

# Theory of Electron Transport in Small Semiconductor Devices Using the Pauli Master Equation

M. V. FISCHETTI\*

*IBM Research Division, Thomas J. Watson Research Center, P. O. Box 218, Yorktown Heights, NY 10598, USA*

It is argued that the Pauli master equation can be used to simulate electron transport in very small electronic devices under steady-state conditions. Written in a basis of suitable wavefunctions and with the appropriate open boundary conditions, this equation removes some of the approximations which render the Boltzmann equation unsatisfactory at small length-scales. The main problems consist in describing the interaction of the system with the reservoirs and in assessing the range of validity of the equation: Only devices smaller than the size of the electron wavepackets injected from the contacts can be handled. Two one-dimensional examples solved by a simple Monte Carlo technique are presented.

*Keywords:* Master equation, electron transport, small devices

## 1. INTRODUCTION

As noted by Frensky discussing quantum transport in open systems [1], the Pauli master equation [2], (PME) could constitute an intuitive description of electron transport in semiconductor devices of size comparable to the electron wavelength: One could capture the wavelike nature of transport, lost in the Boltzmann-transport-equation (BTE) picture, so bypassing the weak-field, long devices limitations. In retaining the weak-scattering and completed-collision limits one would not reach a full-fledged quantum description, as required to study very fast time transients [3, 4]. Yet, as devices approach the sub-50 nm length-scale, the

advantage of avoiding the concept of point-like electrons appears overwhelming.

Although arguments have been raised against the correctness of the PME [1], it is applicable to very *small devices* and to *steady-state* phenomena: *Small devices*, because the PME rests on the absence of off-diagonal elements of the density matrix  $\rho$ . In turn, their absence implies that the contacts inject 'plane waves', which corresponds physically to the injection of spatially highly delocalized wavepackets. Hence, the device must be smaller than the 'dephasing length'  $\lambda_\phi$  in the contacts [5]. For heavily doped Si, this is of the order of 50 nm, which is of technological interest. *Steady state*, since any nontrivial time transient

---

\*Corresponding author.

(e.g., relaxation of photo-excited carriers, fast switching of a device) would excite those off-diagonal terms of  $\rho$  which the PME cannot handle. Indeed, as also note by Frenley [1], the PME is inconsistent with current continuity in this case.

## 2. MASTER EQUATION AND CONTACTS

Let's consider a semiconductor region unbounded in the  $(x, y)$ -plane, homogeneous in this plane, in contact with two reservoirs at  $z=0$  and  $z=L$ . Let's write the single-particle Hamiltonian as:

$$-\frac{\hbar^2 \nabla^2}{2m^*} - eV(z) + H_{\text{int}} + H_{\text{res}} = H_0 + H_{\text{int}} + H_{\text{res}}, \quad (1)$$

within the envelope/effective-mass approximation,  $m^*$  being the effective mass.  $H_{\text{int}}$  describes the electron-phonon interaction or interactions with impurities, etc.  $H_{\text{res}}$  represents the interaction between the device and the external reservoirs. The time evolution of the system is given by the 'transport' equation for the 'reduced' (electronic) density matrix  $\rho$ :

$$\frac{\partial \rho}{\partial t} = \frac{i}{\hbar} [\rho, H] + \left( \frac{\partial \rho}{\partial t} \right)_{\text{res}}. \quad (2)$$

The last term on the right-hand-side describes the effect of the reservoirs. Assuming  $H_{\text{int}} = \alpha H'$ , we can derive a transport equation from Eq. (2), by considering only terms to the leading order in  $\alpha$ , and following the 'standard' derivation of Kohn and Luttinger [7], but now using as basis functions not plane waves, but the eigenstates  $|\mu \mathbf{K}\rangle$  of  $H_0$ , such that  $H_0 |\mu \mathbf{K}\rangle = [E_\mu + \hbar^2 \mathbf{K}^2 / (2m^*)] |\mu \mathbf{K}\rangle = E_\mu(\mathbf{K}) |\mu \mathbf{K}\rangle$ . This simplifies Eq. (2), since the commutator now contains only the interaction term  $H_{\text{int}}$ . The wave-functions are simply of the form  $\zeta_\mu(z) \exp(i \mathbf{K} \cdot \mathbf{R}) / (2\pi)$ , where  $\mathbf{K}$  and  $\mathbf{R}$  are the wavevector and position on the  $(x, y)$ -plane, and the functions  $\zeta_\mu(z)$  obey the one-dimensional

