

Simulation at the Start of the New Millenium: Crossing the Quantum-Classical Threshold

D. K. FERRY

*Department of Electrical Engineering and Center for Solid State Electronics Research,
Arizona State University, Tempe, AZ 85287-5706, USA*

It is clear that continued scaling of semiconductor devices will bring us to a regime with gate lengths less than 50 nm within another decade. The questions that must be addressed in simulation are difficult. Pushing to dimensional sizes such as this will probe the transition from classical to quantum transport, and *there is no present approach to this regime that has proved effective*. Contrary to the classical case in which electrons are negligibly small, the finite extent of the momentum space available to the electron set size limitations on the minimum wave packet – this is of the order of a few nanometers – and leads to the *effective potential*. The latter is an approach to find the equivalent classical potential, by which the actual potential is modified by quantum effects. The use of the effective potential for analyzing the effect of quantization on semiconductor devices will be discussed. The manner in which this leads to new formulations for quantum transport will be discussed.

Keywords: Effective potential; Quantum transport; Device simulation

1. INTRODUCTION

As the density of integrated circuits continues to increase, there is a need to shrink the dimensions of the devices of which they are comprised. Smaller circuit dimensions lead to more transistors on a single die. Advances in lithography have driven device dimensions to the deep-submicron range. Currently, 0.18 μm is the state-of-the-art process technology, but even smaller dimensions are expected in the near future. Groups from Toshiba and Lucent Bell Labs have fabricated *n*-channel MOSFETs with effective gate lengths below 25 nm, thus demonstrating that these feature sizes are feasible.

In this regime, the transport is expected to be dominated by quantum effects throughout the active region, even though quantum transport for these small (and inhomogeneous) devices is not well established within a consistent conceptual framework [1]. Nevertheless, several approaches to simulation of semiconductor devices have appeared in which the transport is handled quantum mechanically [2]. An additional problem arises in small structures, where one must begin to worry about the effective size of the carriers themselves [3]. In recent work, we have discussed the arguments for various sizes for electrons in semiconductor devices [4]. There are several reasons why this becomes important in small devices.

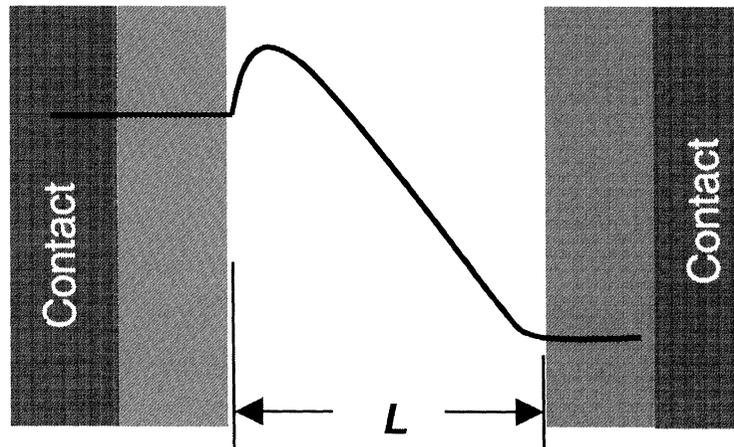


FIGURE 1 A conceptual device, with an applied source-drain bias. The source is at the left and the drain is at the right. The dashed areas are the “transition” regions, which must be considered as part of the active device.

Consider, for example, the conceptual device in Figure 1. The active channel length L is bounded by two contact transition regions, which in turn are coupled to the source and the drain contacts [5]. It is in these contact transition regions that the carriers lose their coherence completely so that the actual contacts can be considered as being in equilibrium. The drain region is the distance over which the energetic channel carriers equilibrate and can be as much as 10 nm [6]. On the other hand, the transition near the source is smaller, but represents an absolute minimum of the device resistance which cannot be reduced [7].

