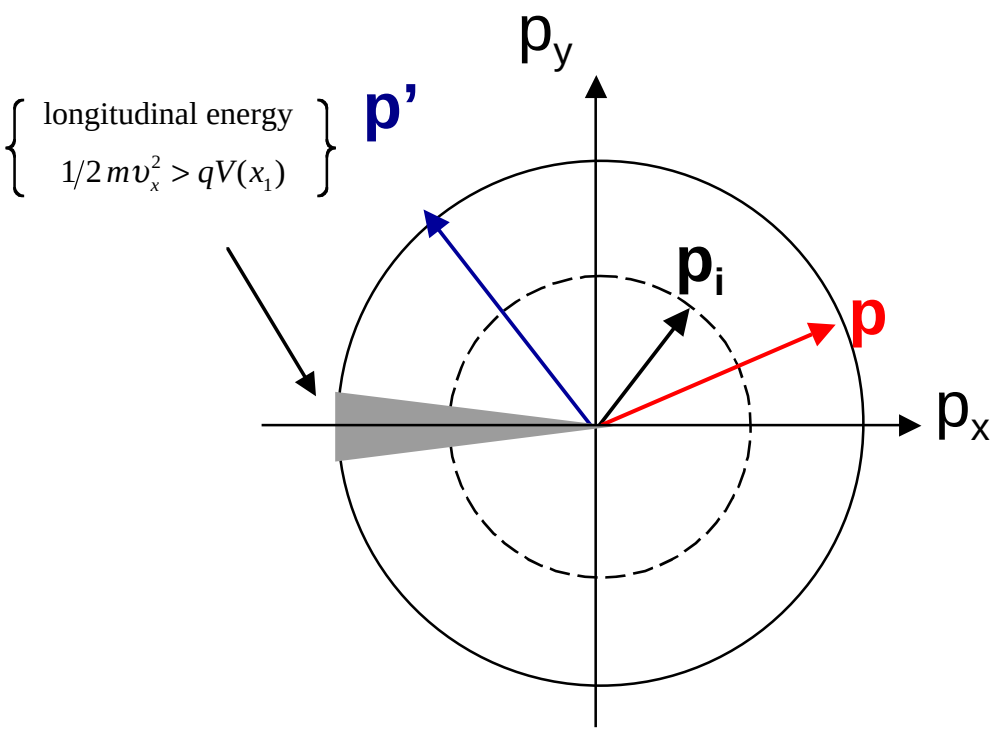

Fig. 4.6b Illustration of the scattering event of Fig. 4.6a in momentum space. Carriers that backscatter into the shaded region have sufficient longitudinal energy to surmount the barrier and return to the source.

Figure 4.7 is a plot of eq. (4.34) vs. position. It shows that when electrons gain more kinetic energy in the channel field [$qV(x_1)$] than the kinetic energy they were injected into the channel with, ($E_i - E_b$), then when they scatter they are unlikely to have sufficient kinetic energy associated with the x-directed velocity to surmount the barrier and return to the source. The answer to our question of how to define the critical length in eq. (4.31), therefore is: the critical length for channel backscattering is the distance over which the carrier gains an energy in the channel electric field that is equal to the kinetic energy that it was injected with. Nondegenerate carriers, are injected with a kinetic energy of k_BT, so the critical layer is the part of the channel over which the first k_BT/q of potential drop occurs, which is typically a small part of the channel. Above thresholds, carriers are degenerate and their kinetic energy exceeds k_BT (for $n_S = 10^{13}$ cm^{-2}, the Fermi level is located xxx meV above the bottom of the subband). Power supply voltages are still several times higher than the

injected kinetic energy, so the portion of the channel that contributes to backscattering to the source is a small fraction of the channel.

Fig. 4.7 Plot of eq. (4.34) for an injected kinetic energy of 100 meV showing that the transition between a high probability of backscattering to the source and a low probability occurs when the energy gained in the channel electric field equals the kinetic energy that the carrier was injected into the channel with.

4.6 Discussion

According to our scattering theory, the on-current of a MOSFET is controlled by the factor

$$\mathcal{B} = \frac{T}{2-T} = \frac{\lambda}{2\ell + \lambda}, \tag{4.35}$$

where we refer to $\mathcal{B}$ as the ballistic figure of merit for the device. This factor approaches one and the on-current approaches it's ballistic limit when

$$\lambda >> 2\ell. \tag{4.36}$$

Equation (4.36) explains why present-day devices can operate relatively close to the ballistic limit even though they are several mean-free-paths long. To deliver near-ballistic currents, the mean-free-path only has to be much longer than twice the critical length, which is a fraction of the channel length. Note also that the fact that a device delivers nearly the ballistic current does not imply that scattering is weak. Near the drain end of the channel, where the kinetic energy is high, the scattering rate increases and the mean-free-path is very short. We have seen, however, that carriers that scatter near the drain have little chance to return to the source.

