Wigner function approach to nano device simulation

Hans Kosina

TU Vienna, Institute for Microelectronics,
Gusshausstrasse 27–29, A-1040 Vienna, Austria
E-mail: kosina@iue.tuwien.ac.at

Abstract: Coherent transport in mesoscopic devices is well described by the Schrödinger equation supplemented by open boundary conditions. When electronic devices are operated at room temperature, however, a realistic transport model needs to include carrier scattering. In this work, the kinetic equation for the Wigner function is employed as a model for dissipative quantum transport. Carrier scattering is treated in an approximate manner through a Boltzmann collision operator. The development of Monte-Carlo algorithms for this quantum kinetic equation is complicated by the fact that, as opposed to the semi-classical case, the integral kernel is no longer positive. This so-called negative sign problem requires the introduction of new numerical techniques in order to obtain stable Monte-Carlo methods. A particular method for the solution of the stationary Wigner equation is presented. Applications to single barrier and double barrier structures are discussed.

Keywords: device simulation; quantum transport; Wigner function; Monte Carlo method; resonant tunnelling diode.

Reference to this paper should be made as follows: Kosina, H. (2006) 'Wigner function approach to nano device simulation', *Int. J. Computational Science and Engineering*, Vol. 2, Nos. 3/4, pp.100–118.

Biographical notes: Hans Kosina received the 'Diplomingenieur' degree in Electrical Engineering, a PhD in Technical Sciences and the 'venia docendi' in Microelectronics from the Technical University of Vienna in 1987, 1992 and 1998, respectively. Currently, he is an Associate Professor at the Institute for Microelectronics at the same University. His current research interests include modelling of semiconductor devices, nanoelectronic devices, organic semi-conductors and optoelectronic devices, development of novel Monte Carlo algorithms for classical and quantum transport problems and computer aided engineering in ULSI-technology.

1 Introduction

Modelling of electronic transport in mesoscopic systems requires a theory that describes open, quantum-statistical systems driven far from thermodynamic equilibrium. Several formulations of quantum transport have been employed practically, such as those based on the density matrix, non-equilibrium Green's functions and the Wigner function.

A quantum-mechanical phase-space distribution was introduced by Eugene Wigner in 1932. The purpose was the formulation of a quantum correction for the thermodynamic equilibrium of a many-body system by means of a quasiprobability function. In more recent times, the definition of the Wigner function has been generalised as a Fourier transform of a many-body Green's function (Mahan, 1983).

The Wigner function is a real-valued but not necessarily positive definite quasidistribution and represents a quantum generalisation of Boltzmann's N-particle distribution. The Wigner function formalism is attractive as it allows the expression of quantum dynamics in a phase-space formulation, directly comparable with the classical analogue. A phase-space approach may appear more intuitive compared with the more abstract density matrix and Green's function approaches. The method of quasidistributions has proved especially useful in providing reductions to classical physics and kinetic regimes under suitable conditions.

To discuss the physical interpretation of a quasidistribution, let us consider the simple case of a one-particle distribution. Starting with the classical case, the distribution $f_{\mathrm{cl}}(\mathbf{p}, \mathbf{r}, t)$ is proportional to the probability density of finding a particle of momentum $\mathbf{p}$ and position $\mathbf{r}$ in the phase-space volume $\mathrm{d}^3p\,\mathrm{d}^3r$. This is a purely classical interpretation, directly conflicting with the uncertainty principle. The quantum mechanical quasidistribution $f_{\mathrm{w}}(\mathbf{p}, \mathbf{r}, t)$, however, is not positive definite and has to be interpreted as a joint density of $\mathbf{p}$ and $\mathbf{r}$ (Tatarskii, 1983). Only the marginal distributions are positive definite, that is, integrating $f_{\mathrm{w}}(\mathbf{p}, \mathbf{r}, t)$ over momentum space gives the probability density in $\mathbf{r}$-space and vice versa.

An excellent review of quantum-mechanical phase-space distributions in scattering theory has been given by Carruthers and Zachariasen (1983). This work deals with potential scattering, the two-body problem and the

N-body problem. A coupled hierarchy for reduced distribution functions and its truncation to the Boltzmann-Vlasov equation is presented. Tatarskii (1983) concentrates on quantum-mechanical systems in a pure state and investigates the representation of quantum mechanics by phase-space distributions. He points out that not every function that solves the Wigner equation describes a pure state. Therefore, initial conditions for the Wigner equation have to be subjected to a supplementary restriction. Today, phase-space quantisation is considered to be a third autonomous and logically complete formulation of quantum mechanics beyond the conventional ones based on operators in Hilbert space or path integrals (Cutright and Zachos, 2001; Zachos, 2002). This formulation is free of operators and wave functions. Observables and matrix elements are computed through phase-space integrals of c-number functions weighted by a Wigner function.

Important quantum mechanical properties of electronic transport in semi-conductor structures are often those associated not with the degeneracy of the Fermi system but rather with quantum interference effects (Rammer, 1991). A wide variety of electronic quantum transport problems of interest are essentially one-particle in nature. In such cases, a full many-body description of the problem is not necessary and a description of electronic transport that makes use of the one-particle approximation can be used from the very outset. However, even when the electron–electron interaction effects are of interest, certain approximations do exist, allowing their description on a one-particle level (Rammer, 1991). Therefore, we shall consider in the following only electronic systems with one-particle degrees of freedom.

1.1 History and state of the art review

Reports on finite-difference solutions of the one-particle Wigner equation for device applications are due to Ravaioli et al. (1985), Kluksdahl et al. (1987) and date back to the mid 1980s. Frensley (1986a,b, 1987) was the first who introduced boundary conditions on the Wigner function to model open quantum systems. Later, self-consistency was added to the Wigner equation solvers (Frensley, 1989; Kluksdahl et al., 1989). Main and Haddad included a reduced Boltzmann scattering operator in transient Wigner function-based simulations (Mains and Haddad, 1988). Research on finite-difference solution methods for the Wigner equation culminated in 1990 when the review papers of Frensley (1990) and Buot and Jensen (1990) appeared.

The 1990s have seen further extensions and applications of the finite-difference Wigner function method. High-frequency operation of resonant tunnelling diodes has been studied by Jensen and Buot (1990, 1991), and the transient response by Gullapalli et al. (1994) and Biegel and Plummer (1997) and later by Woolard et al. (2002). A new finite-difference discretisation scheme has been proposed by Mains and Haddad (1994).

In 2002, implementations of Monte Carlo methods for solving the Wigner device equation were reported (Nedjalkov et al., 2002; Shifren and Ferry, 2002). Although with the finite-difference method, scattering was restricted to the relaxation time approximation and the momentum space to one dimension, the Monte Carlo method allows scattering processes to be included on a more detailed level, assuming a three-dimensional momentum-space (Kosina et al., 2003a,b; Shifren et al., 2003). Issues such as choosing proper up-winding schemes, restrictions on matrix size and momentum space resolution are largely relaxed or do not exist when using the Monte Carlo method. Construction of new Monte Carlo algorithms is complicated by the fact that the kernel of the integral equation to solve is not positive semidefinite. As a consequence, the commonly applied Markov chain Monte Carlo method shows a variance exponentially increasing with time, prohibiting its application to realistic structures or larger evolution times (Nedjalkov et al., 1996, 2002; Rossi et al., 1994). Because of this so-called negative sign problem, the concept of Wigner paths alone (Bordone et al., 2003; Pascoli et al., 1998) is not sufficient to construct a stable Monte Carlo algorithm. Instead, additional measures have to be introduced that prevent a runaway of the particle weights and hence of the variance (Kosina et al., 2003a,b; Shifren et al., 2003). Note that in Shifren et al. (2003), the statistical weights are termed affinities.

Large basic research efforts on the Monte Carlo modelling of electron–phonon interaction based on the Wigner function formalism have been reported Rossi et al. (1994) Bordone et al. (2003) Bertoni et al. (1999), Bordone et al. (1999) and Jacoboni et al. (2001).

The effect of a spatially varying effective mass in Wigner device simulations has been demonstrated by Tsuchiya et al. (1991) and Shih et al. (1994). A non-parabolic version of the Wigner equation has been derived by Bufler and Schlösser (1994). Multiband models have been reported by Miller and Neikirk (1991), Demaio et al. (2002) and Unlu et al. (2004).

A Wigner equation including a magnetic field has been solved by Wu and Wu (1992). The gauge-invariant formulation of the Wigner equation has been given by Levinson (1970) and a discussion can be found in various works (Carruthers and Zachariasen, 1983; Ferry and Goodnick, 2001; Holland and Kypriandis, 1986; Levanda and Fleurov, 2001; Sonego, 1991). Two-time and frequency-dependent Wigner functions are considered by Mahan (1983), Levanda and Fleurov (2001), Hänsch (1991) and Brunetti et al. (2000).

In various types of field-effect transistors carrier transport takes place in inversion layers, quantum wells or thin semi-conductor films. To model such devices, the effect of quantum confinement needs to be included in the transport model. An expansion of the three-dimensional wave function in the basis set of transverse wave functions is employed in the mode space approach (Venugopal et al., 2002). The problem is significantly simplified if coupling of the modes or subbands, can be neglected. In this case a set of decoupled Wigner equations, one for each subband, can be used (Croitoru et al., 2003, 2004). If the cross section of the semiconductor film changes rapidly, intermode coupling becomes important (Venugopal et al., 2004). Inclusion of intermode coupling in the Wigner function formalism for confined systems is an open issue.

Finally, we note that the Wigner function formalism is often used to derive reduced transport models, such as

the quantum hydrodynamic model (Degond and Ringhofer, 2003; Gardner, 1994; Gardner and Ringhofer, 1996; Zhou and Ferry, 1992) or to find quantum corrections to classical models, such as the ensemble Monte Carlo method (Tsuchiya and Ravaioli, 2001) or the spherical harmonics expansion method (Goldsman et al., 2000; Han et al., 2000).

2 The Wigner function formalism

In the Schrödinger picture, a physical system is quantum-mechanically described by a state vector $|\Psi(t)\rangle$ as function of time t. Often, the precise quantum-mechanical state of a system is not known, but rather some statistical information about the probabilities for the system being in one of a set of states. Suppose that there is a set of orthonormal states $\{|\Psi_1\rangle, |\Psi_2\rangle, \ldots\}$ and that the probabilities that the system is in one of these states are $\{p_1, p_2, \ldots\}$. Then, the expectation value of operator $\hat{\mathrm{A}}$ associated with the observable A is given by

$$\langle \mathrm{A} \rangle = \sum_i p_i \ \langle \Psi_i | \, \hat{\mathrm{A}} \, | \Psi_i \rangle \tag{1}$$

which is a quantum and statistical average. Introducing the density operator $\hat{\rho}$ as

$$\hat{\rho} = \sum_i p_i \ |\Psi_i\rangle \langle \Psi_i| \tag{2}$$

the expectation value becomes

$$\langle \mathrm{A} \rangle = \mathrm{Tr}(\hat{\rho}\hat{\mathrm{A}}) = \mathrm{Tr}(\hat{\mathrm{A}}\hat{\rho}) \tag{3}$$

Equations (1) and (3) require the operator $\hat{\mathrm{A}}$ to be self-adjoint. Equation (3) can be easily verified by expressing the trace of some operator $\hat{\mathrm{X}}$ in the basis $\{|\Psi_i\rangle\}$.

$$\mathrm{Tr}(\hat{\mathrm{X}}) = \sum_i \langle \Psi_i | \, \hat{\mathrm{X}} \, | \Psi_i \rangle \tag{4}$$

The fact that the probabilities sum up to unity, $\sum_i p_i = 1$, is expressed by the fact that the trace of the density operator is also unity, $\mathrm{Tr}(\hat{\rho}) = 1$. If the system is in a pure state $|\Psi_i\rangle$ it holds $p_i = 1$ and $p_j = 0 \ \forall \ j \neq i$ and the density operator is idem-potent, $\hat{\rho}^2 = \hat{\rho}$. Otherwise, the system is in a mixed state and $\hat{\rho}$ does not obey the idem-potency condition. From the Schrödinger equation for the state vector and the definition of $\hat{\rho}$, we immediately obtain the Liouville-von Neumann equation for the evolution of the density operator.

$$i\hbar \frac{\partial \hat{\rho}}{\partial t} = [\hat{\mathrm{H}}, \hat{\rho}] \tag{5}$$

Introducing the one-particle approximation (Rammer, 1991) implies that the electron system is modelled as consisting of many, non-interacting electrons. In the next step, one chooses the coordinate representation, where the set of basis vectors is given by the electron position eigenstates $|\mathbf{r}\rangle$. The eigenstates of the system are then represented by the wavefunctions $\Psi_i(\mathbf{r}, t) = \langle \mathbf{r} | \Psi_{\mathrm{i}}(\mathrm{t}) \rangle$ and the density operator by the density matrix $\rho(\mathbf{r}_1, \mathbf{r}_2, t)$.

$$\rho(\mathbf{r}_1, \mathbf{r}_2, t) = \langle \mathbf{r}_1 | \, \hat{\rho}(t) \, | \mathbf{r}_2 \rangle = \sum_i p_i \ \Psi_i(\mathbf{r}_1, t) \ \Psi_i^*(\mathbf{r}_2, t) \tag{6}$$

The Liouville-von Neumann equation in coordinate representation is found as

$$\frac{\partial \rho(\mathbf{r}_1, \mathbf{r}_2, t)}{\partial t} = (H_{\mathbf{r}_1} - H_{\mathbf{r}_2}) \rho(\mathbf{r}_1, \mathbf{r}_2, t) \tag{7}$$

2.1 The Wigner function

The Wigner function is obtained from the density matrix by means of the Wigner–Weyl transformation. This transformation consists of a change of independent coordinates to diagonal and cross-diagonal coordinates

$$\mathbf{r} = \frac{1}{2}(\mathbf{r}_1 + \mathbf{r}_2), \qquad \mathbf{s} = \mathbf{r}_1 - \mathbf{r}_2 \tag{8}$$

followed by a Fourier transformation with respect to $\mathbf{s}$ (Frensley, 1990). The variables $\mathbf{r}_1$ and $\mathbf{r}_2$ may be expressed in terms of the new ones.

$$\mathbf{r}_1 = \mathbf{r} + \frac{\mathbf{s}}{2}, \qquad \mathbf{r}_2 = \mathbf{r} - \frac{\mathbf{s}}{2} \tag{9}$$

Then, the elementary definition of the Wigner distribution is given by the following transformation of the density matrix.

$$f_{\mathrm{w}}(\mathbf{k}, \mathbf{r}, t) = \int \rho \left(\mathbf{r} + \frac{\mathbf{s}}{2}, \mathbf{r} - \frac{\mathbf{s}}{2}, t \right) \ \mathrm{e}^{-i\mathbf{k}\mathbf{s}} \, \mathrm{d}\mathbf{s} \tag{10}$$

The Wigner function (10) is real-valued, but not positive semidefinite. In terms of the wave functions, the definition (10) becomes

$$f_{\mathrm{w}}(\mathbf{k}, \mathbf{r}, t) = \sum_i p_i \int \Psi_i \left(\mathbf{r} + \frac{\mathbf{s}}{2}, t \right) \Psi_i^* \left(\mathbf{r} - \frac{\mathbf{s}}{2}, t \right) \ \mathrm{e}^{-i\mathbf{k}\mathbf{s}} \, \mathrm{d}\mathbf{s} \tag{11}$$

The normalisation of the Wigner function results from the normalisation of the wave functions.

$$\frac{1}{(2\pi)^3} \int \mathrm{d}\mathbf{r} \int \mathrm{d}\mathbf{k} \ f_{\mathrm{w}}(\mathbf{k}, \mathbf{r}, t) = 1 \tag{12}$$

Here, the $\mathbf{k}$-integration can be performed first, giving $\int \mathrm{e}^{-i\mathbf{k}\mathbf{s}} \mathrm{d}\mathbf{k} = (2\pi)^3 \delta(\mathbf{s})$. The normalisation (12) ensures that the quantity $N f_{\mathrm{w}}$, where N is the number of electrons in the system, will approach the classical distribution function f_{cl} in the classical limit (Jacoboni et al., 2001).