Schrödinger equation for the potential  $V(z)$ , subject to the boundary conditions  $V(0) = V_L$  and  $V(L) = V_R$ . Taking  $-eV_L > -eV_R$ , for any given energy  $E_\mu > -eV_L$ , two independent solutions represent waves incident from the left (+) or right (-), and partially reflected and transmitted, with reflection and transmission amplitudes  $r^\pm$  and  $t^\pm$ . For  $-eV_L > E_\mu > -eV_R$ , the nondegenerate solutions represent non-current-carrying states. In the following the index  $\sigma = \pm$  will label left- and right-traveling waves in the energy range  $E_\mu > eV_L$ , and  $\langle z | \mu \sigma \rangle = \zeta_\mu^\sigma(z)$ .

To the leading order in  $\alpha$ , the diagonal elements,  $\rho_{\mu \mathbf{K} \sigma}$  (symbol used in place of  $\rho_{\mu \mathbf{K} \sigma \mu \mathbf{K} \sigma}$ ) of the density matrix obey the Eqs. [7–9]:

$$\frac{\partial \rho_{\mu \mathbf{K} \sigma}}{\partial t} = \sum_{\nu \mathbf{K}' \sigma' \neq \mu \mathbf{K} \sigma} [W_{\mu \mathbf{K} \sigma; \nu \mathbf{K}' \sigma'} \rho_{\nu \mathbf{K}' \sigma'} - W_{\nu \mathbf{K}' \sigma'; \mu \mathbf{K} \sigma} \rho_{\mu \mathbf{K} \sigma}] + \left( \frac{\partial \rho_{\mu \mathbf{K} \sigma}}{\partial t} \right)_{\text{res}}. \quad (3)$$

here  $W_{\nu \mathbf{K}' \sigma'; \mu \mathbf{K} \sigma}$  represents the probability per unit time of a transition from the state  $|\mu \mathbf{K} \sigma\rangle$  to the state  $|\nu \mathbf{K}' \sigma'\rangle$  using Fermi golden rule. While the derivations provided in [7–9] can be followed very easily, now one must also assume that the reservoirs do not alter the ratio  $\alpha$  between off-diagonal and diagonal matrix elements. Van Hove [8] has shown the validity of Eq. (3) in describing approach to equilibrium, in the case of closed systems (for which the term  $(\partial \rho / \partial t)_{\text{res}}$  is absent), provided the initial state is chosen as 'quasi-diagonal': At  $t=0$  we must have  $\rho_{\mu \mathbf{K} \sigma \nu \mathbf{K}' \sigma'} = 0$  for  $E_\mu(\mathbf{K}) - E_\nu(\mathbf{K}') > \delta$ , where  $\delta/\hbar$  is of the order of the relaxation rate,  $1/\tau$ . Similarly, Eq. (3) is valid as long as the contacts do not 'inject' off-diagonal elements, that is  $(\partial \rho_{\mu \mathbf{K} \sigma \nu \mathbf{K}' \sigma'} / \partial t)_{\text{res}} = 0$  for  $E_\mu(\mathbf{K}) - E_\nu(\mathbf{K}') > \delta$ . This is the case for very delocalized wavepackets entering the contact, plane waves being the ideal limiting case. In this limit, the effect of the contacts can be modeled phenomenologically by assuming that the left reservoir attempts to restore charge neutrality by injecting a flux of

particles of any  $\mathbf{K}$  into the right-bound component of the states associated to waves incident from the left,  $\zeta_{\mu}^{+}$ , fixed by the Fermi level at the cathode,  $E_F^L$ . Accounting also for the flux out of the device via either contact, one can write:

$$\left(\frac{\partial \rho_{\mu+}}{\partial t}\right)_{\text{res}} = [f^{(L)}(E_{\mu}) - \rho_{\mu+} - T_{\mu}^{-} \rho_{\mu-}] \frac{\hbar k_{\mu}^{-}}{m^{*}} |A_{\mu}^{+}|^2. \quad (4)$$

