PAPER *Special Issue: SISPAD'99*

Effect of the Tunneling Rates on the Conductance Characteristics of Single-Electron Transistors

Andreas SCHOLZE[†], **Andreas SCHENK**[†], *and* **Wolfgang FICHTNER**[†], *Nonmembers*

SUMMARY We present calculations of the linear-response conductance of a SiGe based single-electron transistor (SET). The conductance and the discrete charging of the quantum dot are calculated by free-energy minimization. The free-energy calculation takes the discrete level-spectrum as well as complex many-body interactions into account. The tunneling rates for tunneling through the source and lead barrier are calculated using Bardeen's transfer Hamiltonian formalism [1]. The tunneling matrix elements are calculated for transitions between the zero-dimensional states in the quantum dot and the lowest subband in the one-dimensional constriction. We compare the results for the conductance peaks with those from calculations with a constant tunneling rate where the shape of the peaks is only due to energetic arguments.

key words: single-electron tunneling, Coulomb blockade, conductance oscillations, transfer Hamiltonian formalism

1. Introduction

Single-electron tunneling in semiconductor nanodevices such as single-electron transistors is an interesting subject for application-oriented research towards new device principles. Even though an ever-increasing number of SET structures based on single-electron tunneling and Coulomb blockade (CB) are available now, the interpretation of the spacing and the heights of the CB peaks is still controversial. Main issues concerning, for instance, the statistics of the peak spacing still remain unresolved (for recent work see [2] and [3]). Most of the analysis of experimental data is carried out using the so-called *orthodox theory* of CB and its extensions towards semiconductor quantum dots with a discrete level spectrum, the *constant interaction model* (CI). The basic assumptions of this theory allow to predict the main trends in the line shape and the peak height as well as the spacing of the CB peaks within certain regimes which are given by the range of the quanum dot level spacing, $\Delta\varepsilon$, and the thermal energy, $k_{\mathrm{B}}T$ [4]. The basis of this analysis is an equation for the linear-response current of a quantum dot connected to a reservoir via two tunneling barriers

$$G = \frac{e^2}{k_{\mathrm{B}}T} \sum_{\{n_i\}} P_{\mathrm{eq}}(\{n_i\}) \sum_k \delta_{n_k,0} \frac{\Gamma_k^{\mathrm{s}} \Gamma_k^{\mathrm{d}}}{\Gamma_k^{\mathrm{s}} + \Gamma_k^{\mathrm{d}}} f(\epsilon) \qquad (1)$$

with

$$\epsilon = F(\{n_i\}, n_k = 1, N+1) - F(\{n_i\}, n_k = 0, N) - E_{\mathrm{F}},$$

where $F(\{n_i\}, N)$ is the total (Helmholtz) free energy of the system with N electrons in the quantum dot and the occupation configuration $\{n_i\}$, $n_i = 0, 1$, E_{F} is the Fermi energy and $P_{\mathrm{eq}}(\{n_i\})$ is the Gibbs distribution function of the electron population of a quantum dot in equilibrium with the reservoirs, i. e.

$$P_{\mathrm{eq}}(\{n_i\}) = \frac{\exp\{-\beta[F(\{n_i\}, N) - E_{\mathrm{F}}N]\}}{\sum_{\{n_j\}} \exp\{-\beta[F(\{n_j\}, N) - E_{\mathrm{F}}N]\}}, \qquad (2)$$

with $\beta = (k_{\mathrm{B}}T)^{-1}$, the inverse of the thermal energy.

The tunneling rates for the drain- and the source-side barrier are denoted Γ_k^{d} and Γ_k^{s}. They are often assumed to be constant. Such a treatment is only able to reveal the effect of the free energies on both line shape and peak hight of the conductance peaks. In this paper we present calculations of the tunneling rates using Bardeen's *transfer Hamiltonian* formalism [1]. The tunneling rates are inserted into Eq. (1) and the effect of the tunneling rates on the heights of the CB peaks is investigated for a SiGe test structure.