Sometimes it is convenient to use the inverse Fourier transform of (10).

$$\rho \left(\mathbf{r} + \frac{\mathbf{s}}{2}, \mathbf{r} - \frac{\mathbf{s}}{2}, t \right) = \frac{1}{(2\pi)^3} \int f_{\mathrm{w}}(\mathbf{k}, \mathbf{r}, t) \ \mathrm{e}^{i\mathbf{k}\mathbf{s}} \, \mathrm{d}\mathbf{k} \tag{13}$$

Changing variables gives a transformation that inverts the Wigner-Weyl transformation.

$$\rho(\mathbf{r}_1, \mathbf{r}_2, t) = \frac{1}{(2\pi)^3} \int f_{\mathrm{w}} \left(\mathbf{k}, \frac{\mathbf{r}_1 + \mathbf{r}_2}{2}, t \right) \mathrm{e}^{i\mathbf{k}(\mathbf{r}_1 - \mathbf{r}_2)} \, \mathrm{d}\mathbf{k} \tag{14}$$

An important feature of the phase-space approach is the possibility of expressing quantum-mechanical expectation

values in the same way as it is done in classical statistical mechanics, employing integration over the phase-space. The expectation values of operators of the form $A(\hat{\mathbf{r}})$ and $B(\hat{\mathbf{k}})$, where $\hat{\mathbf{k}} = \hat{\mathbf{p}}/\hbar$, are given as follows.

$$\begin{aligned}\langle \mathrm{A}(\hat{\mathbf{r}})\rangle &= \frac{1}{(2\pi)^3}\int f_{\mathrm{w}}(\mathbf{k},\mathbf{r},t)A(\mathbf{r})\,\mathrm{d}\mathbf{k}\,\mathrm{d}\mathbf{r}\\ &= \sum_i p_i \int A(\hat{\mathbf{r}})\,|\Psi_i(\mathbf{r},t)|^2\,\mathrm{d}\mathbf{r}\end{aligned} \tag{15}$$

$$\begin{aligned}\langle \mathrm{B}(\hat{\mathbf{k}})\rangle &= \frac{1}{(2\pi)^3}\int f_{\mathrm{w}}(\mathbf{k},\mathbf{r},t)B(\mathbf{k})\,\mathrm{d}\mathbf{k}\,\mathrm{d}\mathbf{r}\\ &= \sum_i p_i \int B(\mathbf{k})\,|\Phi_i(\mathbf{k},t)|^2\,\mathrm{d}\mathbf{k}\end{aligned} \tag{16}$$

If the classical observable $C(\mathbf{k},\mathbf{r})$ is a function of both momentum and position, the definition of a corresponding Hermitian operator $\hat{\mathrm{C}}$ is not unique. In this case, the Weyl quantisation can be applied. Thereby, the function C is expressed through its Fourier transform c.

$$C(\mathbf{k},\mathbf{r}) = \int c(\mathbf{a},\mathbf{b})\,\mathrm{e}^{i(\mathbf{k}a+\mathbf{r}b)}\,\mathrm{d}\mathbf{a}\,\mathrm{d}\mathbf{b} \tag{17}$$

The operator $\hat{\mathrm{C}}$ is defined by the following rule of correspondence.

$$\hat{\mathrm{C}} = \int c(\mathbf{a},\mathbf{b})\,\mathrm{e}^{i(\hat{\mathbf{k}}a+\hat{\mathbf{r}}b)}\,\mathrm{d}\mathbf{a}\,\mathrm{d}\mathbf{b} \tag{18}$$

Then, the expectation value of $\hat{\mathrm{C}}$ is given by the phase-space integral.

$$\mathrm{Tr}(\hat{\mathrm{C}}\hat{\rho}) = \int C(\mathbf{k},\mathbf{r})\,f_{\mathrm{w}}(\mathbf{k},\mathbf{r},t)\,\mathrm{d}\mathbf{k}\,\mathrm{d}\mathbf{r} \tag{19}$$

To proceed with (18), one may employ the Baker-Campbell-Hausdorff formula,

$$\mathrm{e}^{\hat{\mathrm{A}}+\hat{\mathrm{B}}} = \mathrm{e}^{\hat{\mathrm{A}}}\,\mathrm{e}^{\hat{\mathrm{B}}}\,\mathrm{e}^{-[\hat{\mathrm{A}},\hat{\mathrm{B}}]/2} \tag{20}$$

which is generally valid when $[\hat{\mathrm{A}},[\hat{\mathrm{A}},\hat{\mathrm{B}}]] = [\hat{\mathrm{B}},[\hat{\mathrm{A}},\hat{\mathrm{B}}]] = 0$, or in particular when $[\hat{\mathrm{A}},\hat{\mathrm{B}}]$ is a c-number.

2.2 *Marginal distributions*

The Wigner function (10) can assume negative values. Only the marginal distributions of $f_{\mathrm{w}}(\mathbf{k},\mathbf{r},t)$ are positive semi-definite and have the meaning of probability distributions in real space and momentum space, respectively.

$$n(\mathbf{r}) = \frac{1}{(2\pi)^3}\int f_{\mathrm{w}}(\mathbf{k},\mathbf{r},t)\,\mathrm{d}\mathbf{k} = \sum_i p_i\,|\Psi_i(\mathbf{r},t)|^2 \tag{21}$$

$$p(\mathbf{k}) = \frac{1}{(2\pi)^3}\int f_{\mathrm{w}}(\mathbf{k},\mathbf{r},t)\,\mathrm{d}\mathbf{r} = \sum_i p_i\,|\Phi_i(\mathbf{k},t)|^2 \tag{22}$$

Here, $\Phi_i(\mathbf{k},t)$ denotes the momentum representation of the state vector $|\Psi_i\rangle$. The integration in (22) can easily be carried out after changing variables, using (8).

$$\begin{aligned}&\int \mathrm{d}\mathbf{r}\int \mathrm{d}\mathbf{s}\;\Psi_i\left(\mathbf{r}+\frac{\mathbf{s}}{2},t\right)\;\Psi_i^*\left(\mathbf{r}-\frac{\mathbf{s}}{2},t\right)\;\mathrm{e}^{-i\mathbf{k}\mathbf{s}}\\ &\quad= \int \mathrm{d}\mathbf{r}_1\int \mathrm{d}\mathbf{r}_2\Psi_i(\mathbf{r}_1,t)\Psi_i^*(\mathbf{r}_2,t)\mathrm{e}^{-i\mathbf{k}(\mathbf{r}_1-\mathbf{r}_2)}\\ &\quad= (2\pi)^3|\Phi_i(\mathbf{k},t)|^2\end{aligned} \tag{23}$$

The marginal distributions (21) and (22) can also be expressed as the diagonal elements of the density matrix.

$$\frac{1}{(2\pi)^3}\int f_{\mathrm{w}}(\mathbf{k},\mathbf{r},t)\mathrm{d}\mathbf{k} = \langle\mathbf{r}|\,\hat{\rho}\,|\mathbf{r}\rangle = \rho(\mathbf{r},\mathbf{r}) \tag{24}$$

$$\frac{1}{(2\pi)^3}\int f_{\mathrm{w}}(\mathbf{k},\mathbf{r},t)\mathrm{d}\mathbf{r} = \langle\mathbf{k}|\,\hat{\rho}\,|\mathbf{k}\rangle = \sigma(\mathbf{k},\mathbf{k}) \tag{25}$$

Here, $|\mathbf{k}\rangle$ denotes the electron momentum eigenstate with eigenvalue $\hbar\mathbf{k}$ and σ the density matrix in momentum representation. Note that the latter can be used for a dual definition of the Wigner function (Rossi et al., 1994; Fannjiang et al., 2002).

$$f_{\mathrm{w}}(\mathbf{k},\mathbf{r},t) = \int \sigma\left(\mathbf{k}+\frac{\mathbf{l}}{2},\mathbf{k}-\frac{\mathbf{l}}{2},t\right)\,\mathrm{e}^{ir1}\,\mathrm{d}\mathbf{l} \tag{26}$$

This definition follows, for example, from (11), when the Ψ_i are replaced by

$$\Psi_i(\mathbf{r},t) = (2\pi)^{-3/2}\int \Phi_i(\mathbf{k}',t)\mathrm{e}^{i\mathbf{k}'\,r}\,\mathrm{d}\mathbf{k}' \tag{27}$$

Other marginal distributions than the elementary ones, (21) and (22), have to be constructed with care. Only Hermitian operators give real marginal distributions. For the current density, this operator would be $(\hat{\mathbf{k}}\hat{\rho}+\hat{\rho}\hat{\mathbf{k}})/2$. Expressing $\hat{\rho}$ in terms of the wave functions, we get the elementary current definition from wave mechanics.

$$\begin{aligned}\mathbf{j}(\mathbf{r}) &= \frac{\hbar}{2m^*}\,\langle\mathbf{r}|\,\hat{\mathbf{k}}\hat{\rho}+\hat{\rho}\hat{\mathbf{k}}\,|\mathbf{r}\rangle\\ &= \frac{\hbar}{2m^*}\sum_i p_i\big(\langle\mathbf{r}|\,\hat{\mathbf{k}}\,|\Psi_i\rangle\,\langle\Psi_i|\mathbf{r}\rangle+\langle\mathbf{r}|\Psi_i\rangle\,\langle\Psi_i|\,\hat{\mathbf{k}}\,|\mathbf{r}\rangle\big)\\ &= \frac{\hbar}{2im^*}\sum_i p_i\big[\Psi_i^*(\mathbf{r})\nabla\Psi_i(\mathbf{r})-\Psi_i(\mathbf{r})\nabla\Psi_i^*(\mathbf{r})\big]\end{aligned} \tag{28}$$

Choosing the momentum representation of $\hat{\rho}$,

$$\begin{aligned}\mathbf{j}(\mathbf{r}) = \frac{\hbar}{2m^*}\int \mathrm{d}\mathbf{k}_1\int \mathrm{d}\mathbf{k}_2\,\big(&\langle\mathbf{r}|\,\hat{\mathbf{k}}\,|\mathbf{k}_1\rangle\,\langle\mathbf{k}_1|\,\hat{\rho}\,|\mathbf{k}_2\rangle\,\langle\mathbf{k}_2|\mathbf{r}\rangle\\ &+\langle\mathbf{r}|\mathbf{k}_1\rangle\,\langle\mathbf{k}_1|\,\hat{\rho}\,|\mathbf{k}_2\rangle\,\langle\mathbf{k}_2|\,\hat{\mathbf{k}}\,|\mathbf{r}\rangle\big)\end{aligned}$$

we get the current density expressed in terms of the Wigner function.

$$\begin{aligned}\mathbf{j}(\mathbf{r}) &= \frac{\hbar}{2m^*}\int \mathrm{d}\mathbf{k}_1\int \mathrm{d}\mathbf{k}_2\,\sigma(\mathbf{k}_1,\mathbf{k}_2)\,(\mathbf{k}_1+\mathbf{k}_2)\,\mathrm{e}^{i(\mathbf{k}_1-\mathbf{k}_2)\mathbf{r}}\\ &= \frac{1}{(2\pi)^3}\int \frac{\hbar}{m^*}\mathbf{k}\,f_{\mathrm{w}}(\mathbf{k},\mathbf{r},t)\,\mathrm{d}\mathbf{k}\end{aligned} \tag{29}$$

Here, the Wigner function has been introduced using (26). The current density is given by the first-order moment of the Wigner function, in full analogy with the classical phase-space definition.

For the definition of the energy density we discuss several options. Starting from the trace operation for the statistical average, one would consider the symmetrised operator $(\hat{\mathbf{k}}^2\hat{\rho} + \hat{\rho}\hat{\mathbf{k}}^2)/2$ and derive the marginal distribution.

$$\begin{aligned} w_1(\mathbf{r}) &= \frac{\hbar^2}{4m^*}\left(\langle\mathbf{r}|\,\hat{\mathbf{k}}^2\hat{\rho}\,|\mathbf{r}\rangle + \langle\mathbf{r}|\,\hat{\rho}\hat{\mathbf{k}}^2\,|\mathbf{r}\rangle\right) \\ &= -\frac{\hbar^2}{4m^*}\sum_i p_i\left[\Psi_i^*(\mathbf{r})\nabla^2\Psi_i(\mathbf{r}) + \Psi_i(\mathbf{r})\nabla^2\Psi_i^*(\mathbf{r})\right] \\ &= \sum_i p_i[E_i - V(\mathbf{r})]|\Psi_i(\mathbf{r})|^2 \end{aligned} \tag{30}$$

The last expression in (30) is obtained with the help of the stationary Schrödinger equation. Apparently, w_1 describes the kinetic energy density, as the potential energy term $V(\mathbf{r})n(\mathbf{r})$ is subtracted from the total energy term. This energy density can become negative in tunnelling regions, where for one or more states $E_i < V(\mathbf{r})$ holds. In a derivation similar to (29), one finds the Wigner representation of w_1.

$$\begin{aligned} w_1(\mathbf{r}) &= \frac{\hbar^2}{4m^*}\int \mathrm{d}\mathbf{k}_1 \int \mathrm{d}\mathbf{k}_2\, \sigma(\mathbf{k}_1,\mathbf{k}_2)\,(\mathbf{k}_1^2+\mathbf{k}_2^2)\,\mathrm{e}^{i(\mathbf{k}_1-\mathbf{k}_2)\mathbf{r}} \\ &= \frac{1}{(2\pi)^3}\int \frac{\hbar^2}{2m^*}\left(|\mathbf{k}|^2 - \frac{1}{4}\nabla_{\mathbf{r}}^2\right) f_{\mathrm{w}}(\mathbf{k},\mathbf{r},t)\,\mathrm{d}\mathbf{k} \end{aligned} \tag{31}$$

To ensure positiveness of the energy density, in (Kosik, 2004) the Hermitian operator $\hat{\mathbf{k}}\hat{\rho}\hat{\mathbf{k}}$ is considered. Its marginal distribution can be shown to be positive semi-definite.

$$w_2(\mathbf{r}) = \frac{\hbar^2}{2m^*}\,\langle\mathbf{r}|\,\hat{\mathbf{k}}\hat{\rho}\hat{\mathbf{k}}\,|\mathbf{r}\rangle = \frac{\hbar^2}{2m^*}\sum_i p_i|\,\langle\mathbf{r}|\,\hat{\mathbf{k}}\,|\Psi_i\rangle\,|^2 \tag{32}$$

$$= \frac{\hbar^2}{2m^*}\sum_i p_i|\nabla\Psi_i(\mathbf{r})|^2 \geq 0 \tag{33}$$

The Wigner representation of w_2 is obtained as

$$\begin{aligned} w_2(\mathbf{r}) &= \frac{\hbar^2}{4m^*}\int \mathrm{d}\mathbf{k}_1 \int \mathrm{d}\mathbf{k}_2\, \sigma(\mathbf{k}_1,\mathbf{k}_2)\,(\mathbf{k}_1^2-\mathbf{k}_2^2)\,\mathrm{e}^{i(\mathbf{k}_1-\mathbf{k}_2)\mathbf{r}} \\ &= \frac{1}{(2\pi)^3}\int \frac{\hbar^2}{2m^*}\left(|\mathbf{k}|^2 + \frac{1}{4}\nabla_{\mathbf{r}}^2\right) f_{\mathrm{w}}(\mathbf{k},\mathbf{r},t)\,\mathrm{d}\mathbf{k} \end{aligned} \tag{34}$$

Conditions for obtaining non-negative marginal distributions are theoretically discussed by Wlodarz (1999). The Weyl correspondence (18) gives the definition of the energy density as the second-order moment of the Wigner function.

$$w_3(\mathbf{r}) = \frac{1}{(2\pi)^3}\int \frac{\hbar^2}{2m^*}|\mathbf{k}|^2 f_{\mathrm{w}}(\mathbf{k},\mathbf{r},t)\,\mathrm{d}\mathbf{k} \tag{35}$$

It can be seen that (35) is just the arithmetic mean of (31) and (34), $w_3 = (w_1 + w_2)/2$. Therefore, (35) represents the marginal distribution of the symmetrised operator $(\hat{\mathbf{k}}^2\hat{\rho} + 2\hat{\mathbf{k}}\hat{\rho}\hat{\mathbf{k}} + \hat{\rho}\hat{\mathbf{k}}^2)/4$.