where  $\rho_{\mu\sigma} = \sum_{\mathbf{K}} \rho_{\mu\mathbf{K}\sigma}$ ,  $T_{\mu}^{\mp} = |A_{\mu}^{\mp} t_{\mu}^{\mp}|^2 / |A_{\mu}^{\pm}|^2$ ,  $\hbar k_{\mu}^{\pm} = [2m^{*}(E_{\mu} + eV_{R/L})]^{1/2}$ , and the  $A$ 's are normalization constants. Similar expressions also hold for states traveling from the right. The occupation of the states at thermal equilibrium, as determined by the 'left' reservoir, has been expressed in terms of the equilibrium Fermi function (integrated over the 'transverse' wavevectors  $\mathbf{K}$ ),  $f^{(th,L)}(E_{\mu})$ , possibly shifted in  $k$ -space, in order to account for the open nature of the contacts [1, 6]. By analogy, a similar expression (with the terms  $\rho_{\mu+} - T_{\mu}^{-} \rho_{\mu-}$  replaced by  $\rho_{\mu}$ ) is also used to express the linear relaxation of the of states with energy  $-eV_R < E_{\mu} < -eV_L$ . Finally, the net current flowing through the device is given by:

$$J = -e \sum_{\mu} \left\{ \rho_{\mu+} \frac{\hbar k_{\mu}^{+}}{m^{*}} |t_{\mu}^{+} A_{\mu}^{+}|^2 - \rho_{\mu-} \frac{\hbar k_{\mu}^{-}}{m^{*}} |t_{\mu}^{-} A_{\mu}^{-}|^2 \right\}, \quad (5)$$

where obviously only the doubly-degenerate states with  $E_{\mu} > -eV_L$  contribute to the sum.

Criticisms have been raised against Eq. (3) based on the violation of continuity [1]. Yet, it should be noted that at steady state the reservoirs and the reverse scattering-induced transitions provide the 'missing' current and current continuity is trivially satisfied. However, continuity is indeed violated when using the PME to describe 'fast' time-dependent phenomena: For example, any time dependent perturbation  $\delta V(t)$  will cause an electron in  $|\mu\mathbf{K}\sigma\rangle$  at  $t = 0$  to have nonzero amplitude also in a state  $|\nu\mathbf{K}'\sigma'\rangle$ . At long times the requirement that this amplitude be negligible translates into the condition  $|e\partial\delta V/\partial t| \ll \hbar/\tau_{\phi}^2$ , expressing the fact

that the applied bias cannot change appreciably over a relaxation time, so that the PME is unable to handle nontrivial time-dependent phenomena.

### 3. MONTE CARLO SOLUTION OF ONE DIMENSIONAL EXAMPLES

In order to implement Eq. (3) in one-dimensional situations, starting from a given potential  $V(z)$ , such as a Thomas-Fermi solution, the Schrödinger equation is solved with vanishing Neumann boundary conditions. The eigenvalues  $\{E_{\mu}\}$  will be spaced densely enough (to accurately reproduce the bulk density of states in the reservoirs) if the length  $L$  of the device is selected appropriately. The extended states can be finally decomposed into left- and right-traveling waves following the 'quantum transmitting boundary method' by Lent and Kirkner [10]. The levels  $|\mu\rangle$  are populated following Pötz [6] and the Poisson equation can be solved again, using well-known potential-damping schemes to improve convergence. Iterating this Poisson/Schrödinger scheme yields a solution in the ballistic case. Scattering is introduced by calculating all transition rates  $W_{\nu\mathbf{K}'\sigma';\mu\mathbf{K}\sigma}$  between the levels, and adding the PME as a third equation over which one must iterate. A Monte Carlo algorithm is used here. Scattering and injection/exit from contacts are treated stochastically in order to determine the new populations  $\rho_{\mu\mathbf{K}\sigma}$ . Typically, 300,000 'electrons' are employed, using a time step  $\delta t \approx 10^{-16}$  s in the Monte Carlo algorithm. After about 100–500 Monte Carlo steps, the Poisson and the Schrödinger equations are solved again until convergence is obtained. Finally, the current flowing through the device is obtained from Eq. (5), while information about carrier density, kinetic energy, and velocity as a function of position inside the device can be obtained from the expectation values of the corresponding operators as traces over their products with the density matrix.