2. Free-energy minimization

We employ a free-energy minimization scheme, i. e. the chemical potential μ of the quantum dot is changed and the self-consistent ground states and the free energies $F(\{n_k\})$ are calculated for several values of μ, i. e.

$$F(\{n_k\}) = \sum_k n_k \varepsilon_k^0 + E_{\mathrm{xc}} + \frac{1}{2} \sum_{i=1}^{M} Q_i V_i - \sum_{i=1}^{M} \int_0^t \mathrm{d}t'\, I_i(t') V_i(t'). \qquad (3)$$

The first sum over the bare (non-interacting) dot levels ε_k^0 with a particular occupation configuration $\{n_k\}$ is the interaction-free kinetic energy of the quantum dot T_{s} and E_{xc} is the exchange-correlation energy of the electrons in the local-density approximation (LDA) which corresponds to a (local) exchange-correlation potential V_{xc}. The last two terms are the classical contributions to the free energy. The summations are over $i = 1 \ldots M$ distinct elements with the total equilibrium charges Q_i and the voltages V_i. The I_i are the currents provided by the external circuitry (voltage sources).

[†]The authors are with the Institut für Integrierte Systeme, ETH Zürich, CH-8092 Zürich, Switzerland

The device is separated into elements which contain a continuous space charge ρ (dot, leads, donor layer) and metal plates which are equipotential regions (gates, contacts). Consequently, the electrostatic energy U can be written as

$$U = \frac{1}{2}\int_{\hat{\Omega}} \mathrm{d}\mathbf{r}\, \rho(\mathbf{r})\phi(\mathbf{r}) + \frac{1}{2}\sum_{i\in\text{gates}} Q_i V_i. \qquad (4)$$

The potential ϕ is the electrostatic potential, i. e. the solution of the Poisson equation with the charge density ρ subject to boundary conditions at the gates and contacts. The total energy contribution of the quantum dot electrons to the free energy is

$$E_{\text{tot}}(\{n_k\}, Q_{\text{dot}}, V_{\text{dot}}) = \sum_k n_k \varepsilon_k^0 + \frac{1}{2}Q_{\text{dot}}V_{\text{dot}} + E_{\text{xc}}. \qquad (5)$$

The self-consistent single-particle energies ε_k are calculated solving a three-dimensional Kohn-Sham equation with the effective potential $\mathcal{V}_{\text{eff}}(\mathbf{r}) = \Delta E_{\text{c}}(\mathbf{r}) - q\phi(\mathbf{r}) + V_{\text{xc}}(\mathbf{r})$ where ΔE_{c} is the conduction-band offset. This leads to the following form of the ε_k

$$\varepsilon_k = \varepsilon_k^0 + \langle\psi_k| - q\phi(\mathbf{r})|\psi_k\rangle + \langle\psi_k|V_{\text{xc}}(\mathbf{r})|\psi_k\rangle. \qquad (6)$$

Small thermal variations in the level occupancies have a negligible effect on the self-consistent results (ε_k, ϕ, ρ, and Q_i) which are implicit functions of the electron number and the applied voltages [5]. This is especially true at low temperatures where the level spacing is hardly affected by the temperature. Therefore, the occupation configuration dependence of the last two terms is ignored and the discrete occupation numbers n_k are replaced by the non-integer occupation numbers according to the Fermi-Dirac distribution. With this approximation the interaction-free kinetic energy is obtained as

$$T_{\text{s}}(\{n_k\}) = \sum_k n_k\varepsilon_k + q\int_{\Omega} \mathrm{d}\mathbf{r}\, n(\mathbf{r})\phi(\mathbf{r}) - \int_{\Omega} \mathrm{d}\mathbf{r}\, n(\mathbf{r})V_{\text{xc}}(\mathbf{r}). \qquad (7)$$

The above result is inserted into Eq. (5) where the electrostatic energy is replaced according to

$$\frac{1}{2}Q_{\text{dot}}V_{\text{dot}} \to -\frac{1}{2}q\int_{\Omega} \mathrm{d}\mathbf{r}\, n(\mathbf{r})\phi(\mathbf{r}). \qquad (8)$$

The final expression for E_{tot} becomes

$$E_{\text{tot}}(\{n_k\}, \rho, \phi) = \sum_k n_k\varepsilon_k + \frac{1}{2}q\int_{\Omega} \mathrm{d}\mathbf{r}\, n(\mathbf{r})\phi(\mathbf{r}) + E_{\text{xc}} - \int_{\Omega} \mathrm{d}\mathbf{r}\, n(\mathbf{r})V_{\text{xc}}(\mathbf{r}). \qquad (9)$$

The integration is over the quantum dot area only and the electron density n is always positive. The total energy of the quantum dot, E_{tot}, is included in the Helmholtz free energy given by Eq. (3). This leads to the final form of the free energy $F(\{n_k\})$.