All three definitions of the energy density give the same statistical average $\langle\epsilon\rangle = \mathrm{Tr}\big(\epsilon(\hat{\mathbf{k}})\hat{\rho}\big)$. The differences among the definitions are in the ∇^2 term, which vanishes after the $\mathbf{r}$-integration. However, only the density w_1 seems to have a clear physical interpretation as the kinetic energy density.

2.3 The Wigner equation

In this section, we consider a system consisting of one electron interacting with a potential distribution $V_{\mathrm{tot}}(\mathbf{r})$. This potential is assumed to be a superposition of some potential $V(\mathbf{r})$ and a uniform electric field: $V_{\mathrm{tot}}(\mathbf{r}) = V(\mathbf{r}) - \hbar\mathbf{F}\,\mathbf{r}$, with $\hbar\mathbf{F} = -e\mathbf{E}$. Although the existence of a field term is not physically motivated at this point, it is introduced here to demonstrate its treatment in the Wigner function formalism. The potential $V(\mathbf{r})$ comprises the electrostatic potential and the band-edge profile of the semi-conductor. A uniform effective mass m^* is assumed. In the usual coordinate representation, the Hamiltonian of the system is then given by

$$H = H_0 + V(\mathbf{r}) - \hbar\mathbf{F}\mathbf{r} \tag{36}$$

with

$$H_0 = -\frac{\hbar^2}{2m^*}\nabla_{\mathbf{r}}^2 \tag{37}$$

Electron–phonon interaction is neglected here. The evolution equation for the Wigner function is found by taking the time derivative of the defining Equation (10) and substituting the Liouville-von Neumann equation (7) on the right-hand side.

$$\begin{aligned} &\frac{\partial}{\partial t}f_{\mathrm{w}}(\mathbf{k},\mathbf{r},t) \\ &= \frac{1}{i\hbar}\int (H_{\mathbf{r}_1} - H_{\mathbf{r}_2})\rho\left(\mathbf{r}+\frac{\mathbf{s}}{2},\mathbf{r}-\frac{\mathbf{s}}{2},t\right)\mathrm{e}^{-i\mathbf{k}\mathbf{s}}\,\mathrm{d}\mathbf{s} \end{aligned}$$

In the following, the three parts of the Hamiltonian (36) will be separately transformed. Unlike in Section 2.2, where calculations were done in momentum representation, we choose below the configuration representation to carry out the transformations (Bertoni et al., 1999).

The free-electron Hamiltonian is given by H_0. To calculate the Wigner transform of H_0, we have to transform the gradients first. Differentiating the density matrix with respect to the new variables $\mathbf{r}$ and $\mathbf{s}$

$$\nabla_{\mathbf{r}}\rho\left(\mathbf{r}+\frac{\mathbf{s}}{2},\mathbf{r}-\frac{\mathbf{s}}{2},t\right) = \nabla_{\mathbf{r}_1}\rho + \nabla_{\mathbf{r}_2}\rho \tag{38}$$

$$\nabla_{\mathbf{s}}\rho\left(\mathbf{r}+\frac{\mathbf{s}}{2},\mathbf{r}-\frac{\mathbf{s}}{2},t\right) = \frac{1}{2}\nabla_{\mathbf{r}_1}\rho - \frac{1}{2}\nabla_{\mathbf{r}_2}\rho \tag{39}$$

gives the relations

$$\begin{aligned} \nabla_{\mathbf{r}_1} + \nabla_{\mathbf{r}_2} &= \nabla_{\mathbf{r}} \\ \nabla_{\mathbf{r}_1} - \nabla_{\mathbf{r}_2} &= 2\nabla_{\mathbf{s}} \\ \nabla_{\mathbf{r}_1}^2 - \nabla_{\mathbf{r}_2}^2 &= 2\nabla_{\mathbf{r}}\nabla_{\mathbf{s}} \end{aligned}$$

Now the free-electron term transforms to a diffusion term. For the sake of brevity, we write $\rho_{\mathbf{r},\mathbf{s}} = \rho(\mathbf{r} + \mathbf{s}/2, \mathbf{r} - \mathbf{s}/2, t)$ in the following.

$$\begin{aligned} &\frac{1}{i\hbar}\int -\frac{\hbar^2}{2m^*}\left(\nabla_{\mathbf{r}_1}^2 - \nabla_{\mathbf{r}_2}^2\right)\rho_{\mathbf{r},\mathbf{s}}\mathrm{e}^{-i\mathbf{k}\mathbf{s}}\,\mathrm{d}\mathbf{s} \\ &\quad= -\frac{\hbar}{im^*}\nabla_{\mathbf{r}}\int\left(\nabla_{\mathbf{s}}\rho_{\mathbf{r},\mathbf{s}}\right)\mathrm{e}^{-i\mathbf{k}\mathbf{s}}\,\mathrm{d}\mathbf{s} \\ &\quad= -\frac{\hbar\mathbf{k}}{m^*}\nabla_{\mathbf{r}}\int\rho_{\mathbf{r},\mathbf{s}}\mathrm{e}^{-i\mathbf{k}\mathbf{s}}\,\mathrm{d}\mathbf{s} \\ &\quad= -\frac{\hbar\mathbf{k}}{m^*}\nabla_{\mathbf{r}}f_{\mathrm{w}}(\mathbf{k},\mathbf{r},t) \end{aligned} \tag{40}$$

Next, we transform the potential term.

$$\frac{1}{i\hbar}\int\left[V\left(\mathbf{r}+\frac{\mathbf{s}}{2}\right)-V\left(\mathbf{r}-\frac{\mathbf{s}}{2}\right)\right]\rho_{\mathbf{r},\mathbf{s}}\mathrm{e}^{-i\mathbf{k}\mathbf{s}}\,\mathrm{d}\mathbf{s} = \int V_{\mathrm{w}}(\mathbf{k}-\mathbf{k}',\mathbf{r})\,f_{\mathrm{w}}(\mathbf{k}',\mathbf{r},t)\,\mathrm{d}\mathbf{k}' \tag{41}$$

This transformation is readily found by replacing $\rho_{\mathbf{r},\mathbf{s}}$ on the left-hand side by the inverse Fourier transformation (13). The remaining integral over $\mathbf{s}$ is denoted by V_{w} and referred to as the Wigner potential.

$$V_{\mathrm{w}}(\mathbf{q},\mathbf{r}) = \frac{1}{(2\pi)^3 i\hbar}\int\left[V\left(\mathbf{r}+\frac{\mathbf{s}}{2}\right) - V\left(\mathbf{r}-\frac{\mathbf{s}}{2}\right)\right]\mathrm{e}^{-i\mathbf{q}\mathbf{s}}\,\mathrm{d}\mathbf{s} \tag{42}$$

Using the simple relation $-(\mathbf{F}\mathbf{r}_1 - \mathbf{F}\mathbf{r}_2) = -\mathbf{F}\mathbf{s}$, the constant-field term transforms as

$$\frac{1}{i\hbar}\int(-\hbar\mathbf{F}\mathbf{s})\rho_{\mathbf{r},\mathbf{s}}\mathrm{e}^{-i\mathbf{k}\mathbf{s}}\,\mathrm{d}\mathbf{s} = -\frac{1}{\hbar}\mathbf{F}\nabla_{\mathbf{k}} f_{\mathrm{w}}(\mathbf{k},\mathbf{r},t) \tag{43}$$

Collecting the above results gives the Wigner equation for the system Hamiltonian (36).

$$\left(\frac{\partial}{\partial t}+\frac{\hbar\mathbf{k}}{m^*}\nabla_{\mathbf{r}}+\mathbf{F}\nabla_{\mathbf{k}}\right)f_{\mathrm{w}}(\mathbf{k},\mathbf{r},t) = \int V_{\mathrm{w}}(\mathbf{k}-\mathbf{k}',\mathbf{r})\,f_{\mathrm{w}}(\mathbf{k}',\mathbf{r},t)\,\mathrm{d}\mathbf{k}' \tag{44}$$

The terms are arranged so to form the classical Liouville operator on the left-hand side. The interaction of the electron with the potential distribution $V(\mathbf{r})$ is described by the potential operator on the right-hand side. As can be seen, the Wigner function in $\mathbf{k}$ and $\mathbf{r}$ depends in a non-local manner on the Wigner function in all other momentum points $\mathbf{k}'$ and through V_{w} also on the potential at all other locations $\mathbf{r} \pm \mathbf{s}/2$.

2.4 *Electron–phonon interaction*

The Wigner equation has frequently been solved using the finite-difference method (Biegel and Plummer, 1996; Frensley, 1990), assuming the phenomenological relaxation time approximation for dissipative transport. Recently developed Monte Carlo methods allowed phonon scattering to be included semi-classically in quantum device simulations (Kosina et al., 2003a,b; Shifren and Ferry, 2002). Use of a Boltzmann scattering operator acting on the Wigner distribution was originally suggested by Frensley (1990). A rigorous treatment of electron–phonon interaction in the Wigner formalism based on a one-electron, many-phonon Hamiltonian was reported by Rossi et al. (1994), Bertoni et al. (1999) and Bordone et al. (1999). However, for the purpose of numerical simulation of mesoscopic devices this formalism turned out to be rather complex. Introducing the weak scattering limit and assuming the phonon system to be in thermodynamic equilibrium, an equation for the reduced Wigner function is obtained (Brunetti et al., 1989; Nedjalkov et al., 2002; Rossi et al., 1994). Since this equation takes into account the finiteness of the interaction time, it is time reversible and includes effects such as collisional broadening, collisional retardation and the intracollisional field effect (Gurov et al., 2002; Nedjalkov et al., 2001). For a uniform electric field, the equation reduces to the Levinson equation (1970). Introducing the classical limit, the scattering operator of the equation for the reduced Wigner function reduces to the Boltzmann scattering operator. This limit implies that the system is considered on a timescale much larger than the timescale of the lattice vibrations (Ringhofer et al., 2004). The resulting Wigner equation is of the form

$$\left(\frac{\partial}{\partial t}+\frac{\hbar\mathbf{k}}{m^*}\nabla_r\right)f_{\mathrm{w}} = \Theta_{\mathrm{w}}[f_{\mathrm{w}}] + Q[f_{\mathrm{w}}] \tag{45}$$

where Θ_{w} denotes the potential operator,

$$\Theta_{\mathrm{w}}[f_{\mathrm{w}}](\mathbf{k},\mathbf{r},t) = \int V_{\mathrm{w}}(\mathbf{k}-\mathbf{k}',\mathbf{r})\,f_{\mathrm{w}}(\mathbf{k}',\mathbf{r},t)\,\mathrm{d}\mathbf{k}' \tag{46}$$

and Q the Boltzmann scattering operator,

$$Q[f_{\mathrm{w}}](\mathbf{k},\mathbf{r},t) = \int\left[S(\mathbf{k},\mathbf{k}')\,f_{\mathrm{w}}(\mathbf{k}') - S(\mathbf{k}',\mathbf{k})\,f_{\mathrm{w}}(\mathbf{k})\right]\mathrm{d}\mathbf{k}' \tag{47}$$

The scattering rates is given by Fermi's golden rule, with M denoting the interaction matrix element and $\hbar\omega_{\mathbf{q}}$ the phonon energy.

$$S(\mathbf{k}',\mathbf{k}) = \frac{V}{(2\pi)^3}\sum_{\nu=\pm1}\frac{2\pi}{\hbar}M^2(\mathbf{q})\left(N_{\mathbf{q}}+\frac{1}{2}-\frac{\nu}{2}\right) \times\delta[\epsilon(\mathbf{k}')-\epsilon(\mathbf{k})+\nu\hbar\omega_{\mathbf{q}}] \tag{48}$$

Consequences of the classical limit are that scattering events are instantaneous in time, energy is conserved exactly as expressed by the δ-function in (48) and the equation is time-irreversible.

2.5 *Classical force*

For a one-dimensional device, the definition of the Wigner potential (**??**) reduces to

$$V_{\mathrm{w}}(q,x) = \frac{1}{2\pi i\hbar}\int_{-\infty}^{\infty}\left[V\left(x+\frac{s}{2}\right) - V\left(x-\frac{s}{2}\right)\right]\mathrm{e}^{-iqs}\,\mathrm{d}s \tag{49}$$

This potential is assumed to be constant outside the simulation domain. When a voltage is applied to the device, the integrand of (**??**) becomes constant for large s: $V(x+s/2)-V(x-s/2)\to(-e)V_{\mathrm{app}}$. Since the integrand does not vanish for $s\to\infty$ the Fourier integral will diverge at $q=0$, giving rise to a singularity in $V_{\mathrm{w}}(q,x)$. Therefore, one should exclude the interval $|q| < q_{\mathrm{c}}/2$ from the domain of integration, where q_{c} is some small wave number (Gehring and Kosina, 2004). The potential operator (46) is rewritten as

$$\Theta_{\mathrm{w}}[f_{\mathrm{w}}](k,x,t) = \int_{-\infty}^{\infty}V_{\mathrm{w}}(q,x)\,f_{\mathrm{w}}(k-q,x,t)\,\mathrm{d}q \tag{50}$$

In the the integral over the small wave numbers the term $f_{\mathrm{w}}(k-q)\simeq -q\,\partial_k f_{\mathrm{w}}(k)$ can be linearised, yielding a local classical force term.

$$\Theta_{\rm cl}[f](k,x,t) = \frac{1}{\hbar}\frac{\partial V_{\rm cl}(x)}{\partial x}\,\frac{\partial f_{\rm w}(k,x,t)}{\partial k} \tag{51}$$

$$V_{\rm cl}(x) = \frac{1}{2\pi}\int_{-q_{\rm c}}^{q_{\rm c}} \hat{V}(q)\mathrm{e}^{iqx}\mathrm{d}q \tag{52}$$

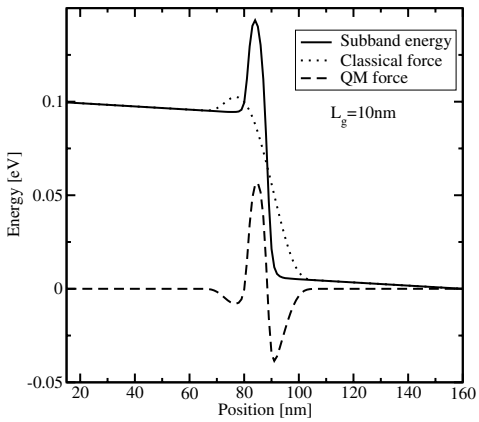
In this equation, $\hat{V}(q)$ is the Fourier transform of the potential. Equation (52) motivates the following spectral decomposition of the potential profile into a slowly varying, classical component (52) and a rapidly varying, quantum mechanical component.