Because scattering rates generally increase with energy, the question of what mean-free-path to use in eq. (4.35) arises – the very short mean-free-path near the drain or the relatively long mean-free-path near the source? We have seen that scattering in the low field region near the source is what controls the on-current, so the mean-free-path near the source is the appropriate one to use. Because carrier near the source have gained little kinetic energy from the channel electric field, the mean-free-path is close to its near-equilibrium value. Mobility is a concept that is well-defined in a bulk semiconductor under near equilibrium conditions; it has no clear meaning in a short high-field region where transport is very far from equilibrium [12]. But the near-equilibrium mobility is a bulk semiconductor is related to the near-equilibrium mean-free-path according to eq. (4.29). So the mobility measured in a long

channel MOSFET is a measure of the near-equilibrium mean-free-path, which is the quantity that controls the on-current of a nanoscale MOSFET for which mobility itself has no clear meaning. It is in this sense that we can say that increasing the mobility will increase the on-current of a nanoscale MOSFET [17]. These considerations explain the experimental correlation of on-current with mobility for short channel MOSFETs [18, 29].

We have seen that present day devices operate well below the ballistic limit because $T < 1$. It is of interest to ask how T scales as device dimensions shrink. According to eq. (4.31)

$$T = \frac{1}{1 + \lambda / \ell}. \tag{4.37}$$

As device dimensions shrink, the critical length, ℓ, will also shrink. At the same time, however, the increased doping and the stronger normal electric fields associated with the thinner gate oxide will reduce the mean-free-path, λ. The result is that we expect T to be roughly constant with scaling. This expectation is conformed by analyzing device data for the past 15 years. Devices are scaled in a way that maintains a constant transmission coefficient.

The question is whether it is possible to increase T by building devices differently then arises. Equation (4.37) shows that there are two possibilities. One is to raise the carrier mobility (mean-free-path), and a second possibility is to reduce the critical length, ℓ. The use of new materials (e.g. SiGe) or strain silicon shows promise for increasing mobility [20]. The use of fully depleted SOI MOSFET has the beneficial side effect in increasing carrier mobility. Reducing ℓ by increasing the electric field near the source would increase DIBL. Well-designed MOSFETs" keep DIBL low by keeping the electric field near the source low, which has the side effect of keeping ℓ long. Another possibility is to build in an electric field with lateral doping gradients or by the use if dual workfunction gates. Still another possibility is to make use of heterojunctions at the source end of the channel. These types of "source engineering" ideas could lead to higher transmission coefficients.

In Sec. 4.4, we showed that the scattering theory for the linear region could be written in the conventional form, except that L was replaced by $L + \lambda$. It turns out that the high V_{DS} current, eqn. (4.17), can also be expressed in a more conventional form. When $V_{DS} < V_{Dsat}$, the applied potential drops approximately linearly, so

$$\ell \approx \left(\frac{k_B T / q}{V_{DS}} \right) L . \tag{4.38}$$

If eq. (4.38) is used with eq. (4.35) in the drain current expression, eq. (4.17), we can express the drain current as

$$I_D = \frac{W C_{ox} \upsilon_T \, m\left(V_{GS} - V_T\right) V_{DS}}{1 + m V_{DS}}, \tag{4.39}$$

where

$$m = \frac{\mu}{2 L \upsilon_T} \tag{4.40}$$

and we have assumed nondegenerate conditions. Equation (4.35) is identical to a more conventional short channel model developed by Veeraraghavan and Fossum [21] except that the saturated velocity is replaced by the thermal velocity and a bulk charge term in [21] is missing in (4.39). Although the expressions are similar, the scattering approach and the conventional approach are based on much different physical pictures. Beginning from the velocity saturation view, we would be tempted to say that if velocity overshoot occurs, the on-current should be enhanced. In the scattering view, however, the limiting velocity is set by source injection. Velocity overshoot occurs within the channel where it can influence the self-consistent electrostatic potential. Since this affects the potential at the source, velocity overshoot influence the current indirectly.

For very long channel devices, V_{Dsat} approaches $V_{GS} - V_T$. If we substitute $V_{Dsat} = \left(V_{GS} - V_T\right)$ and assume that L is large, eq. (4.35) becomes

$$I_D = \mu C_{ox} \frac{W}{2L}\left(V_{GS} - V_T\right)^2, \tag{72}$$

which is the expected result for a long channel transistor. We conclude that the scattering model produces results much like the conventional model in both the long and short-channel limits [22].

4.7 Summary

Figure 4.8 summarizes the essential physical picture that we have developed for the nanoscale MOSFET. Near thermal equilibrium conditions are maintained at the source end of the channel so that $Q_n = C_{ox}\left(V_{GS} - V_T\right)$ (two-dimensional electrostatics lowers V_T from its value in an MOS capacitor or long channel MOSFET). Since carriers are in thermal equilibrium at the source, the maximum velocity at the source end of the channel is limited to the thermal velocity, $\tilde{\upsilon}_T$. Velocity overshoot cannot occur at the source, but $\tilde{\upsilon}_T$ can exceed 1.0×10^7 cm/s by as much as 50% when the carriers are degenerate.