As a first example, let's consider a simple *nin* resistor at 300 K consisting of a 300 nm-long

region between  $z = 0$  and  $z = L$  with  $n$ -type regions 100 nm-long  $n$ -type doped to  $10^{17} \text{ cm}^{-3}$ , and an intrinsic region also 100 nm long. A uniform effective mass of  $m^* = 0.32 m_0$  (where  $m_0$  is the free electron mass) and a dielectric constant  $\epsilon = 11.7 \epsilon_{\text{vac}}$  (where  $\epsilon_{\text{vac}}$  is the permittivity of vacuum) are used. Nonpolar scattering with optical phonons is included using a photon energy 60 meV, an optical deformation potential of  $5 \times 10^8 \text{ eV/cm}$ , a crystal density  $\rho_x = 2.33 \text{ g/cm}^3$ . These parameters imply a ‘dephasing length’  $\lambda_\phi$  in reservoirs of the same material of the order of 785 nm, larger than the active region of the device, about 100 nm long.

Figure 1 shows the potential, carrier density, kinetic energy, and velocity at the end of the iteration. The dashed lines are the corresponding

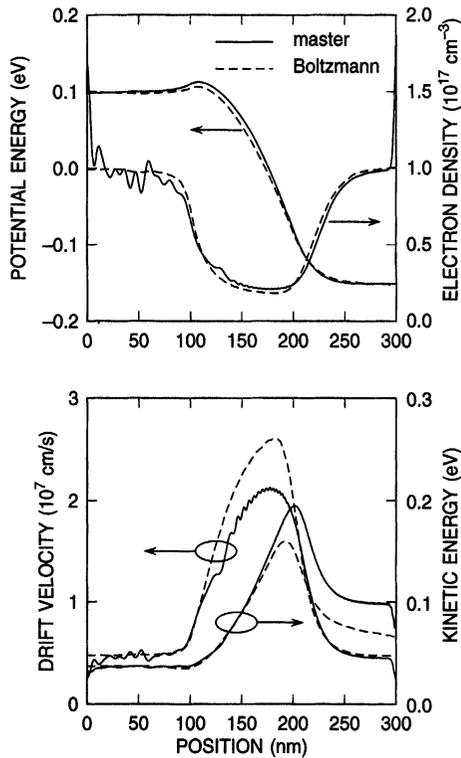


FIGURE 1 Potential and charge density (top frame), kinetic energy and drift-velocity (bottom frame) as a function of position  $z$  inside an  $nin$  device calculated using the master (solid lines) and Boltzmann (dashed lines) equation. The charge accumulation at the device boundaries is an artifact caused by boundary conditions employed to solve the Schrödinger equation.

quantities obtained from a conventional Monte Carlo solution of the BTE using identical parameters. The current density is about  $6.78 \times 10^4 \text{ A/cm}^2$ , about 10% smaller than the value obtained using the BTE,  $7.58 \times 10^4 \text{ A/cm}^2$ . This difference originates mainly from a slightly higher built-in barrier at the cathode/intrinsic-region junction seen in the ‘quantum’ results, caused by the penetration of wavefunctions into the barrier itself. The penetration of charge into classically forbidden region is also the origin of the redistribution of charge seen in the top frame of Figure 1: The charge density obtained using the PME is larger in the intrinsic region, as wavefunctions penetrate into the gap beyond the classical turning point. This charge redistribution is the major factor responsible for the differences seen in average kinetic energy and drift velocity.