$$V(x) = V_{\rm cl}(x) + V_{\rm qm}(x) \tag{53}$$

An example for this decomposition is shown in Figure 1. A double gate MOSFET with 10 nm gate length is considered and the potential energy of the electrons is taken to be the first subband in the device.This decomposition yields a Wigner equation including both, a local classical force term with $F_{\rm cl} = -\partial_x V_{\rm cl}$ and a non-local potential operator.

$$\left(\frac{\partial}{\partial t} + v_x\frac{\partial}{\partial x} + \frac{F_{\rm cl}}{\hbar}\frac{\partial}{\partial k}\right) f_{\rm w}(k,x,t) = \int V_{\rm w}^{\rm qm}(q,x)\, f_{\rm w}(k-q,x,t)\mathrm{d}q + Q[f_{\rm w}](k) \tag{54}$$

Figure 1 Potential decomposition of the subband potential into a classical part $V_{\rm cl}(x)$ and a quantum-mechanical part $V_{\rm qm}(x)$ for the 10 nm gate length device



Scattering is taken into account through the Boltzmann scattering operator Q. The Wigner potential is calculated from the quantum mechanical potential component, $V_{\rm qm}$.

$$V_{\rm w}^{\rm qm}(q,x) = \frac{1}{2\pi i\hbar}\int_{-\infty}^{\infty}\left[V_{\rm qm}\left(x+\frac{s}{2}\right) - V_{\rm qm}\left(x-\frac{s}{2}\right)\right]\mathrm{e}^{-iqs}\,\mathrm{d}s$$

Realistic devices often consist of a quantum region embedded in an extended classical region. The discrete Fourier transform requires a uniform discretisation of the quantum region with some step size Δx. To obtain a good resolution of the Wigner potential, a fairly large value for N_k, the number of points for the discrete Fourier transform, has to be chosen. Typical values are in the range $10^2 - 10^3$. In this case the domain of the Fourier transform, $L_{\rm FT} = N_k\Delta x$, can become much larger than the simulation domain. For the purpose of the Fourier transform, the potential outside the device region is extrapolated by a constant. In the discrete system, the classical potential component defined by (52) becomes:

$$V_{\rm cl}(x_j) = \frac{\sum_l w_{jl}\, V(x_l)}{\sum_l w_{jl}}, \quad -\frac{N_k}{2} \le l \le \frac{N_k}{2} \tag{55}$$

$$w_{jl} = \frac{\sin\big((j-l)q_{\rm c}\Delta x\big)}{(j-l)q_{\rm c}\Delta x} \tag{56}$$

The cutoff wavenumber $q_{\rm c}$ is conveniently specified by a cutoff wave length as

$$q_{\rm c} = \frac{2\pi}{\lambda_{\rm c}} \tag{57}$$

It can be seen from (56) that $\lambda_{\rm c} = 2\Delta x$ would give $V_{\rm cl}(x_j) = V(x_j)$ and hence the quantum component would vanish. This choice of $\lambda_{\rm c}$ gives the classical limit of the discrete system. For a quantum transport calculation one has to choose $\lambda_{\rm c} \gg 2\Delta x$.

2.6 *Integral form of the Wigner equation*

From the integro-differential form of the Wigner equation, a path-integral formulation can be derived. The equation to be transformed reads

$$\left(\frac{\partial}{\partial t} + \mathbf{v}(\mathbf{k})\nabla_r + \mathbf{F}(\mathbf{r})\nabla_k\right) f_{\rm w}(\mathbf{k},\mathbf{r},t) = \int [S(\mathbf{k},\mathbf{k}') + V_{\rm w}(\mathbf{k}-\mathbf{k}',\mathbf{r}) + \alpha(\mathbf{k},\mathbf{r})\,\delta(\mathbf{k}-\mathbf{k}')] \times f_{\rm w}(\mathbf{k}',\mathbf{r},t)\,\mathrm{d}\mathbf{k}' - [\lambda(\mathbf{k},\mathbf{r}) + \alpha(\mathbf{k},\mathbf{r})] f_{\rm w}(\mathbf{k},\mathbf{r},t) \tag{58}$$

where λ denotes the total scattering rate, $\lambda(\mathbf{k}) = \int S(\mathbf{k}',\mathbf{k})\,\mathrm{d}\mathbf{k}'$ and a fictious scattering mechanism of the form $\alpha\delta(\mathbf{k}-\mathbf{k}')$ is introduced, referred to as self-scattering (Jacoboni and Reggiani, 1983). Because of the δ-function, this mechanism does not change the state of the electron and hence does not affect the solution of the equation. For the sake of brevity, we define an integral kernel Γ and the symbols μ and U.

$$\mu(\mathbf{k},\mathbf{r}) = \lambda(\mathbf{k},\mathbf{r}) + \alpha(\mathbf{k},\mathbf{r})$$

$$\Gamma(\mathbf{k},\mathbf{k}',\mathbf{r}) = \frac{S(\mathbf{k},\mathbf{k}') + V_{\rm w}(\mathbf{k}-\mathbf{k}',\mathbf{r}) + \alpha(\mathbf{k},\mathbf{r})\,\delta(\mathbf{k}-\mathbf{k}')}{\mu(\mathbf{k}',\mathbf{r})}$$

$$U(\mathbf{k},\mathbf{r},t) = \int \Gamma(\mathbf{k},\mathbf{k}',\mathbf{r})\,\mu(\mathbf{k}',\mathbf{r})\; f_{\rm w}(\mathbf{k}',\mathbf{r},t)\,\mathrm{d}\mathbf{k}'$$

The Liouville operator in (58) is treated by the method of characteristics. One introduces path variables $\mathbf{K}(t)$ and $\mathbf{R}(t)$ and takes the total time derivative of $f_{\rm w}$.

$$\frac{\mathrm{d}}{\mathrm{d}t} f_{\rm w}(\mathbf{K}(t),\mathbf{R}(t),t) = \left(\frac{\partial}{\partial t} + \frac{\mathrm{d}\mathbf{K}(t)}{\mathrm{d}t}\nabla_{\mathbf{k}} + \frac{\mathrm{d}\mathbf{R}(t)}{\mathrm{d}t}\nabla_{\mathbf{r}}\right) f_{\rm w}$$

The right-hand side equals the Liouville operator if the path variables satisfy the following equations of motion.

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{K}(t) = \mathbf{F}(\mathbf{R}(t)) \quad \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{R}(t) = \mathbf{v}(\mathbf{K}(t)) \tag{59}$$

Now we assume some phase-space point $\mathbf{k}$, $\mathbf{r}$ and some time t to be given. A phase-space trajectory with the initial condition $\mathbf{K}(t' = t) = \mathbf{k}$ and $\mathbf{R}(t' = t) = \mathbf{r}$ is obtained by formal integration.

$$\mathbf{K}(t') = \mathbf{k} + \int_t^{t'} \mathbf{F}(\mathbf{R}(y))\,\mathrm{d}y, \quad \mathbf{R}(t') = \mathbf{r} + \int_t^{t'} \mathbf{v}(\mathbf{K}(y))\,\mathrm{d}y$$

Note that $\mathbf{k}, \mathbf{r}, t$ are treated as constants in the following derivation, only t' is a variable. Introducing the functions

$$\tilde{f}_\mathrm{w}(t') = f_\mathrm{w}(\mathbf{K}(t'), \mathbf{R}(t'), t')$$
$$\tilde{\mu}(t') = \mu(\mathbf{K}(t'), \mathbf{R}(t'))$$
$$\tilde{U}(t') = U(\mathbf{K}(t'), \mathbf{R}(t'), t')$$

allows (58) to be rewritten as an ordinary differential equation of first order.

$$\frac{\mathrm{d}}{\mathrm{d}t'}\tilde{f}_\mathrm{w}(t') + \tilde{\mu}(t')\tilde{f}_\mathrm{w}(t') = \tilde{U}(t') \tag{60}$$

When multiplied by an integrating factor of the form $\exp\left[\int_0^{t'} \tilde{\mu}(y)\mathrm{d}y\right]$, the equation can easily be integrated in time.

$$\frac{\mathrm{d}}{\mathrm{d}t'}\exp\left[\int_0^{t'} \tilde{\mu}(y)\mathrm{d}y\right]\tilde{f}_\mathrm{w}(t') = \exp\left[\int_0^{t'} \tilde{\mu}(y)\mathrm{d}y\right]\tilde{U}(t') \tag{61}$$

The choice of the upper and lower bounds of time integration depends on whether the problem under consideration is time-dependent or stationary.

The ordinary differential equation Equation (61), which is the result of treating the Liouville operator by the method of characteristics, has the same structure as the corresponding differential equation for the Boltzmann equation. Therefore, we can refer to the work on the Boltzmann equation regarding the details of the time integration of (61) (Kosina and Nedjalkov, 2003; Kosina et al., 2000).

2.6.1 The time-dependent equation

The upper bound of the time integration should be $t' = t$ to obtain $\tilde{f}_\mathrm{w}(t) = f_\mathrm{w}(\mathbf{k}, \mathbf{r}, t)$, the value of the unknown at the given phase-space point. At $t' = 0$, an initial distribution $f_\mathrm{i}(\mathbf{k}, \mathbf{r})$ is assumed to be given. In analogy with the Boltzmann equation (Kosina and Nedjalkov, 2003), the integral form of the Wigner equation is obtained.

$$\begin{aligned} f_\mathrm{w}(\mathbf{k}, \mathbf{r}, t) = &\int_0^t \mathrm{d}t' \int \mathrm{d}\mathbf{k}' \exp\left\{-\int_{t'}^t \mu[\mathbf{K}(y), \mathbf{R}(y)]\mathrm{d}y\right\} \\ &\times \Gamma[\mathbf{K}(t'), \mathbf{k}', \mathbf{R}(t')]\ \mu[\mathbf{k}', \mathbf{R}(t')]\ f_\mathrm{w}[\mathbf{k}', \mathbf{R}(t'), t'] \\ &+ \exp\left\{-\int_0^t \mu[\mathbf{K}(y), \mathbf{R}(y)]\mathrm{d}y\right\} f_\mathrm{i}[\mathbf{K}(0), \mathbf{R}(0)] \end{aligned} \tag{62}$$

This equation states that the Wigner function at time t depends on the Wigner function at some previous time t'. Using (62) in an iterative procedure, with each iteration the time variable would move to smaller values. Therefore, another equation is desirable that describes the evolution of the system in forward time direction. Such an equation is given by the adjoint equation of (62).

$$\begin{aligned} g_\mathrm{w}(\mathbf{k}', \mathbf{r}', t') = &\int_{t'}^\infty \mathrm{d}\tau \int \mathrm{d}\mathbf{k}\ g_\mathrm{w}[\mathbf{K}(\tau), \mathbf{R}(\tau), \tau] \\ &\times \exp\left\{-\int_{t'}^\tau \mu[\mathbf{K}(y), \mathbf{R}(y)]\mathrm{d}y\right\}\Gamma(\mathbf{k}, \mathbf{k}', \mathbf{r}')\ \mu(\mathbf{k}', \mathbf{r}') \\ &+ g_0(\mathbf{k}', \mathbf{r}', t') \end{aligned} \tag{63}$$

The derivation of the adjoint equation (63) is discussed in detail by Kosina et al. (2000) and Kosina and Nedjalkov (2003).

2.6.2 The stationary equation

In a stationary system, the potential and all material parameters are independent of time. A phase-space trajectory is invariant under time translations. This property can be conveniently used to adjust the time reference of each trajectory (Kosina et al., 2003a,b; Nedjalkov et al., 2003). In the stationary case we assume the phase-space point $\mathbf{k}$, $\mathbf{r}$ to be given at $t' = 0$. So the initial condition for the phase-space trajectory is $\mathbf{K}(0) = \mathbf{k}$ and $\mathbf{R}(0) = \mathbf{r}$. For the upper bound of time integration of (61), we choose now $t' = 0$ to obtain $\tilde{f}_\mathrm{w}(0) = f_\mathrm{w}(\mathbf{k}, \mathbf{r})$. The lower time bound has to be chosen such that the functions $\mathbf{K}(t)$ and $\mathbf{R}(t)$ take on values at which the Wigner function is known. In the steady-state, this function is known only at the domain boundary. An appropriate lower time bound is therefore the time when the trajectory enters the simulation domain. This time is denoted by t_b^- and depends on the point $\mathbf{k}$, $\mathbf{r}$ under consideration. The case that the real space trajectory $\mathbf{R}(t)$ never intersects the domain boundary can occur for a classically bound state. Then the trajectory forms a closed loop and the appropriate choice is $t_b^- = -\infty$. Integration of (61) in the time bounds discussed above results in the integral form of the stationary Wigner equation (cf. Kosina et al., 2003a,b).

$$\begin{aligned} f(\mathbf{k}, \mathbf{r}) = &f_0(\mathbf{k}, \mathbf{r}) \\ &+ \int_{t_b^-(\mathbf{k},\mathbf{r})}^0 \mathrm{d}t' \int \mathrm{d}\mathbf{k}' \exp\left\{-\int_{t'}^0 \mu[\mathbf{K}(y), \mathbf{R}(y)]\,\mathrm{d}y\right\} \\ &\times \Gamma[\mathbf{K}(t'), \mathbf{k}', \mathbf{R}(t')]\ \mu(\mathbf{k}', \mathbf{r}')\ f_\mathrm{w}[\mathbf{k}', \mathbf{R}(t')] \end{aligned} \tag{64}$$

$$\begin{aligned} f_0(\mathbf{k}, \mathbf{r}) = &f_\mathrm{b}\left\{\mathbf{K}[t_b^-(\mathbf{k}, \mathbf{r})], \mathbf{R}[t_b^-(\mathbf{k}, \mathbf{r})]\right\} \\ &\times \exp\left\{-\int_{t_b^-(\mathbf{k},\mathbf{r})}^0 \lambda[\mathbf{K}(y), \mathbf{R}(y)]\,\mathrm{d}y\right\} \end{aligned} \tag{65}$$

Here, f_b denotes the boundary distribution. The integral form (64) represents a backward equation. The corresponding forward equation is given by the adjoint equation.

$$g_{\mathrm{w}}(\mathbf{k},\mathbf{r}) = g_0(\mathbf{k},\mathbf{r}) + \int_0^{t_b^+(\mathbf{k}',\mathbf{r})} \mathrm{d}t \int d\mathbf{k}' g_{\mathrm{w}}[\mathbf{K}(\tau),\mathbf{R}(\tau)] \times \exp\left\{-\int_t^\tau \mu[\mathbf{K}(y),\mathbf{R}(y)]dy\right\} \Gamma(\mathbf{k}',\mathbf{k},\mathbf{r})\,\mu(\mathbf{k},\mathbf{r})\,\Theta_D(\mathbf{r}) \tag{66}$$

Θ_D denotes the indicator function of the simulation domain D. The initial condition for the phase-space trajectory is $\mathbf{K}(t) = \mathbf{k}'$ and $\mathbf{R}(t) = \mathbf{r}$.

3 The Monte Carlo method

Monte Carlo is a numerical method that can be applied to solve integral equations. Applying this method to the various integral formulations of the Wigner equation gives rise to a variety of Monte Carlo algorithms, as discussed in the following.

3.1 The general scheme

This section introduces the general scheme of the Monte Carlo method and outlines its application to the solution of integrals and integral equations. To calculate some unknown value m by the Monte Carlo method, one has to find a random variable ξ whose expectation value equals $\mathrm{E}\{\xi\} = m$. The variance of ξ is designated σ^2, with σ being the standard deviation.