We may consider the importance of these transition regions from the size of important lengths and times within a Si MOSFET. These are listed in Table I. Nearly all of these lengths are of comparable size. This means that traditional transport based upon the Boltzmann equation is inappropriate for such devices. On the other hand,

TABLE I Important parameters

Parameter	Value in Si
Gate length	35–100 nm
Inelastic mfp	30–80 nm
Coherence length	20–50 nm
Coherence time	50–100 fs
Transit time	0.1–1 ps

approaches such as the ensemble Monte Carlo technique, in which no distribution function is assumed, provide a viable approach to this regime, *if suitable quantum effects can be incorporated*.

As mentioned above, the effective size of the electron wave packet is one such correction. We will see in the next section that this leads to an effective potential which characterizes the *initial* quantization effects in the channel. Moreover, the use of such an effective potential dramatically changes the way in which we utilize quantum transport equations, such as the Wigner equation of motion [8]. This will be discussed in the subsequent section, and limitations in quantum approaches that are used today will be discussed.

2. THE EFFECTIVE POTENTIAL

In classical mechanics, we typically consider the electron to be represented by a delta function in position. In quantum mechanics, however, the electron is represented by a wave packet, which has non-zero extent. In some sense, this size is a result of the uncertainty relation between position and momentum. One rational connection in this regard is the representation in terms of

Wannier functions. Wannier functions are highly localized functions which arise from summing the Bloch functions over a complete band. That is, there are N Wannier functions, whose extent is roughly of atomic dimensions – this becomes the equivalent of a delta function in space. The truly localized Wannier function, of atomic size, only arises if *all* possible Bloch states are in the summation. If the band is only partially full, however, the summation may be made over the occupied states, and the resulting Wannier function has an extended range. This approach was used, for example, to find the orbit of electrons bound to impurities. Here, we adapt this approach to determine the occupied states from the appropriate quantum distribution function, which then leads to a broadened Wannier function in real space. We assert that this broadened function represents the electron wave packet for this system [3]. We will see later that this is directly connected to the theory of distributions [9], which leads to a formal theory of the delta function [10].

In order to describe the packet in real space, we must account for the contributions to the wave packet from all occupied plane wave states. Even at room temperature, the carriers in the inversion layer in a Si MOSFET are a two-dimensional gas, and all states up to the Fermi energy are occupied. Nevertheless, we may take a non-degenerate approach for which the wave packet becomes a Gaussian. The spatial extent of this packet can be estimated as [3]

$$\delta r = 2\Delta r = \frac{2}{k_F} = \frac{\lambda_F}{\pi}. \quad (1)$$

For a Si inversion layer, we thus find that the effective size of the wave packet is about 3.6 nm at a density of $5 \times 10^{12} \text{ cm}^{-2}$.

The general Hamiltonian term for the potential energy in an inhomogeneous situation is given by [11]

$$H_V = \int d^3\mathbf{r} V(\mathbf{r}) n(\mathbf{r}). \quad (2)$$

This can be used to define the effective potential through

$$\begin{aligned} H_V &= \int d^3\mathbf{r} V(\mathbf{r}) \\ &\times \sum_i \int d^3\mathbf{r}' f(\mathbf{r} - \mathbf{r}') \delta(\mathbf{r}' - \mathbf{r}_i) \\ &= \sum_i \int d^3\mathbf{r}' \delta(\mathbf{r}' - \mathbf{r}_i) V_{\text{eff}}(\mathbf{r}') \\ &= \int d^3\mathbf{r}' V_{\text{eff}}(\mathbf{r}') n(\mathbf{r}'), \end{aligned} \quad (3)$$

where the density is now represented in a classical manner by delta functions. Here, the effective potential is defined by the wave packet itself, with

$$V_{\text{eff}}(\mathbf{r}) = \int d^3\mathbf{r}' V(\mathbf{r} + \mathbf{r}') f(\mathbf{r}'). \quad (4)$$

As mentioned, a particular form for the spread function is (in one dimension)

$$f(x) = \frac{1}{\sqrt{2\pi}a_0} \exp\left(-\frac{x^2}{2a_0^2}\right). \quad (5)$$

While the natural value to take for Si is the 3.6 nm found earlier, this is for “free” carriers in a two-dimensional gas. The quantization in the inversion layer is normal to the interface, and we have to worry about the bound states.