The second example is a simple GaAs/Al<sub>0.3</sub>Ga<sub>0.7</sub>As resonant tunneling structure. Two 150 nm thick layers of GaAs-like semiconductor with  $n$ -type doping ( $2 \times 10^{17} \text{ cm}^{-3}$ ) are separated by two AlGaAs tunnel barriers, each 2.8 nm thick, enclosing a 4.5 nm wide semiconductor well. Two 3 nm-thick undoped ‘spacer’ layers, in turn, enclose the double-barrier region. ‘Conventional’ parameters for GaAs and Al<sub>0.3</sub>Ga<sub>0.7</sub>As are used [12], also including nonpolar scattering with acoustic phonons and polar scattering with longitudinal optical phonons. A lattice temperature of 300 K is used. The contact mobility used to displace the distributions injected by the reservoirs is taken to be  $\mu_{el} = 5000 \text{ cm}^2/\text{Vs}$ , corresponding to a dephasing length in the reservoirs,  $\lambda_\phi \approx 130 \text{ nm}$ , larger than the active region of the device. Since even at room temperature scattering is not sufficiently strong to fully populate the states in the accumulation layer, as discussed by Frensley [1], the curves in Figure 2 show that a large fraction of the applied bias drops over the cathode region to the left of the double-barrier region. Although similar results have been obtained by Frensley [1] and Kluksdahl *et al.* [11], they appear poorly credible, since the potential profile clearly depends on the size of the simulated region.

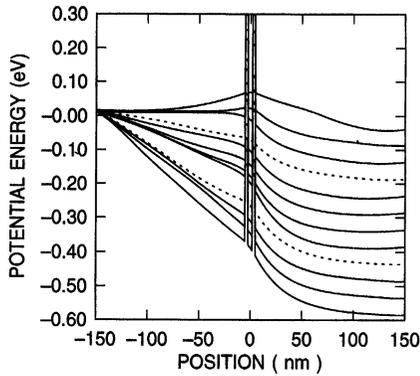


FIGURE 2 Potential energy vs. position in a resonant tunnel diode calculated including electron-phonon scattering. The curves are parametrized by the applied bias from 0.0 to 0.6 V in steps of 0.05 V. Note the 'gaps' at the left of the barriers around 0.2 and 0.45 V. Convergence could not be obtained close to these biases, because of an instability caused by scattering, as discussed in the text. The dotted solutions for biases of 0.2 and 0.45 V are only representative of the average potential during the iteration.

Figure 3 shows the current-voltage characteristics: As the resonance is approached (at about 0.2 V), the associated build-up of charge in the well screens the cathode field, increasing the field across the collector barrier. A critical point is soon

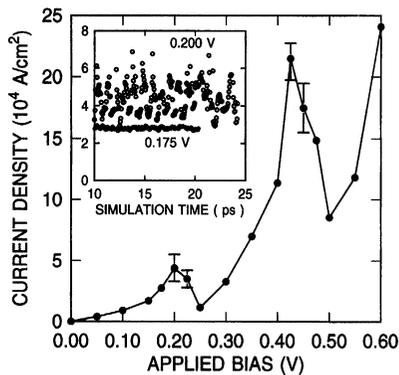


FIGURE 3 Current density in the GaAs/AlGaAs resonant tunneling structure with the inclusion of electron-phonon scattering. Note the 'snap back' of the current at biases of 0.2 and 0.45 V, corresponding to the instability regions of Figure 2. The error bars are used to illustrate the range of the oscillation of the current. The inset shows the Calculated current density vs. simulation time at two bias points, slightly away from (0.175 V) and close to (0.200 V) resonance. Note the instability of the current at the bias point near resonance.

reached at which the rate at which scattering feeds charge into the states localized in the well becomes smaller than the rate at which the well-charge leaks to the right. Charge is removed from the well, and the potential profile now 'snaps', a larger voltage now dropping at the left of the barriers. The population of the current-carrying scattering states, and so the current itself, now drops. As the applied bias is increased further, another 'snapping point' is reached, at about 0.45 V. These critical bias points can be seen as 'gaps' in the cathode potential-profile in Figure 2.

As loosely indicated by the error bars in Figure 3, for biases close to and slightly beyond resonance the situation is unstable: As 'resonant wavefunctions' leak out to the anode, scattering to non-current-carrying states in the collector region is enhanced, thanks to a larger overlap (form) factor. The resulting redistribution of charge alters the potential profile, driving the system away from resonance. As this happens, the charge in the anode side of the barriers drops once more, and the system is now driven back towards a resonant condition, and the cycle repeats once more. The inset of Figure 3 illustrates some features of this instability. The good convergence properties of the computed current density for a bias slightly away from resonance (0.175 V) are compared to the unstable behavior for a bias close to the first resonance (0.200 V). In the latter case, with some imagination one may detect an oscillatory behavior between two states [11, 13].