Now consider N independent random variables $\xi_1, \xi_2, \ldots, \xi_N$ with distributions identical to that of ξ. Consequently, their expectation values and their variance are equal.

$$\mathrm{E}\{\xi_i\} = m, \quad \mathrm{Var}\{\xi_i\} = \sigma^2, \quad i = 1, 2, \ldots, N \tag{67}$$

Expectation value and variance of the sum of all these random variables are given by

$$\mathrm{E}\{\xi_1 + \xi_2 + \cdots + \xi_N\} = \mathrm{E}\{\xi_1\} + \cdots + \mathrm{E}\{\xi_N\} = Nm \tag{68}$$

$$\mathrm{Var}\{\xi_1 + \xi_2 + \cdots + \xi_N\} = \mathrm{Var}\{\xi_1\} + \cdots + \mathrm{Var}\{\xi_N\} = N\sigma^2 \tag{69}$$

Using the properties $\mathrm{E}\{c\xi\} = c\mathrm{E}\{\xi\}$ and $\mathrm{Var}\{c\xi\} = c^2\mathrm{Var}\{\xi\}$, one obtains from (68) and (69)

$$\mathrm{E}\left\{\frac{1}{N}(\xi_1 + \xi_2 + \cdots + \xi_N)\right\} = m \tag{70}$$

$$\mathrm{Var}\left\{\frac{1}{N}(\xi_1 + \xi_2 + \cdots + \xi_N)\right\} = \frac{\sigma^2}{N} \tag{71}$$

Therefore, the random variable

$$\overline{\xi} = \frac{1}{N}\sum_{i=1}^{N} \xi_i \tag{72}$$

has the same expectation value as ξ and an N times reduced variance. A Monte Carlo simulation of the unknown m consists of drawing one random number $\overline{\xi}$. Indeed, this is equivalent to drawing N values of the random variable ξ and evaluating the sample mean (72).

The Monte Carlo method gives an estimate of both the result and the error. According to the central limit theorem, the sum $\rho_N = \xi_1 + \xi_2 + \cdots + \xi_N$ of a large number of identical random variables is approximately normal. For this reason, the following three-sigma rule holds only approximately

$$P\{|\rho_N - Nm| < 3\sqrt{N\sigma^2}\} \approx 0.997 \tag{73}$$

In this equation, the expectation value and the variance of ρ_N are given by (68) and (69), respectively. Dividing the inequality by N and using $\overline{\xi} = \rho_N/N$ we arrive at an equivalent inequality and the probability will not change:

$$P\left\{|\overline{\xi} - m| < 3\frac{\sigma}{\sqrt{N}}\right\} \approx 0.997 \tag{74}$$

This formula indicates that the sample mean $\overline{\xi}$ will be approximately equal to m. The error of this approximation will most probably not exceed the value $3\sigma/\sqrt{N}$. This error evidently approaches zero as N increases (Sobol, 1984).

3.1.1 Monte Carlo integration

We apply the Monte Carlo method to the evaluation of an integral.

$$m = \int_a^b \phi(x)\,\mathrm{d}x \tag{75}$$

For this purpose, the integrand has to be decomposed into a product $\phi = p\,\psi$, where p is a density function, which means that p is non-negative and satisfies $\int_a^b p(x)\,\mathrm{d}x = 1$. Integral (75) becomes

$$m = \int_a^b p(x)\psi(x)\,\mathrm{d}x \tag{76}$$

and denotes the expectation value $m = \mathrm{E}\{\Psi\}$ of some random variable $\Psi = \psi(X)$. Now the general scheme described in the previous section can be applied. First, a sample $x_1, \ldots, x_N$ is generated from the density p. Then the sample $\psi_1, \ldots, \psi_N$ is obtained by evaluating the function ψ: $\psi_i = \psi(x_i)$. The sample mean

$$m \simeq \overline{\psi} = \frac{1}{N}\sum_{i=1}^{N} \psi_i \tag{77}$$

approximates the expectation value. To employ the error estimation (74), the variance of Ψ can be approximately evaluated by the sample variance

$$\sigma^2 \simeq \overline{\sigma}^2 = \frac{1}{N-1}\sum_{i=1}^{N} (\psi_i - \overline{\psi})^2 \tag{78}$$

Because the factorisation of the integrand is not unique, different random variables can be introduced depending on the choice of the density p. All of them have the same expectation value but different variance.

3.1.2 Integral equations

The kinetic equations considered in this work can be formulated as integral equations of the form

$$f(x) = \int K(x, x') f(x')\, \mathrm{d}x' + f_0(x) \tag{79}$$

where the kernel K and the source term f_0 are given functions. Equations of this form are known as Fredholm integral equations of the second kind. In the particular cases of the Boltzmann equation and the Wigner equation the unknown function f represents the phase-space distribution function. The multidimensional variable x stands for $(\mathbf{k}, \mathbf{r}, t)$ in the transient case and for $(\mathbf{k}, \mathbf{r})$ in the steady-state.

Substituting (79) recursively into itself gives the Neumann series, which, if convergent, is a formal solution to the integral equation (Byron and Fuller, 1992).

$$f = f^{(0)} + f^{(1)} + f^{(2)} + \cdots \tag{80}$$

The iteration terms are defined recursively beginning with $f^{(0)}(x) = f_0(x)$.

$$f^{(n+1)}(x) = \int K(x, x') f^{(n)}(x')\, \mathrm{d}x' \,, n = 0, 1, 2, \ldots \tag{81}$$

The series (80) yields the function value in some given point x. However, in many cases one is interested in mean values of f rather than in a point-wise evaluation. Such a mean value represents a linear functional and can be expressed as an inner product.

$$(f, A) = \int f(x) A(x) \mathrm{d}x \tag{82}$$

It is to note that (79) is a backward equation. The corresponding forward equation is given by the adjoint equation,

$$g(x') = \int K^{\dagger}(x', x) g(x)\, \mathrm{d}x + A(x') \tag{83}$$

where the kernel is defined by $K^{\dagger}(x', x) = K(x, x')$. Multiplying (79) by $g(x)$ and (83) by $f(x')$, and integrating over x and x', respectively, results in the equality

$$(f, A) = (g, f_0) \tag{84}$$

By means of (84), one can calculate a statistical mean value not only from f, but also from g, the solution of the adjoint equation. The given function A has to be used as the source term of the adjoint equation. The link with the numerical Monte Carlo method is established by evaluating the terms of the Neumann series by Monte Carlo integration, as pointed out in the previous section.

Note that usage of (84) precludes a point-wise evaluation of the distribution function using a forward algorithm, because $A(x) = \delta(x)$ cannot be treated by the Monte Carlo method. The probability for a continuous random variable x' to assume a given value x is zero. Only the probability of finding x' within a small but finite volume around x is non-zero.

3.2 Particle models

Each term of the Neumann series of the adjoint equation describes a sequence of alternating free flight and scattering events. A transition consisting of a free flight with initial state $\mathbf{k}_\mathrm{i}$ at time t_i and a scattering process to the final state $\mathbf{k}_\mathrm{f}$ at time t_f is described by the following expression. For the sake of brevity, the $\mathbf{r}$-dependence of Γ and μ is omitted in the following.

$$\begin{aligned} P(\mathbf{k}_\mathrm{f}, t_\mathrm{f}, \mathbf{k}_\mathrm{i}, t_\mathrm{i}) &= \Gamma[\mathbf{k}_\mathrm{f}, \mathbf{K}_\mathrm{i}(t_\mathrm{f})]\ \mu[\mathbf{K}_\mathrm{i}(t_\mathrm{f})] \\ &\quad \times \exp\left\{-\int_{t_\mathrm{i}}^{t_\mathrm{f}} \mu[\mathbf{K}_\mathrm{i}(\tau)]\, \mathrm{d}\tau\right\} \end{aligned} \tag{85}$$

In a Monte Carlo simulation, the time of the next scattering event, t_f, is generated from an exponential distribution, given by the terms $\mu \exp()$ in (85). Then, a transition from the trajectory end point $\mathbf{K}_\mathrm{i}(t_\mathrm{f})$ to the final state $\mathbf{k}_\mathrm{f}$ is realised using the kernel Γ. In contrast to the classical case, where P would represent a transition probability, such an interpretation is not possible in the case of the Wigner equation because P is not positive semi-definite. The problem originates from the Wigner potential, which assumes positive and negative values. However, because of its antisymmetry with respect to $\mathbf{q}$, the Wigner potential can be reformulated in terms of one positive function V_w^{+} (Kosina et al., 2003a,b).

$$V_\mathrm{w}^{+}(\mathbf{q}, \mathbf{r}) = \max(0, V_\mathrm{w}(\mathbf{q}, \mathbf{r})) \tag{86}$$

$$V_\mathrm{w}(\mathbf{q}, \mathbf{r}) = V_\mathrm{w}^{+}(\mathbf{q}, \mathbf{r}) - V_\mathrm{w}^{+}(-\mathbf{q}, \mathbf{r}) \tag{87}$$

Then, the kernel Γ is rewritten as a sum of transition probabilities.

$$\begin{aligned} \Gamma(\mathbf{k}, \mathbf{k}') &= \frac{\lambda}{\mu} s(\mathbf{k}, \mathbf{k}') + \frac{\alpha}{\mu} \delta(\mathbf{k}' - \mathbf{k}) \\ &\quad + \frac{\gamma}{\mu} [w(\mathbf{k}, \mathbf{k}') - w^{*}(\mathbf{k}, \mathbf{k}')] \end{aligned} \tag{88}$$

The transition probabilities are given by

$$s(\mathbf{k}, \mathbf{k}') = \frac{S(\mathbf{k}', \mathbf{k})}{\lambda(\mathbf{k}')} \tag{89}$$

$$w(\mathbf{k}, \mathbf{k}') = \frac{V_\mathrm{w}^{+}(\mathbf{k} - \mathbf{k}')}{\gamma} \tag{90}$$

$$w^{*}(\mathbf{k}, \mathbf{k}') = w(\mathbf{k}', \mathbf{k}) \tag{91}$$

The normalisation factor associated with the Wigner potential is defined as

$$\gamma(\mathbf{r}) = \int V_\mathrm{w}^{+}(\mathbf{q}, \mathbf{r})\, \mathrm{d}\mathbf{q} \tag{92}$$

In the following, different variants of generating the final state $\mathbf{k}_\mathrm{f}$ from the kernel Γ will be discussed.

3.2.1 The Markov chain method

In analogy to the simple integral (76), we have now to decompose the kernel P into a transition probability p and the remaining function P/p. More details on the Markov chain method can be found by Ermakow (1975) and Hammersley

and Handscomb (1964). With respect to (85), one could use the absolute value of Γ as a transition probability. Practically, it is more convenient to use the absolute values of the components of Γ, giving the following transition probability.

$$p(\mathbf{k}_\mathrm{f}, \mathbf{k}') = \frac{\lambda}{\nu}\, s(\mathbf{k}_\mathrm{f}, \mathbf{k}') + \frac{\alpha}{\nu}\, \delta(\mathbf{k}_\mathrm{f} - \mathbf{k}') + \frac{\gamma}{\nu}\, w(\mathbf{k}_\mathrm{f}, \mathbf{k}') + \frac{\gamma}{\nu}\, w^*(\mathbf{k}_\mathrm{f}, \mathbf{k}') \tag{93}$$

The normalisation factor is $\nu = \lambda + \alpha + 2\gamma$. The free-flight time is generated from the exponential distribution appearing in (85).

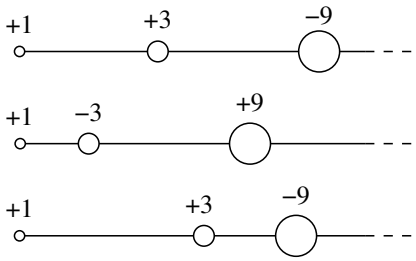
$$p_\mathrm{t}(t_\mathrm{f}, t_\mathrm{i}, \mathbf{k}_\mathrm{i}) = \mu[\mathbf{K}_\mathrm{i}(t_\mathrm{f})] \exp\left\{-\int_{t_\mathrm{i}}^{t_\mathrm{f}} \mu[\mathbf{K}_\mathrm{i}(\tau)]\, \mathrm{d}\tau\right\} \tag{94}$$

For the sake of brevity, the state at the end of the free flight is labelled $\mathbf{k}' = \mathbf{K}_\mathrm{i}(t_\mathrm{f})$ in the following. To generate the final state $\mathbf{k}_\mathrm{f}$, one of the four terms in (93) is selected with the associated probabilities λ/ν, α/ν, γ/ν, and γ/ν, respectively. Apparently, these probabilities sum up to one. If classical scattering is selected, $\mathbf{k}_\mathrm{f}$ is generated from s. If self-scattering is selected, the state does not change and $\mathbf{k}_\mathrm{f} = \mathbf{k}'$ holds. If the third or fourth term is selected, the particle state is changed by scattering from the Wigner potential and $\mathbf{k}_\mathrm{f}$ is selected from w or w^*, respectively. The particle weight has to be multiplied by the ratio

$$\frac{\Gamma}{p} = \pm\left(1 + \frac{2\gamma}{\lambda + \alpha}\right) \tag{95}$$

where the minus sign applies if $\mathbf{k}_\mathrm{f}$ has been generated from w^*. For instance, for a quantum mechanical system, where the classical scattering rate λ is less than the Wigner scattering rate γ, the self-scattering rate α can be chosen in such a way that $\lambda + \alpha = \gamma$. Then, the multiplier (95) evaluates to ± 3. An ensemble of particles would evolve as shown schematically in Figure 2. As the multiplier (95) is always greater than one, the absolute value of the particle weight will inevitably grow with the number of transitions on the trajectory. To solve the problem of growing particle weights, one can split particles. In this way, an increase in particle weight is transformed to an increase in particle number.

Figure 2 With the Markov chain method, the number of numerical particles is conserved. The magnitude of the particle weight increases with each event, and the sign of the weight changes randomly according to a given probability distribution



3.2.2 *Pair generation method*

The basic idea of splitting is refined so to avoid fractional weights. Different interpretations of the kernel are presented that conserve the magnitude of the particle weight. Choosing the initial weight to be $+1$, all generated particles will have weight $+1$ or -1. This is achieved by interpreting the potential operator in (45) as a generation term of positive and negative particles. We consider the kernel (88).

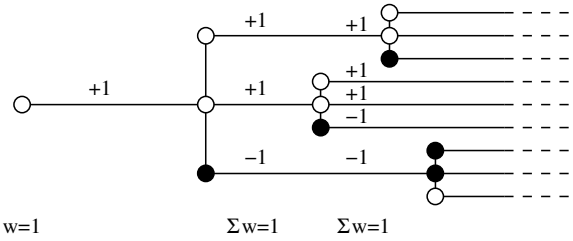
$$\Gamma(\mathbf{k}_\mathrm{f}, \mathbf{k}') = \frac{\lambda}{\mu}\, s(\mathbf{k}_\mathrm{f}, \mathbf{k}') + \frac{\alpha}{\mu}\, \delta(\mathbf{k}_\mathrm{f} - \mathbf{k}') + \frac{\gamma}{\mu}\left[w(\mathbf{k}_\mathrm{f}, \mathbf{k}') - w^*(\mathbf{k}_\mathrm{f}, \mathbf{k}')\right] \tag{96}$$

If the Wigner scattering rate γ is larger than the classical scattering rate λ, the self-scattering rate α has to be chosen large enough to satisfy the inequality $\gamma/\mu \leq 1$. Typical choices are $\mu = \mathrm{Max}(\lambda, \gamma)$ or $\mu = \lambda + \gamma$. These expressions also hold for the less interesting case $\gamma < \lambda$, where quantum interference effects are less important than classical scattering effects. In the following, we discuss the case $\gamma > \lambda$, where quantum effects are dominant. We choose the self-scattering rate equal to $\alpha = \gamma$ and regroup the kernel as

$$\Gamma(\mathbf{k}_\mathrm{f}, \mathbf{k}') = \frac{\lambda}{\mu}\, s(\mathbf{k}_\mathrm{f}, \mathbf{k}') + \left(1 - \frac{\lambda}{\mu}\right)\left[\delta(\mathbf{k}_\mathrm{f} - \mathbf{k}') + w(\mathbf{k}_\mathrm{f}, \mathbf{k}') - w^*(\mathbf{k}_\mathrm{f}, \mathbf{k}')\right] \tag{97}$$

As in the classical Monte Carlo method, the distribution of the free-flight duration is given by the exponential distribution (94). At the end of a free flight, classical scattering is selected with probability $p_\mathrm{s} = \lambda/\mu$. With the complementary probability, $1 - p_\mathrm{s}$, a self-scattering event and a pair generation event occur. The weight of the state generated from w^* is multiplied by -1. The weights of the states generated from w and from self-scattering do not change. Therefore, the magnitude of the initial particle weight is conserved, as shown in Figure 3.