Feynman and Kleinert have used an effective potential of this form in an energy minimization procedure for bound carriers in a parabolic well. When the levels are not well defined (due to thermal broadening), then we may approximate the spread by

$$a_0^2 = \frac{\hbar^2}{8mk_B T}. \quad (6)$$

This gives a value of 0.64 nm in Si inversion layers.

We can estimate the validity of this approach by comparing the role played by the effective potential in a quantum well such as that at the interface in a MOSFET. The approach we take is to compare the effective potential derived from

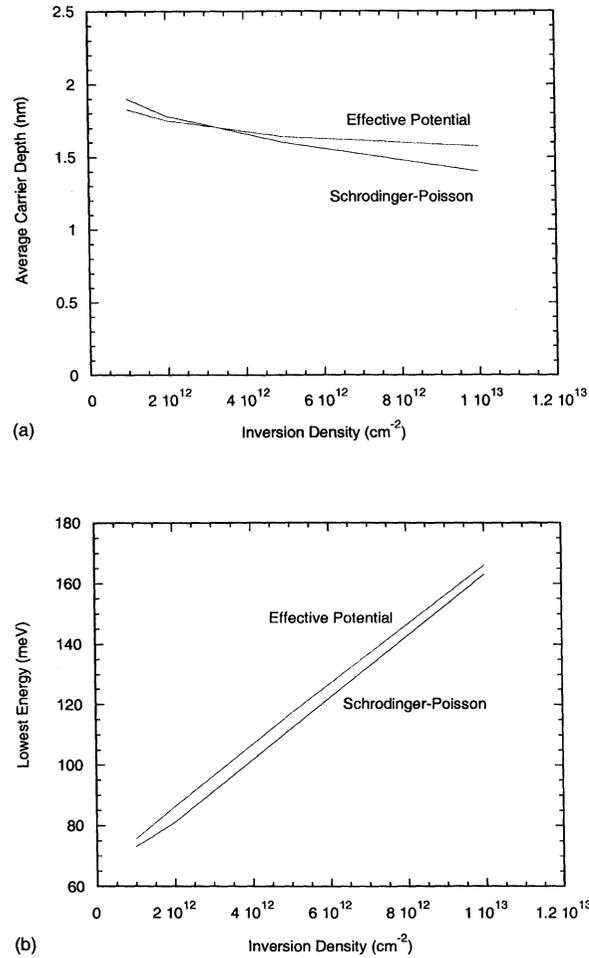


FIGURE 2 The average inversion depth (a) and quantization energy (b) for carriers in a Si MOSFET. A value of $a_0 = 0.5$ nm has been used.

a triangular quantum well with the fully self-consistent solution of a coupled Schrödinger-Poisson solver for the inversion layer. In Figure 2, we show the average depth of the inversion density and the quantized energy level of the well itself. It can be seen that the agreement is good.

3. APPLICATIONS IN QUANTUM TRANSPORT EQUATIONS

We turn at this point to the Wigner function, which is a Fourier transform, in the difference

coordinate, of the density matrix [2]. Generally, use of the Wigner function has been hampered by the fact that it has non-compact support – it takes on values at spatial regions where the wave function itself is undefined. This has led to a very complicated equation of motion, particularly in the connection of the spatially varying potential energy with the momentum of the Wigner function.

The most usual form of the Wigner equation of motion can easily be derived by techniques reviewed by Moyal many years ago [12]. In this derivation, a delta function is used to uncouple the potential from the Wigner function in the commutator driving terms. In the approach built around the finite extent of the wave packet, we use (5) as a generalized distribution with which to represent the delta function. Classically, one should take the limit as $a_0 \rightarrow 0$ in order to recover the delta function. In quantum statistical mechanics, however, we need not take the full limit. Rather, we need only limit it to the size of the fundamental wave packet. When this is done, we can write the field driving term in the Wigner equation of motion as

$$-\frac{\partial V_{\text{eff}}}{\partial x} \frac{\partial f_W(x, p, t)}{\partial p}. \quad (7)$$

This is essentially the classical term that appears in the Boltzmann equation with the exception that the actual potential is replaced by the effective potential (4). However, for strict validity of the approximation used, we also require that a_0 not be too small; in fact, we need

$$a_0 \bar{p} > \hbar. \quad (8)$$

Here, \bar{p} is the “uncertainty” of the momentum, such as the thermal spread of this value. In this sense, a_0 should be close to the thermal Debye wavelength. It is well known that the Wigner function is non-negative, but these negative regions vanish if the function is averaged over phase space regions of a size given by the uncertainty relation.