#### 4. SUMMARY AND CONCLUSIONS

The main conclusion of this paper is that there is an interesting class of problems for which the Pauli master provides a solution of the electron transport problem overcoming some of the strongest limitations of the semiclassical Boltzmann equation: For devices with active regions smaller than a few hundreds of nm, the master equation makes it possible to account for the wave nature of the electrons, although only in the weak scattering

limit and in steady state or slow (adiabatic) time transients. The major price one pays is the inability to handle fast time-transients and strong scattering situations, collisional broadening and coherent transport in the femtosecond time scale being the most notable examples. The advantages consist in the ability to treat inelastic scattering processes, arbitrarily strong electric fields (including intra-collisional field effects) tunneling phenomena, elastic and not, and, in general, all those effects which depend on the wavelike nature of the electrons.

### Acknowledgements

The author is deeply indebted to W. R. Frensley, A. Kumar, R. Landauer, S. E. Laux, P. J. Price, N. Sano and F. Stern for many discussions, suggestions, and criticisms, and to L. Reggiani for having drawn attention to the important work by van Hove.

### References

- [1] Frensley, W. R. (1990). "Boundary conditions for open quantum systems driven far from equilibrium", *Rev. Mod. Phys.*, **62**, 745–791.
- [2] Pauli, W. (1928). In "Festschrift zum 60. Geburtstag A. Sommerfeld" (Hirzel, Leipzig), p. 30.
- [3] Jacoboni, C. (1992). "Comparison between quantum and classical results in hot-electron transport", *Semicond. Sci. Technol.*, **7**, B6–B11.
- [4] See, for instance, Khun, T. and Rossi, F. (1992). "Monte Carlo simulation of ultrafast processes in photoexcited semiconductors: Coherent and incoherent dynamics", *Phys. Rev. B*, **46**, 7496–7514.
- [5] Thanks are due to R. Landauer for making this suggestion, with the strong warning, however, that no formal proof has ever been given that this is indeed true.
- [6] Pötz, W. (1989). "Self-consistent model of transport in quantum well tunneling structures", *J. Appl. Phys.*, **66**, 2458–2466.
- [7] Kohn, W. and Luttinger, J. M. (1957). "Quantum Theory of Electrical Transport Phenomena", *Phys. Rev.*, **108**, 590–611.
- [8] Léon, Van Hove (1955). "Quantum-mechanical perturbations giving rise to a statistical transport equation", *Physica*, **XXI**, 517–540.
- [9] Jones, W. and March, N. H. (1973). "Theoretical Solid State Physics, Volume 2: Non-equilibrium and Disorder" (Wiley-Interscience, Bristol), p. 736 ff. and Appendix A6.1.
- [10] Lent, G. S. and Kirkner, D. J. (1990). "The quantum transmitting boundary method", *J. Appl. Phys.*, **67**, 6353–6359.
- [11] Klusdahl, N. C., Kriman, A. M., Ferry, D. K. and Ringhofer, C. (1989). "Self-consistent study of the resonant-tunneling diode", *Phys. Rev. B*, **39**, 7720–7735.
- [12] Fischetti, M. V. (1991). "Monte Carlo Simulation of Transport in Technologically Significant Semiconductors of the Diamond and Zinc-Blende Structure", *IEEE Trans. Electron Devices*, **38**, 634–649.
- [13] Buot, F. A. and Rajagopal, A. K. (1995). "Theory of novel nonlinear quantum transport effects in resonant tunneling structures", *Mat. Sci. Eng. B*, **35**, 303–317.

### Author Biography

**Massimo V. Fischetti** is a Research Staff Member at the IBM T. J. Watson Research Center in Yorktown Heights, New York. His research interests include electronic transport in semiconductors and insulators and the simulation of semiconductor devices. He is a Fellow of the American Physical Society.



**Hindawi**

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