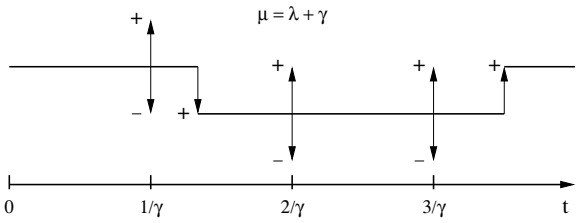
Figure 3 With the pair generation method the magnitude of the particle weight is conserved, but one initial particle generates a cascade of numerical particles. At all times mass is exactly conserved



In this algorithm, classical scattering and pair generation cannot occur at the same time, as shown in Figure 4.

In the pair-generation method described, the weight of the generated particles is ± 1, because the generation rate used equals 2γ. If a generation rate larger than 2γ or a fixed time-step less than $1/2\gamma$ were used, the magnitude of the generated weight would be less than one. This approach has been followed by Shifren et al. (2002), where the resulting fractional weights are termed affinities. On the other hand, a generation rate less than 2γ would result in an under-sampling of the physical process. Then, the magnitude of the generated weights would be generally greater than one.

Figure 4 Trajectory of a sample particle resulting from the pair-generation method



3.3 The negative sign problem

In the following, we analyse the growth rates of particle weights and particle numbers associated with the different Monte Carlo algorithms. In the Markov chain method discussed in Section 3.2.1, the weight increases at each scattering event by the multiplier (95). The growth rate of the weight can be estimated for the case of constant coefficients γ and μ. Because free-flight times are generated with rate μ, the mean free-flight time will be $1/\mu$. During a given time interval t, on average $n = \mu t$ scattering events will occur. The total weight is then estimated asymptotically for $t \gg 1/\mu$.

$$|W(t)| = \left(1 + \frac{2\gamma}{\mu}\right)^n = \left(1 + \frac{2\gamma t}{n}\right)^n \simeq \exp(2\gamma t) \qquad (98)$$

This expression shows that the growth rate is determined by the Wigner scattering rate γ independently of the classical and the self-scattering rates. The growth rate 2γ is equal to the L_1 norm of the Wigner potential.

In the pair generation method, the potential operator

$$\Theta_{\mathrm{w}}[f_{\mathrm{w}}](\mathbf{k}) = \int V^{+}(\mathbf{q})[f_{\mathrm{w}}(\mathbf{k}-\mathbf{q}) - f_{\mathrm{w}}(\mathbf{k}+\mathbf{q})]\,\mathrm{d}\mathbf{q} \qquad (99)$$

is interpreted as a generation term. It describes the creation of two new states, $\mathbf{k}-\mathbf{q}$ and $\mathbf{k}+\mathbf{q}$. The pair generation rate is equal to γ. When generating the second state, the sign of the statistical weight is changed. It should be noted that the Wigner equation strictly conserves mass, as can be seen by taking the zeroth-order moment of (45): $\partial n/\,\partial t + \mathrm{div}\ \mathbf{J} = 0$. Looking at the number of particles regardless of their statistical weights, that is, counting each particle as positive, would correspond to using the following potential operator.

$$\Theta^{*}_{w}[f_{\mathrm{w}}](\mathbf{k}) = \int V^{+}_{\mathrm{w}}(\mathbf{q})[f_{\mathrm{w}}(\mathbf{k}-\mathbf{q}) + f_{\mathrm{w}}(\mathbf{k}+\mathbf{q})]\,\mathrm{d}\mathbf{q} \quad (100)$$

Using (100), a continuity equation for numerical particles is obtained as $\partial n^*/\partial t + \mathrm{div}\ \mathbf{J}^* = 2\gamma(\mathbf{r})n^*$. Assuming a constant γ, the generation rate in this equation will give rise to an exponential increase in the number of numerical particles N^*.

$$N^*(t) = N^*(0)\exp(2\gamma t) \qquad (101)$$

This discussion shows that the appearance of an exponential growth rate is independent of the details of the particular Monte Carlo algorithm and must be considered to be a fundamental consequence of the non-positive kernel.

3.4 Particle annihilation

The discussed particle models are instable, because either the particle weight or the particle number grows exponentially in time. Using the Markov chain method, it has been demonstrated that tunnelling can be treated numerically by means of a particle model (Nedjalkov et al., 2002). However, because of the exponentially increasing particle weight at the very short timescale $(2\gamma)^{-1}$, application of this algorithm turned out to be restricted to single-barrier tunnelling and small barrier heights only. This method can be useful for devices where quantum effects are weak and the potential operator is a small correction to the otherwise classical transport equation.

A stable Monte Carlo algorithm can be obtained by combining one of the particle generation methods with a method to control the particle number. One can assume that two particles of opposite weight and a sufficiently small distance in phase-space annihilate each other. The reason is that the motions of both particles are governed by the same equation. Therefore, when they come close to each other at some time instant, the two particles have approximately the same initial condition and thus a common probabilistic future. In an ensemble Monte Carlo method, a particle removal step should be performed at given time steps. During the time step, the ensemble is allowed to grow to a certain limit, then particles are removed and the initial size of the ensemble is restored. In this work, the problem has been solved for the stationary transport problem. In the algorithm, the trajectory of only one sample particle is followed, whereas other numerical particles are temporarily stored on a phase-space grid. Due to the opposite sign, particle weights annihilate to a large extent in the cells of the grid. The total residual weight in each cell has to be minimised, as it represents a measure for the numerical error of the method (Kosina et al., 2003a,b).

4 Simulation results

The Monte Carlo method for solving the Wigner equation has been applied to different nanoelectronic devices. Since quantum ballistic transport can be described by both the collision-less Wigner equation and the Schrödinger equation, this case can be used for a consistent comparison. The equivalence of both approaches is demonstrated for the case of a double gate MOSFET. Effects of scattering are then discussed for resonant tunnelling structures.

4.1 Comparison of Wigner and Schrödinger equation-based results

To demonstrate the validity of the numerical method presented, quantum ballistic transport in a double gate MOSFET is studied. The test device has a gate length of 10 nm and a silicon film thicknesses of 3 nm. Metal gates with mid-gap work function are assumed. To obtain potential profiles for the quantum transport study a classical simulation using `Minimos-NT` (Institut für Mikroelektronik) is performed. To account for the effect of transverse confinement due to the small silicon film thickness, the first longitudinal subband ($m_1 = 0.91m_0$) is calculated in a post-processing step by solving the Schrödinger equation. Due to the computational requirements of the quantum

Monte Carlo simulation, the potential is not determined self-consistently with the quantum mechanical problem in this study. Figure 5 shows the resulting subband edges along the channel of the MOSFET at a drain voltage of 0.4 V and gate voltages of 0.2 and 0.4 V. Assuming a cutoff length of $\lambda_c = 60$ nm, the spectral decomposition described in Section 2.5 is applied to obtain the classical and the quantum mechanical potential profiles.

Figure 6 compares carrier concentrations for different gate voltages resulting from the quantum Monte Carlo and a classical ballistic Monte Carlo method. In the channel the quantum mechanical concentration is significantly higher than the classical concentration. The additional concentration originates from electrons able to tunnel through the energy barrier.

Figure 5 Potential profiles used in the ballistic simulations represent the lowest subband edges of a double-gate MOSFET of 10 nm gate length

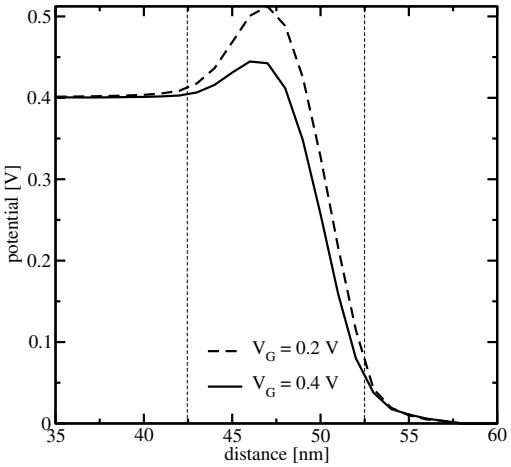
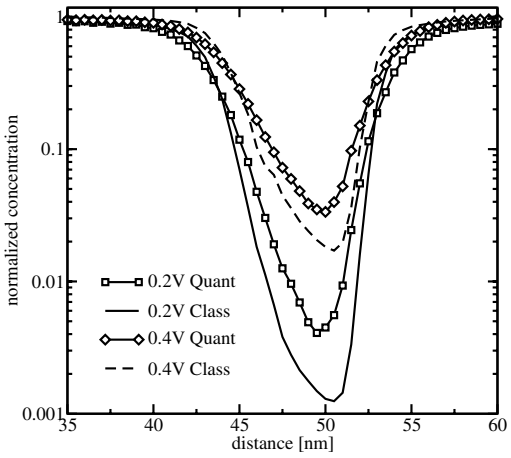
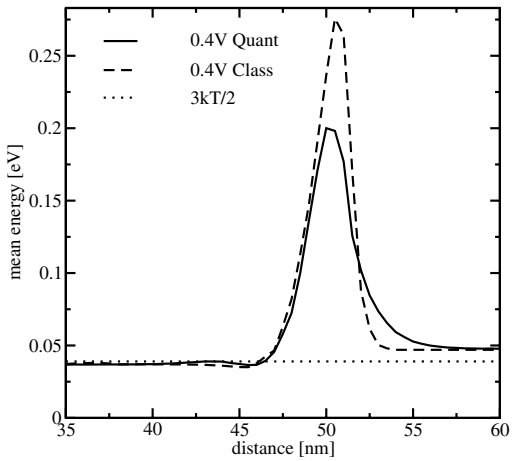


Figure 6 Carrier concentrations from classical ballistic and the quantum Monte Carlo simulations for the potential profiles shown in Figure 5



Mean energies are calculated as the second moments of the Wigner function and the classical ballistic distribution function, respectively (Figure 7). In the source region both the quantum mechanical and classical mean energies are very close to the equilibrium value of $3kT/2$. The difference in mean energy is most pronounced in the region of the potential barrier. Since tunnelling electrons formally possess a negative energy, it can be expected that in a tunnelling region the quantum mechanical mean energy will be lower than the classical one. In the drain region the average energy profiles decay to the same value, which is higher than the equilibrium value. Because both the classical and quantum transport models under consideration are ballistic, electrons reaching the drain region either by surmounting the potential barrier or tunnelling, acquire considerable kinetic energy due to the applied drain-source voltage. Since scattering is neglected, the high momenta of these electrons are preserved, which leads to a significant contribution to the mean energy throughout the drain area. The opposite effect, although much weaker, can be observed in the source region. Because of a lacking back flow of electrons above the energy barrier, the negative, high-momentum tail of the Maxwellian distribution is truncated, resulting in a mean energy slightly below the equilibrium value (Figure 7).

Figure 7 Mean electron energy at $V_G = 0.4$ V, computed as the second moment of the Wigner function and the classical ballistic distribution function



The ballistic current-voltage characteristics, as calculated by the Wigner and the classical Monte Carlo methods, are shown in Figure 8. Current is normalised with respect to the classical on-current. For all bias points, the quantum mechanical current is significantly higher than the classical thermionic current. The current increase is caused by carriers tunnelling through the potential barrier. To justify the correctness of the Wigner Monte Carlo simulations, the current is also determined from a solution of the Schrödinger equation with open boundary conditions for the same potential profile. Results of the Wigner Monte Carlo simulations are in good agreement with the solutions of the Schrödinger equation, as shown in Figure 8.

For a quantum ballistic transport problem, the numerical solution of the Schrödinger equation will always be computationally more efficient than the solution of the Wigner equation. The important advantage of the Wigner function-based method, however, is that dissipative processes due to scattering can be included.

4.2 The effect of scattering

Virtually all published results of Wigner function–based device modelling focus on resonant tunneling diodes (Sun

et al., 1998; Mizuta and Tanoue, 1995). In this section, two different devices are discussed. Their parameter values are collected in Table 1, where RTD1 (Shifren and Ferry, 2002) is a device from the literature. The semi-classical scattering model includes polar optical, acoustic deformation potential and ionised-impurity scattering. Parameter values for GaAs have been assumed.

Figure 8 Classical ballistic and quantum ballistic current, both obtained by Monte Carlo simulations (open symbols). Currents are normalised to the classical on-current at $V_G = 0.4V$. Solid lines show the analytically calculated thermionic emission current and the current from a numerical solution of the open boundary Schrödinger equation, respectively

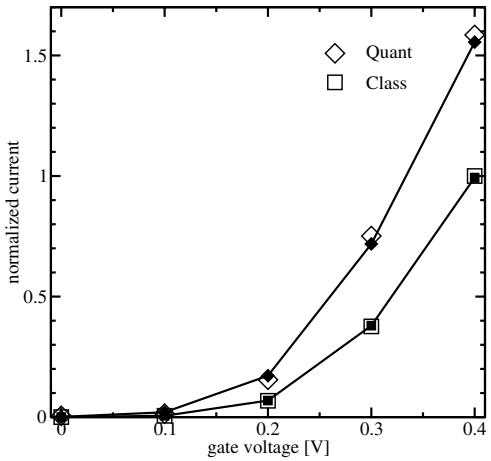


Table 1 Parameter values of the simulated resonant tunnelling diodes. The lattice constant of GaAs is $a_0 = 0.565$ nm

Device name	*Barrier height (eV)*	*Barrier width (nm)*	*Well width (nm)*	*Device length (nm)*	*Contact doping (cm^{-3})*
RTD1	0.3	3.0	5.0	200.0	10^{16}
RTD2	0.47	3.0	4.0	270.0	10^{18}

In RTD1, the potential changes linearly in a region of 40 nm length, starting 10 nm before the emitter barrier and extending 19 nm after the collector barrier, as shown in Figure 9. The Wigner potential is discretised using $N_k = 640$ equidistant k_x points and $\Delta x = 0.5$ nm spacing in x-direction. Assuming a cutoff length of $L_c = 80$ nm, one would require at least $N_k = L_c/\Delta x = 160$. This minium value is often used in finite-difference simulations for the Wigner equation, but in the Monte Carlo simulation we use the considerably larger value stated above in order to get a better resolution of the energy domain. The annihilation mesh is three-dimensional. In x-direction, the grid covers the region where the Wigner potential is non-zero. Because of the cylindrical symmetry of the Wigner function, only two momentum coordinates have to be considered. The mesh extends to an energy of 6 eV in both axial and radial k-direction.