Here, (8) insures that the smoothed Wigner function we use is averaged over these regions so that the uncertainty has been incorporated. In some sense, this approach is comparable to a WKB approximation, although this has been asserted for the Wigner function previously.

Recognition of the wave packet size is fundamental in the collision term as well. Complications are created by the fact that the collision takes a finite time to complete, and quantum transitions are non-local. Indeed, it can be estimated that optical phonon collisions take 2–5 fs to complete [13]. While this is quite short, an electron traveling at 3×10^7 cm/s will cover 0.6–1 nm in this time. This distance is *not* negligible in a 35 nm MOSFET or in a resonant tunneling diode! Consider for example the conceptual problem of Figure 3, in which a particle approaches a potential barrier and *absorbs* a phonon near the barrier, as indicated in the figure.

Classically, the transition indicated in Figure 3 is forbidden. First, it is nonlocal in space. Secondly, absorbing a phonon leads to an *increase* in the momentum wave vector, but here the final

state has a *lower* value of momentum wave vector due to the rapid increase in the potential energy. The horizontal lines are meant to be the spatial extent of the initial and final wave packets (which are both moving). In quantum mechanics, the energy that is conserved is the total energy, not just the kinetic energy. In Figure 3, it is the intracollisional field effect [14] which rapidly changes the momenta of the two states with position (time). Quite beyond the estimates above, an electron approaching the top of a 0.3 eV barrier (in the cladding layer) is traveling at more than 10^8 cm/s, and will cover 3–5 nm (which may be the entire barrier thickness) during the collision. If we use the non-zero spatial resolution of the delta functions through (5), then the in-scattering term, for phonon emission, in the Wigner equation of motion can be written as

$$\left. \frac{\partial f_w}{\partial t} \right|_{\text{coll}} = \frac{2}{\hbar^2} \sum_q \frac{|H_{eL}|^2}{\delta} (N_q + 1) \times \int \frac{f_w(x + x', \xi, t) f(x') dx'}{g(\xi, x + x')^2 + \delta^2}, \quad (9)$$

where $f(x')$ is given by (5), and the energy resolution term

$$g(\xi, x + x') = E(\xi) + V(x + x') + \hbar\omega_0 - E(k) - V(x) \quad (10)$$

provides the energy conservation of the system. It should be pointed out that the Lorentzian form used here is an approximation. Moreover, this term cannot be treated in isolation. In fact, it is the product of the generalized distribution $f(x')$ and this Lorentzian that provides the probability function to find the final state, which is defined by both a position and a wave vector. The broadening term is the same as found previously [14], and is given, within the above approximation, by

$$\delta = \frac{\hbar q \tau_{\text{coll}}}{m^*} \frac{\partial V_{\text{eff}}}{\partial x}. \quad (11)$$

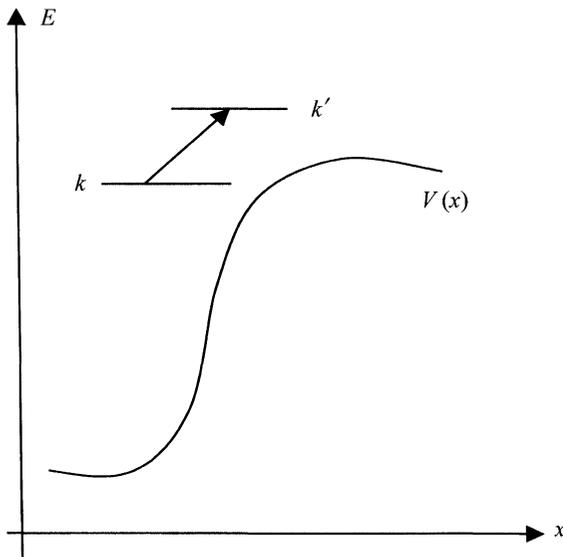


FIGURE 3 Conceptual picture of an electron absorbing a phonon near a potential barrier.