In practice, the average carrier velocity at the source is much less than $\tilde{\upsilon}_T$ because of carrier backscattering. For steady-state current, the backscattering that matters is the portion that occurs within a critical distance, ℓ, from the beginning of the channel. The length of this critical region is roughly the distance over which the potential drops by *the* average kinetic energy of the injected carriers. Backscattering within the critical region is largely controlled by the low-field, near-equilibrium mobility of inversion layer carriers. Carriers that backscatter deeper within the channel are likely to emerge from the drain where they contribute to the DC drain current. (Scattering throughout the channel influences the device transit time, but the transit time is very short in a nanoscale transistor.) Deep within the channel, carrier transport is complex, and strong velocity overshoot occurs. Velocity overshoot lowers the carrier density near the drain, which influences the self-consistent

electric field, and therefore has an indirect effect on the drain current. (In a small device, this indirect effect may be strong.)

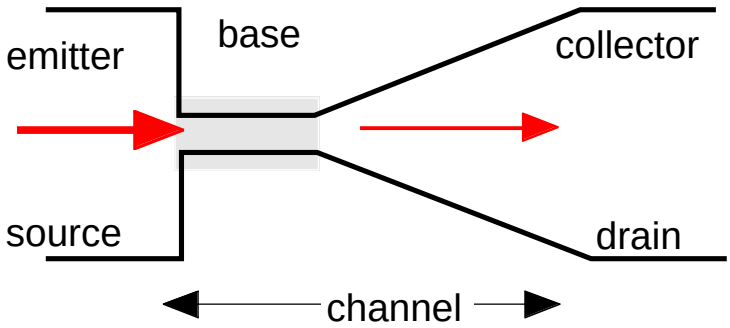

Fig. 4.8 Essential physical picture of steady-state carrier transport in the nanoscale MOSFET.

Chapter 4 References

[1] 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.

[2] A. Lochtefield and D. Antoniadis, "On experimental determination of carrier velocity in deeply scaled NMOS: How close to the thermal limit?, *IEEE Electron Dev. Lett.*, **22**, pp. 95-97, 2001.

[3] G. Timp, J. Bude, et al., Intern. Electron De. Meeting, Tech. Digest, pp. 55-58, 1999.

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

[7] Zhibin Ren, "Nanoscale MOSFETs: Physics, Simulation, and Design" Ph.D. thesis, Purdue University, Dec. 2001.

[8] Alexi Svizhenko and M.P. Anantram, "Influence of scattering on nanotransistor design," unpublished, 2002.

[9] Robert F. Pierret, *Advanced Semiconductor Fundamentals*, Addision-Wesley, Reading, Massachusetts, 1987.

[10] Anisur Rahman and Mark Lundstrom, "A Compact Model for the Nanoscale Double Gate MOSFET," to appear in *IEEE Trans. Electron Dev.*, March 2001.

[11] M.S. Lundstrom, "Elementary scattering theory of the MOSFET," *IEEE Electron Dev. Lett.*, **18**, pp. 361-363, 1997.

[12] Mark Lundstrom, *Fundamentals of Carrier Transport*, 2nd Ed., Cambridge University Press, Cambridge, UK, in press.

[13] E.H. Rhoderick, *Metal-Semiconductor Contacts*, Clarendon Press, Oxford, UK, 1978.

[14] J.P. McKelvey, R.L. Longini, and T.P. Brody, "Alternative approach to the solution of added carrier transport problems in semiconductors, " *Phys. Rev.*, **123**, pp. 51-57, 1961.

[15] W. Shockley, "Diffusion and drift of minority carrier in semiconductors for comparable capture and scattering mean free paths," *Phys. Rev.,* **125**, pp. 1570-1576, 1962.

[16] P.J. Price, "Monte Carlo calculation of electron transport in solids," *Semiconductors and Semimetals*, **14**, pp. 249-334, 1979.

[17] Mark Lundstrom, "On the Mobility Versus Drain Current Relation for a Nanoscale MOSFET," IEEE Electron Dev. Lett.,. **22**, No. 6, pp. 293-295, 2001.

[18] A. Lochtefeld and D.A. Antoniadis, "Investigating the relationship between electron mobility and velocity in deeply scaled NMOS via mechanical stress," IEEE Electron dev. Lett., **22**, pp. 591-593, 2001.

[19] R. Ohba and T. Mizuno, "Nonstationary electron/hole transport in Sub-0.1mm MOS devices: Correlation with mobility and low power CMOS application," *IEEE Trans. Electron Dev*., **48**, pp. 338-343, 2001.

[20] K. Rim, J.L. Hoyt, and J.F. Gibbons, "Fabrication and Analysis of Deep Submicron Strained-Si N-MOSFET's," *IEEE Trans. Electron Dev*., **47**, pp. 1406-1415, 2000.

[21] S. Veeraraghavan and J.G. Fossum, "A Physical Short-Channel Model for the Thin-Film SOI MOSFET Applicable to Device and Circuit CAD," *IEEE Trans. on Electron Dev*., Vol. 35, pp. 1866-1875, 1988.

[22] This connection was first pointed out to me in discussion with Professor J.G. Fossum', Lixin Ge, and Keunwoo Kim at the University of Florida, March, 1999.