Figure 9 Conduction band edge of RTD1 for different voltages. A linear voltage drop is assumed

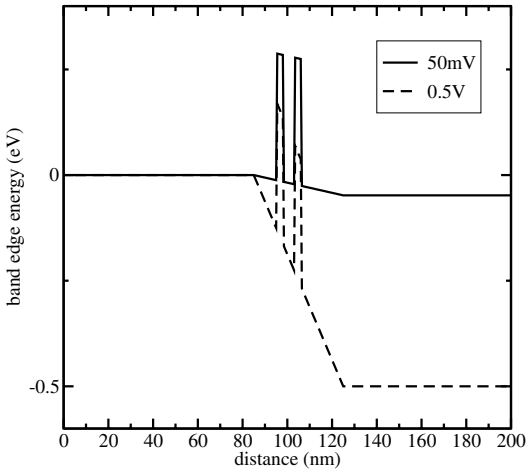
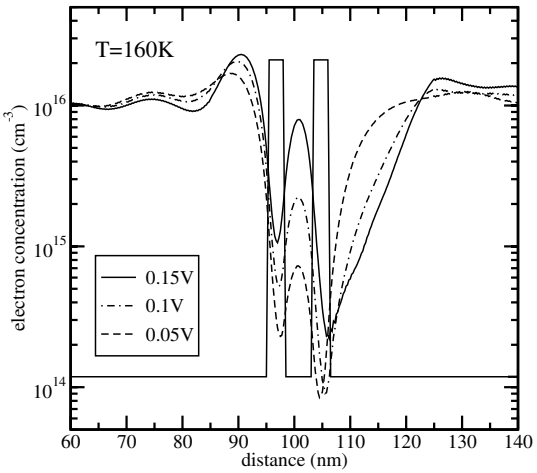


Figure 10 shows the electron concentration in RTD1 at voltages below the resonance voltage. Classical behaviour is observed before and after the double barrier, whereas in the quantum well the behaviour of the solution is non-classical. In front of the barrier an accumulation layer forms, with its maximum concentration increasing with the band bending. In the quantum well, the concentration increases as the resonance is approached. After the barrier a depletion layer forms, which grows with applied voltage. In this region, the concentration at 0.15 V varies exponentially in response to the linear potential (see Figure 9), which is again a classical property.

Figure 10 Electron concentration in RTD1 for voltages less than the resonance voltage



For voltages above the resonance voltage, the concentration in the well drops, whereas the depletion layer continues to grow (Figure 11). The mean kinetic energy of the electrons is depicted in Figure 12. The energy density has been calculated from the second-order moment of the Wigner function (35) and divided by the electron density to get the mean energy per electron. In the zero-field regions, an energy close to the equilibrium energy is obtained, which demonstrates that the energy conservation property of the Wigner potential operator

is also satisfied by the numerical Monte Carlo procedure. One has to keep in mind that the Wigner potential can produce a rather large momentum transfer. For the chosen value for Δx, the related energy transfer can reach values as large as 5 eV, which shows that a large degree of cancellation occurs in the estimator for the mean energy. Electrons injected from the second barrier into the collector space charge region show initially a high kinetic energy.

Figure 11 Electron concentration in RTD1 for voltages greater than the resonance voltage

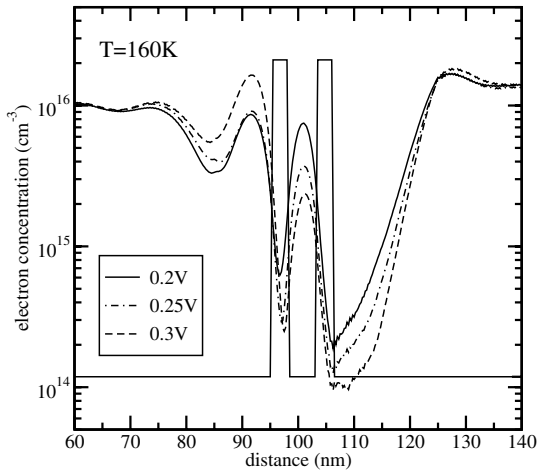
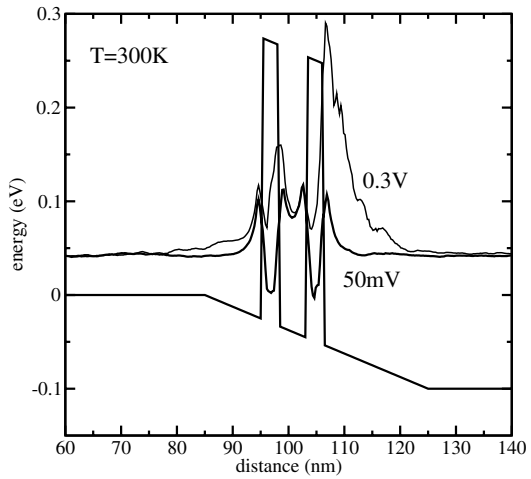


Figure 12 Mean kinetic energy in RTD1 for two different voltages



Phonon scattering strongly affects the current–voltage characteristic of RTD1 (Figure 13). As compared to the coherent case, phonon scattering leads to an increase in the valley current and a resonance voltage shift. The large difference in the valley current can be explained by the electron concentration in off-resonance condition (Figure 14). With phonon scattering included, a significantly higher concentration forms in the emitter notch and injection in the double barrier is increased. This indicates that a quasi bound state forms in the emitter notch. The population of this state increases when scattering is switched on. On the other hand, in resonance condition where the applied voltage is lower, such a bound state does not form and very similar electron concentrations are observed for the coherent and noncoherent case (Figure 15).

Figure 13 Influence of phonon scattering on the current–voltage characteristics of the RTD1

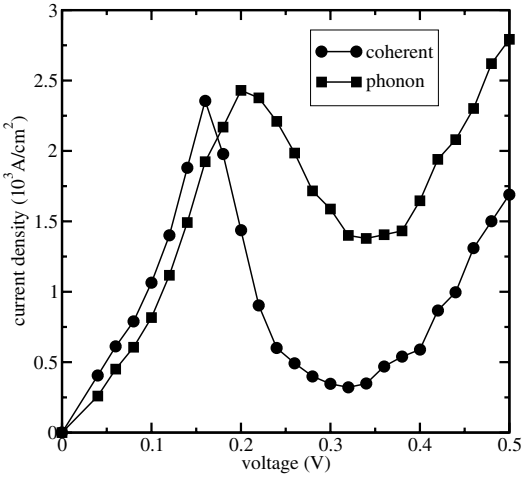
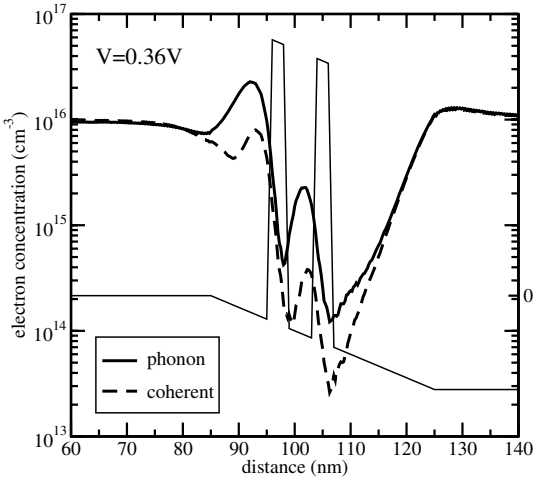


Figure 14 Electron concentration in RTD1 in off-resonance condition



4.3 *Inclusion of extended contact regions*

Since the Wigner equation simplifies to the Boltzmann equation when the potential variation is sufficiently smooth, the proposed quantum Monte Carlo method turns into the semiclassical Monte Carlo method for vanishing Wigner potential. Therefore, one can simulate a quantum region embedded in an extended classical region with the interface between the regions correctly treated in an implicit way. By means of the Wigner generation rate γ, the simulation domain can be decomposed into quantum regions ($\gamma > 0$) and classical regions ($\gamma \simeq 0$). In Figure 16, these regions within RTD1 are marked. The electron concentration and the mean energy are smooth in the extended contact regions and not affected by the strong onset of the Wigner generation rate.

In the simulation of RTD2, the Wigner potential $V_{\mathrm{w}}^{+}(k_x, x)$ is discretised using $N_k = 1200$ equidistant k_x points

and $\Delta x = 0.5\,\text{nm}$ spacing in the x-direction. A cutoff length of $L_c = 60\,\text{nm}$ is assumed. The annihilation mesh consists of 480 points in the longitudinal and 120 points in the perpendicular momentum direction and the real space coordinate is discretised using $\Delta x = 0.5\,\text{nm}$. The electrostatic potential has been computed using the self-consistent Schrödinger-Poisson solver NANOTCAD-1D (Iannaccone et al., 2001). Figure 17 shows the electron concentration profile in the device. At the resonance voltage of 1.2 V, the concentration in the quantum well is considerably higher than in the off-resonance condition at 1.6 V. The concentration in the depletion region left of the barrier depends on the injected current and is thus correlated with the concentration in the well.

Figure 15 Electron concentration in RTD1 in resonance condition

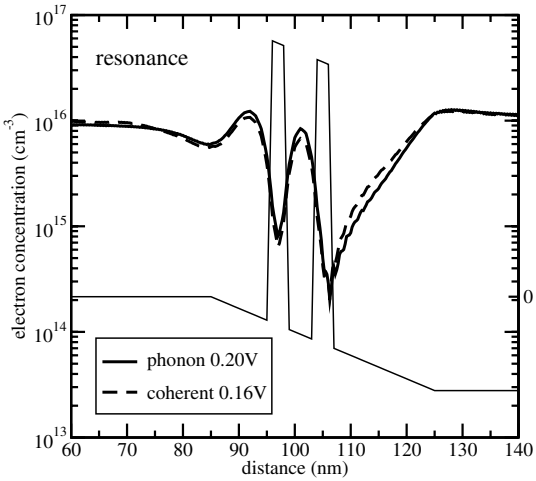


Figure 16 Electron concentration and mean electron energy in RTD1 at $T = 300\,\text{K}$ and 0.1 V applied voltage

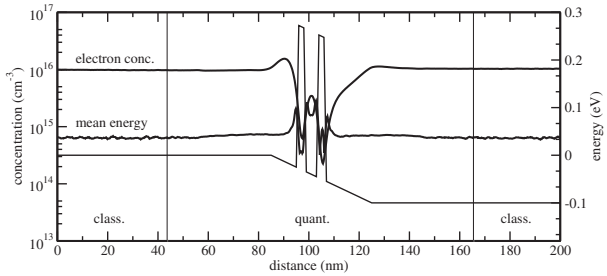
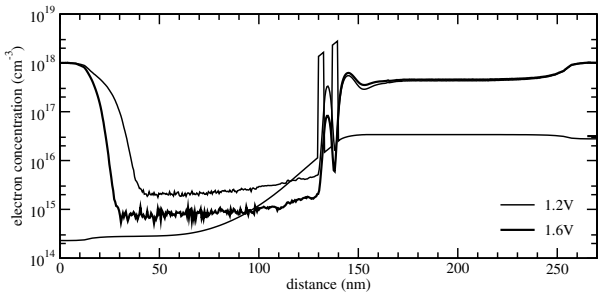


Figure 17 Electron concentration profiles in RTD2



5 Conclusion

The examples presented in Section 4 demonstrate that a numerical solver for the Wigner equation can provide quantitatively correct results. One requirement is that the cutoff length is chosen sufficiently large. The completeness relation of the discrete Fourier transform reflecting Heisenberg's uncertainty principle, $\Delta k_x = \pi/L_c$, shows that a small L_c will result in a coarse grid in momentum space, and resonance peaks in the transmission coefficient might not be resolved properly. In the past, the Wigner equation has been solved most frequently by finite-difference methods. Due to the non-locality of the potential operator, all points in momentum space are coupled, resulting in a very poor sparsity pattern of the matrix. Therefore, increasing the number of grid points in k-space, related to the cutoff length by $N_k = L_c/\Delta x$, is limited by prohibitive memory and computation time requirements. This might be one reason why quantitatively correct solutions were difficult to obtain in the past. We believe that the frequently reported accuracy problems with finite-difference Wigner function-based device simulations result from a too coarse k-space discretisation. As this problem occurs already for one-dimensional geometries, higher dimensional simulations using the finite-difference method are probably out of reach. It is interesting to note that Frensley, who pioneered the finite-difference method for the Wigner equation (Frensley, 1990), later abandoned this method and developed the quantum-transmitting boundary method to describe coherent transport in open systems (Frensley, 1992).

The Monte Carlo method allows the number of k-points to be increased. In this work, the Wigner potential has been discretised using N_k of the order 10^3. However, high-performance resonant tunnelling diodes with very high peak-to-valley current ratio pose still a problem for the Monte Carlo method. In such a device, the density can vary over several orders of magnitude, which often cannot be resolved by the Monte Carlo method. This problem is also well-known from the classical Monte Carlo method. As a solution, one could apply statistical enhancement techniques in such cases. At present, an equidistant k-grid is used for the discretisation of the Wigner potential. Because the transmission coefficient of double-barrier structures may show very narrow resonance peaks, using an equidistant k-grid may not be the optimal choice. However, because of the discrete Fourier transform of the potential involved in the computation of the Wigner potential, the use of a non-equidistant k-grid appears to be problematic.

In a Wigner function-based simulation of one-dimensional heterostructures, fundamental simulation parameters such as the cutoff length are closely linked to physical device parameters such as the spacing from the contacts. This property stems from the choice of plane-wave basis sets in a quantum mechanical regime of broken translational symmetry. Although analytically appealing, this basis set can cause numerical difficulties. Other approaches such as the non-equilibrium Green's function formalism may have the advantage that other basis sets can be used more straightforwardly.

These considerations indicate that from a numerical point of view, the Wigner function formalism might not be the optimal choice for resonant tunnelling simulation. However, because the formalism describes quantum effects and scattering effects with equal accuracy, it appears well suited especially when a quasi-ballistic transport condition without energetically sharp resonances is present. One strength of the Wigner function approach is the treatment of contact regions. Non-equilibrium transport can be simulated in the whole device formed by a central quantum region embedded in extended classical regions. The presented Wigner Monte Carlo method can bridge the gap between classical device simulation and pure quantum ballistic simulations.

Development of Monte Carlo methods for the solution of the Wigner equation is still in the beginning. Research efforts are needed especially with respect to the negative sign problem. The particle generation-annihilation algorithm developed by the authors is just one solution to that problem. Improved variants of this algorithm or even new solution strategies are yet to be devised. Extension of the Monte Carlo methods to higher dimensional device geometries, however, is straightforward.

References

Bertoni, A., Bordone, P., Brunetti, R. and Jacoboni, C. (1999) 'The Wigner function for electron transport in mesoscopic systems', *Journal of Physics: Condensed Matter*, Vol. 11, pp.5999–6012.

Biegel, B. and Plummer, J. (1996) 'Comparison of self-consistency iteration options for the Wigner function method of quantum device simulation', *Physical Review B*, Vol. 54, No. 11, pp.8070–8082.

Biegel, B. and Plummer, J. (1997) 'Applied bias slewing in transient Wigner function simulation of resonant tunneling diodes', *IEEE Transactions on Electron Devices*, Vol. 44, No. 5, pp.733–737.

Bordone, P., Pascoli, M., Brunetti, R., Bertoni, A. and Jacoboni, C. (1999) 'Quantum transport of electrons in open nanostructures with the Wigner-function formalism', *Physical Review B*, Vol. 59, No. 4, pp.3060–3069.

Bordone, P., Bertoni, A., Brunetti, R. and Jacoboni, C. (2003) 'Monte Carlo simulation of quantum electron transport based on Wigner paths', *Mathematics and Computers in Simulation*, Vol. 62, pp.307–314.

Brunetti, R., Bertoni, A., Bordone, P. and Jacoboni, C. (2000) 'Dynamical equation and Monte Carlo simulation of the two-time Wigner function for electron quantum transport', J. Barker and J. Watling (Eds). *International Workshop on Computational Electroncis*, Glasgow, University of Glasgow, pp.131–132.

Brunetti, R., Jacoboni, C. and Rossi, F. (1989) 'Quantum theory of transient transport in semiconductors: a Monte Carlo approach', *Physical Review B*, Vol. 39, No. 15, pp.10781–10790.

Buot, F. and Jensen, K. (1990) 'Lattice Weyl-Wigner formulation of exact many-body quantum-transport theory and applications to novel solid-state quantum-based devices', *Physical Review B*, Vol. 42, No. 15, pp.9429–9457.