Space prohibits the full discussion (and derivation) of these terms, but this will be provided elsewhere [15].

4. CONNECTION WITH OTHER QUANTUM POTENTIALS

We can expand the effective potential when it is a slowly varying function of position. That is, we can take the effective potential from the defining lines in (4) and use a Taylor series expansion as

$$\begin{aligned} W(x) &= \frac{1}{\sqrt{2\pi\alpha}} \int_{-\infty}^{\infty} V(x+\xi) e^{-\xi^2/2\alpha^2} d\xi \\ &\cong \frac{1}{\sqrt{2\pi\alpha}} \int_{-\infty}^{\infty} \left[V(x) + \xi \frac{\partial V}{\partial x} \right. \\ &\quad \left. + \frac{\xi^2}{2} \frac{\partial^2 V}{\partial x^2} + \dots \right] e^{-\xi^2/2\alpha^2} d\xi. \end{aligned} \quad (12)$$

The first term allows us to bring the potential outside the integral, while the second term vanishes due to the symmetry of the Gaussian. The third term becomes the leading correction term, which gives us

$$W(x) = V(x) + \alpha^2 \frac{\partial^2 V}{\partial x^2} + \dots \quad (13)$$

The major purpose of the effective potential is to find a dependence of the density upon this potential as $\exp(-\beta V)$. Using this fact, we can replace the potential in the second term of (13) with the density, and

$$W(x) = V(x) - \frac{\alpha^2}{\beta} \frac{\partial^2 \ln(n/n_0)}{\partial x^2} + \dots \quad (14)$$

This particular form is often connected with a *quantum potential* appearing in hydrodynamic equations that are obtained from the equation of motion for the Wigner function [16] or the density matrix [17]. The second term has been termed the

quantum potential in this approach. We can now replace the density with the square root of the density, and if we ignore terms that are quadratic in the first derivatives and of higher order, (14) can be written as

$$\begin{aligned} W(x) &= V(x) - \frac{2\alpha^2}{\beta} \frac{\partial^2 \ln(\sqrt{n/n_0})}{\partial x^2} + \dots \\ &= V(x) - \frac{2\alpha^2}{\beta\sqrt{n}} \frac{\partial^2 \sqrt{n}}{\partial x^2} + \dots \end{aligned} \quad (15)$$

Within a factor of 2, the second term is now recognized as the density gradient term first discussed for quantum hydrodynamics by de Broglie [18] and Madelung [19], but known more as the Bohm potential [20]. It has become popular to use this correction in current device simulations to try to incorporate some quantum effects, referring to the corrections as the “density gradient potential”.

From the above, we can see that the various forms for the *quantum potential* are really approximations to the full *effective potential*. As a result, use of the latter is to be preferred, since the integral smoothing will reduce fluctuations while the derivative forms amplify fluctuations. Moreover, it is well known that the Bohm potential reproduces the quantization energy of the ground state, and this is found as well in the effective potential. This means that the effective potential is already of a nature to be used for mixed wave functions, whereas the density gradient approaches have severe problems in this case, particularly near nodal points of the composite wave function [1]. By judicious choice of the extent of the wave packet, excellent fit can be obtained for the quantum parameters in *e.g.*, the MOSFET inversion layer.

Acknowledgements

The author would like to express appreciation for helpful discussions with H. L. Grubin, M. Fischetti and D. Vasileska. This work was supported in part by the Office of Naval Research and the Semiconductor Research Corporation.

References

- [1] Ferry, D. K. and Barker, J. R. (1998). "Open Problems in Quantum Simulation in Ultra-Submicron Devices", *VLSI Design*, **8**, 165–72.
- [2] Ferry, D. K. and Grubin, H. L. (1995). "Modeling of Quantum Transport in Semiconductor Devices", In: *Solid State Physics*, **49** (Academic Press, San Diego), 283–448.
- [3] Ferry, D. K. and Grubin, H. L. (1998). "Electrons in Semiconductors: How Big are They"? *Proc. IWCE-6*, Osaka (IEEE Press, New York), pp. 84–7.
- [4] Ferry, D. K. (2000). "The onset of quantization in ultra-submicron semiconductor devices", *Superlatt. Microstruc.*, **27**, 61–6.
- [5] Ferry, D. K. and Barker, J. R. (1999). "Issues in general quantum transport with complex potentials", *Appl. Phys. Lett.*, **74**, 582–5.
- [6] Gross, W. J., Vasileska, D. and Ferry, D. K. (2000). "Ultra-Small MOSFETs: The Importance of the Full Coulomb Interaction on Device Characteristics", *VLSI Design, these proceedings*.
- [7] Datta, S., Assad, F. and Lundstrom, M. S. (1998). "The silicon MOSFET from a transmission viewpoint", *Superlatt. Microstruc.*, **23**, 771–80.
- [8] See, e.g., Grubin, H. L. (2000). "Wigner Function Methods in Modeling of Switching in Resonant Tunneling Diodes", *VLSI Design, these proceedings*.
- [9] See, e.g., Carrier, G. F., Krook, M. and Pearson, C. E. (1966). *Functions of a Complex Variable* (McGraw-Hill, New York), p. 318ff.
- [10] See, e.g., Erdélyi, A. (1961). "From Delta Functions to Distributions", In: *Modern Mathematics for the Engineer*, Edited by Beckenbach, E. F. and Hestenes, M. R. (McGraw-Hill, New York), pp. 5–50.
- [11] See, e.g., Kadanoff, L. P. and Baym, G. (1962). *Quantum Statistical Mechanics* (W. A. Benjamin, Reading, MA), Chapt. 6.
- [12] Moyal, J. E. (1949). "Quantum Mechanics as a Statistical Theory", *Proc. Camb. Phil. Soc.*, **45**, 99–124.
- [13] Bordone, P., Vasileska, D. and Ferry, D. K. (1996). "Collision-duration for optical-phonon emission in semiconductors", *Phys. Rev. B.*, **53**, 3846.
- [14] See, e.g., Barker, J. R. (1978). "High Field Collision Rates in Polar Semiconductors", *Sol.-State Electron.*, **21**, 267; Barker, J. R. (1980). "Quantum Transport Theory", In: *Physics of Nonlinear Transport in Semiconductors*, Edited by Ferry, D. K. et al. (Plenum Press, New York), pp. 126–152.
- [15] Ferry, D. K. and Grubin, H. L. (2001). "Quantum Transport Physics", *Int. J. High Speed Electron. Sys.*, in press.
- [16] Zhou, J.-R. and Ferry, D. K. (1992). "Simulation of Ultra-Small GaAs MESFETs Using Quantum Moment Equations", *IEEE Trans. Electron Dev.*, **39**, 473–8.
- [17] Ferry, D. K. and Zhou, J.-R. (1993). "On the Form of the Quantum Potential for Use in Hydrodynamic Equations for Semiconductor Device Modeling", *Phys. Rev. B*, **48**, 7944–9.
- [18] de Broglie, L. (1926). "Sur la possibilité de relier les phénomènes d'interférence et la diffraction à la théorie des quanta de lumière", *C.R. Acad. Sci. Paris*, **183**, 447; (1927), "La structure atomique de la matière et du rayonnement et la mécanique ondulatoire", *C.R. Acad. Sci. Paris*, **184**, 273.
- [19] Madelung, E. (1926). "Quantentheorie in Hydrodynamischer Form", *Z. Phys.*, **40**, 322.
- [20] Bohm, D. (1952). "A Suggested Interpretation of the Quantum Theory in Terms of 'Hidden' Variables", *Phys. Rev.*, **85**, 166; **85**, 180.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