Buot, F.A. and Jensen, K.L. (1991) 'Intrinsic high-frequency oscillations and equivalent circuit model in the negative differential resistance region of resonant tunneling devices', *COMPEL*, Vol. 10, No. 4, pp.241–253.

Bufler, F. and Schlösser, J. (1994) 'Two-band density matrix theory: the effect of external fields on transition and transport processes', *Journal of Physics: Condensed Matter*, Vol. 6, pp.7445–7460.

Byron, F. and Fuller, R. (1992) *Mathematics of Classical and Quantum Physics*, New York: Dover.

Carruthers, P. and Zachariasen, F. (1983) 'Quantum collision theory with phase-space distributions', *Reviews of Modern Physics*, Vol. 55, No. 1, pp.245–285.

Croitoru, M., Gladilin, V., Fomin, V., Devreese, J., Magnus, W.S.W. and Soree, B. (2003) 'Quantum transport in nanosize silicon-on-insulator metal-oxide-semidonductor field-effect-transistors', *Journal of Applied Physics*, Vol. 93, No. 2, pp.1230–1240.

Croitoru, M., Gladilin, V., Fomin, V., Devreese, J., Magnus, W.S.W. and Soree, B. (2004) 'Quantum transport in nanosize double-gate metal-oxide-semidonductor field-effect-transistor', *Journal of Applied Physics*, Vol. 96, No. 4, pp.2305–2310.

Cutright, T. and Zachos, C. (2001) 'Negative probability and uncertainty relations', *Modern Physics Letters A*, Vol. 16, No. 37, pp.2381–2385.

Degond, P. and Ringhofer, C. (2003) 'Quantum moment hydrodynamics and entropy principle', *Journal of Statistical Physics*, Vol. 112, No. 3, pp.587–628.

Demaio, L., Barletti, L., Bertoni, A., Bordone, P. and Jacoboni, C. (2002) 'Wigner-function approach to multiband transport in semiconductors', *Physica B*, Vol. 314, pp.104–107.

Ermakow, S. (1975) *Die Monte-Carlo-Methode und verwandte Fragen*, München, Wien: R. Oldenburg Verlag.

Frensley, W. (1986a) 'Transient response of a tunneling device obtained from the Wigner function', *Physical Review Letters*, Vol. 57, No. 22, pp.2853–2856.

Frensley, W. (1986b) 'Quantum transport simulation of the resonant tunneling diode', *International Electron Devices Meeting*, Los Angeles, The Institute of Electrical and Electronics Engineers, Inc., pp.571–574.

Frensley, W. (1987) 'Wigner-function model of a resonant-tunneling semiconductor device', *Physical Review B*, Vol. 36, No. 3, pp.1570–1580.

Frensley, W. (1989) 'Effect of inelastic processes on the self-consistent potential in the resonant-tunneling diode', *Solid-State Electronics*, Vol. 32, No. 12, pp.1235–1239.

Frensley, W. (1990) 'Boundary conditions for open quantum systems driven far from equilibrium', *Reviews of Modern Physics*, Vol. 62, No. 3, pp.745–789.

Ferry, D. and Goodnick, S. (2001d) *Transport in Nanostructures*, Cambridge University Press.

Fannjiang, A., Jin, S. and Papanicolaou, G. (2002) 'High frequency behavior of the focusing nonlinear Schrödinger equation with random inhomogeneities', *SIAM Journal of Applied Mathematics*, Vol. 63, No. 4, pp.1328–1358.

Frensley, W. (1992) 'Numerical evaluation of resonant states', *Superlattices and Microstructures*, Vol. 11, No. 3, pp.347–350.

Gullapalli, K., Miller, D. and Neikirk, D. (1994) 'Simulation of quantum transport in memory-switching double-barrier quantum-well diodes', *Physical Review B*, Vol. 49, No. 4, pp.2622–2628.

Gardner, C. (1994) 'The quantum hydrodynamic model for semiconductor devices', *SIAM Journal of Applied Mathematics*, Vol. 54, pp.409–427.

Gardner, C.L. and Ringhofer, C. (1996) 'Smooth quantum potential for the hydrodynamic model', *Physical Review E*, Vol. 53, pp.157–167.

Goldsman, N., Lin, C-K., Han, Z. and Huang, C-K. (2000) 'Advances in the spherical harmonic-Boltzmann-Wigner approach to device simulation', *Superlattices and Microstructures*, Vol. 27, Nos. 2/3, pp.159–175.

Gurov, T., Nedjalkov, M., Whitlock, P.,Kosina, H. and Selberherr, S. (2002) 'Femtosecond relaxation of hot electrons by phonon emission in presence of electric field', *Physica B*, Vol. 314, pp.301–304.

Gehring, A. and Kosina, H. (2004) 'Wigner-function based simulation of classic and ballistic transport in scaled DG-MOSFETs using the Monte Carlo method', *International Workshop on Computational Electronics Abstracts*, (West Lafayette), pp.227–228.

Holland P. and Kypriandis, K. (1986) 'Relativistic generalization of the Wigner function and its interpretation in the causal stochastic formulation of quantum mechanics', *Physical Review A*, Vol. 33, No. 6, pp.4380–4383.

Hänsch, W. (1991) *The Drift-Diffusion Equation and Its Applications in MOSFET Modeling*, Computational Microelectronics, Wien New York: Springer Verlag.

Han, Z., Goldsman, N. and Lin, C. (2000) '2-D quantum transport device modeling by self-consistent solution of the Wigner and Poisson equations', in *International Electron Devices Meeting*, (San Francisco), The Institute of Electrical and Electronics Engineers, Inc., pp.62–65.

Hammersley, J. and Handscomb, D. (1964) *Monte Carlo Methods*. New York: John Wiley.

Institut für Mikroelektronik, Technische Universität Wien, Austria, *MINIMOS-NT 2.1 User's Guide*, 2004.

Iannaccone, G., Macucci, M., Coli, P., Curatola, G., Fiori, G., Gattobigio, M. and Pala, M. (2001) 'Towards nanotechnology computer aided design: the NANOTCAD project', *IEEE Conference on Nanotechnology*, pp.117–122.

Jacoboni, C., Brunetti, R., Bordone, P. and Bertoni, A. (2001) 'Quantum transport and its simulation with the Wigner-function approach', *International Journal of High Speed Electronics and Systems*, Vol. 11, No. 2, pp.387–423.

Jacoboni, C. and Reggiani, L. (1983) 'The Monte Carlo method for the solution of charge transport in semiconductors with applications to covalent materials', *Reviews of Modern Physics*, Vol. 55, No. 3, pp.645–705.

Jensen, K. and Buot, F. (1991) 'Numerical simulation of intrinsic bistability and high-frequency current oscillations in resonant tunneling structures', *Physical Review Letters*, Vol. 66, No. 8, pp.1078–1081.

Kluksdahl, N., Kriman, A., Ferry, D. and Ringhofer, C. (1989) 'Self-consistent study of the resonant-tunneling diode', *Physical Review B*, Vol. 39, No. 11, pp.7720–7735.

Kluksdahl, N., Pötz, W., Ravaioli, U. and Ferry, D. (1987) 'Wigner function study of a double quantum barrier resonant tunneling diode', *Superlattices and Microstructures*, Vol. 3, No. 1, pp.41–45.

Kosina, H., Nedjalkov, M. and Selberherr, S. (2003) 'Quantum Monte Carlo simulation of a resonant tunneling diode including phonon scattering', in M. Laudon and B. Romanowicz, (Eds). *Nanotech*, San Francisco, Computational Publications, pp.190–193.

Kosina, H., Nedjalkov, M. and Selberherr, S. (2003) 'A Monte Carlo method seamlessly linking classical and quantum transport calculations', *Journal of Computational Electronics*, Vol. 2, Nos. 2-4, pp.147–151.

Kosik, R. (2004) 'Numerical challenges on the road to NanoTCAD', Dissertation, Vienna University of Technology.

Kosina, H., Nedjalkov, M. and Selberherr, S. (2000) 'Theory of the Monte Carlo method for semiconductor device simulation', *IEEE Transactions on Electron Devices*, Vol. 47, No. 10, pp.1898–1908.

Kosina, H. and Nedjalkov, M. (2003) 'Particle models for device simulations', *International Journal of High Speed Electronics and Systems*, Vol. 13, No. 3, pp.727–769 (invited).

Kosina, H., Nedjalkov, M. and Selberherr, S. (2003) 'The stationary Monte Carlo method for device simulation - part I: theory', *Journal of Applied Physics*, Vol. 93, No. 6, pp.3553–3563.

Levinson, I. (1970) 'Translational invariance in uniform fields and the equation for the density matrix in the Wigner representation', *Soviet Physics JETP*, Vol. 30, No. 2, pp.362–367.

Levanda, M. and Fleurov, V. (2001) 'A Wigner quasi-distribution function for charged particles in classical electromagnetic fields', *Annals of Physics*, Vol. 292, pp.199–231.

Mahan, G. (1983) *Many-particle physics*, New York and London: Plenum Press.

Mains, R. and Haddad, G. (1988) 'Wigner function modeling of resonant tunneling diodes with high peak-to-valley ratios', *Journal of Applied Physics*, Vol. 64, No. 10, pp.5041–5044.

Mains, R. and Haddad, G. (1994) 'An accurate re-formulation of the Wigner function method for quantum transport modeling', *Journal of Computer and Physises*, Vol. 112, pp.149–161.

Miller, D. and Neikirk, D. (1991) 'Simulation of intervalley mixing in double-barrier diodes using the lattice Wigner function', *Applied Physics Letters*, Vol. 58, No. 24, pp.2803–2805.

Mizuta, H. and Tanoue, T. (1995) *The Physics and Applications of Resonant Tunneling Diodes*, Cambridge University Press.

Nedjalkov, M., Kosik, R., Kosina, H. and Selberherr, S. (2002) 'Wigner transport through tunneling structures - scattering interpretation of the potential operator', *Proceedings International Conference on Simulation of Semiconductor Processes and Devices*, (Kobe, Japan), Publication Office Business Center for Academic Societies Japan, pp. 187–190.

Nedjalkov, M., Dimov, I., Rossi, F. and Jacoboni,C. (1996) 'Convergency of the Monte Carlo algorithm for the Wigner quantum transport equation', *Journal of Mathematical and Computer Modelling*, Vol. 23, Nos. 8/9, pp.159–166.

Nedjalkov, M., Kosik, R., Kosina, H. and Selberherr, S. (2002) 'A Wigner equation with quantum electron-phonon interaction', *Microelectronic Engineering*, Vol. 63, Nos. 1–3, pp.199–203.

Nedjalkov, M., Kosina, H., Selberherr, S. and Dimov, I. (2001) 'A backward Monte Carlo method for simulation of the electron quantum kinetics in semiconductors', *VLSI Design*, Vol. 13, Nos. 1–4, pp.405–411.

Nedjalkov, M., Kosina, H. and Selberherr, S. (2003) 'The stationary Monte Carlo method for device simulation - part II: event biasing and variance estimation', *Journal of Applied Physics*, Vol. 93, No. 6, pp.3564–3571.

Pascoli, M., Bordone, P., Brunetti, R. and Jacoboni, C. (1998) 'Wigner paths for electrons interacting with phonons', *Physical Review B*, Vol. 58, No. 7, pp.3503–3506.

Rammer, J. (1991) 'Quantum transport theory of electrons in solids: a single-particle approach', *Reviews of Modern Physics*, Vol. 63, No. 4, pp.781–817.

Ravaioli, U., Osman, M., Pötz, W., Kluksdahl, N. and Ferry, D. (1985) 'Investigation of ballistic transport through resonant-tunneling quantum wells using Wigner function approach', *Physica B*, Vol. 134, pp.36–40.

Rossi, F., Jacoboni, C. and Nedjalkov, M. (1994) 'A Monte Carlo solution of the Wigner transport equation', *Semiconductor Science and Technology*, Vol. 9, pp.934–936.

Ringhofer, C., Nedjalkov, M., Kosina, H. and Selberherr, S. (2004) 'Semi-classical approximation of electron-phonon scattering beyond Fermi's golden rule', *SIAM Journal of Applied Mathematics*, Vol. 64, No. 6, pp.1933–953.

Shifren, L. and Ferry, D.K. (2002) 'Wigner Function Quantum Monte Carlo', *Physical B*, Vol. 314, pp.72–75.

Shifren, L., Ringhofer, C. and Ferry, D. (2003) 'Inclusion of nonlocal scattering in quantum transport', *Physics Letters A*, Vol. 306, pp.332–336.

Shih, J-J., Huang, H-C. and Wu, G. (1994) 'Effect of mass discontinuity in the Wigner theory of resonant-tunneling diodes', *Physical Review B*, Vol. 50, No. 4, pp.2399–2405.

Sonego, S. (1991) 'Quasiprobabilities and explicitly covariant relativistic quantum theory', *Physical Review A*, Vol. 44, No. 9, pp.5369–5382.

Sobol, I. (1984) *The Monte Carlo Method*, Moscow: Mir Publishers.

Sun, J., Haddad, G., Mazumder, P. and Schulman, J. (1998) 'Resonant tunneling diodes: models and properties', *Proceedings of the IEEE*, Vol. 86, No. 4, pp.641–661.

Tatarskii, V. (1983) 'The Wigner representation of quantum mechanics', *Soviet Physics Uspekhi*, Vol. 26, No. 4, pp.311–327.

Tsuchiya, H., Ogawa, M. and Miyoshi, T. (1991) 'Simulation of quantum transport in quantum devices with spatially varying effective mass', *IEEE Transactions on Electron Devices*, Vol. 38, No. 6, pp.1246–1252.

Tsuchiya, H. and Ravaioli, U. (2001) 'Particle Monte Carlo simulation of quantum phenomena in semiconductor nanostructures', *Journal of Applied Physics*, Vol. 89, No. 7, pp.4023–4029.

Unlu, M.B., Rosen, B., Cui, H-L. and Zhao, P. (2004) 'Multi-band Wigner function formulation of quantum transport', *Physics Letters A*, Vol. 327, pp.230–240.

Venugopal, R., Ren, Z., Datta, S., Lundstrom, M.S. and Jovanovic, D. (2002) 'Simulation of quantum transport in nanoscale transistors: real versus mode-space approach', *Journal of Applied Physics*, Vol. 92, No. 7, pp.3730–3739.

Venugopal, R., Goasguen, S., Datta, S. and Lundstrom, M.S. (2004) 'Quantum mechanical analysis of channel access geometry and series resistance in nanoscale transistors', *Journal of Applied Physics*, Vol. 95, No. 1, pp.292–305.

Wigner, E. (1932) 'On the quantum correction for thermodynamic equilibrium', *Physical Review*, Vol. 40, pp.749–759.

Woolard, D., Zhao, P. and Cui, H. (2002) 'THz-frequency intrinsic oscillations in double-barrier quantum well systems', *Physica B*, Vol. 314, pp.108–112.

Wu, G. and Wu, K-P. (1992) 'Electron transport in a resonant-tunneling diode under the effect of a transverse magnetic field: a quantum theory in the Wigner formalism,' *Journal of Applied Physics*, Vol. 71, No. 3, pp.1259–1264.

Wlodarz, J.J. (1999) 'On marignalization of phase-space distribution functions', *Physics Letters A*, Vol. 264, pp.18–21.

Zachos, C. (2002) 'Deformation quantization: quantum mechanics lives and works in phase-space', *International Journal of Modern Physics A*, Vol. 17, No. 3, pp.297–316.

Zhou, J. and Ferry, D. (1992) 'Simulation of ultra-small GaAs MESFET using quantum moment equations', *IEEE Transactions on Electron Devices*, Vol. 39, No. 3, pp.473–478.