In general, the numbers of quantum-dot electrons obtained at arbitrary values of μ are non-integer. The free energies versus the electron number gives a parabola with the minimum at the equilibrium electron number. The free energies at integer numbers of electrons N are determined by spline interpolation. These values are used in the Gibbs distribution, Eq. (2), to calculate the equilibrium number of electrons in the dot and the conductance according to Eq. (1).

3. Bardeen's transfer Hamiltonian formalism

We use the *transfer Hamiltonian* formalism introduced by Bardeen [1] to calculate the tunneling rates Γ_p between a reservoir and the quantum dot. The quantum dot is separated from the reservoir by a narrow quasi-1D channel. By applying a voltage to an electrostatic gate above the channel a narrow constriction is created (see Fig. (1)). Since the constriction is formed electrostatically, its boundaries are smooth and electron scattering can be neglected. An ideal quantum point contact (QPC) between the reservoir region and the quantum dot is formed. The wavefunction inside the

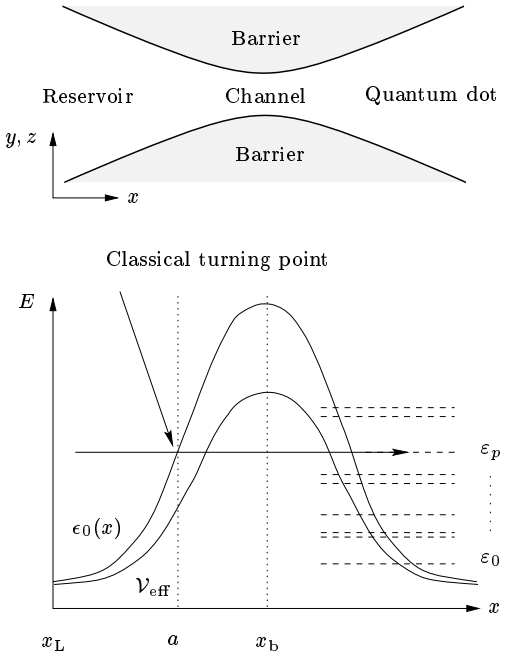

Fig. 1 (a) Schematic view of the constriction. (b) The barrier potential, the energies of the lowest subband and the quantum dot spectrum. a is the classical turning point in the reservoir.

constriction is considered one-dimensional. The width of the constriction can not be considered constant, however. The simplest theory for this case is based on the

adiabatic approximation, which assumes that the cross section of the channel changes so slowly that only negligible scattering between the subbands occurs. Transport is only possible in the x-direction and confinement of carriers occurs in the yz-plane. At each value of x the two-dimensional Schrödinger equation

$$\left[-\frac{\hbar^2}{2m_\perp^*}\left(\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)+ \mathcal{V}_{\rm eff}(x,y,z)-\epsilon_n(x)\right]\phi_n(x;y,z)=0 \quad (10)$$

for the transverse eigenfunctions $\phi_n(x;y,z)$ and the subband energies $\epsilon_n(x)$ is solved where $m_\perp^*$ is the electron effective mass perpendicular to the transport direction. The full wave function ψ satisfying the three-dimensional Schrödinger equation factorizes as

$$\psi \longrightarrow \psi_{nk}(x,y,z)=\phi_n(x;y,z)\xi_k(x), \quad (11)$$

where $\xi_k(x)$ is the solution of the coupled-mode equations

$$\left[-\frac{\hbar^2}{2m_\parallel^*}\frac{\partial^2}{\partial x^2}+\epsilon_n(x)\right]\xi_k(x)=E_{nk}\xi_k(x). \quad (12)$$

The total energy is

$$E \longrightarrow E_{nk}=\epsilon_n(x)+\frac{\hbar^2k^2}{2m_\parallel^*}. \quad (13)$$

Now, the case of an electron moving from a reservoir state labeled E_{nk} to a zero-dimensional quantum-dot state ε_p is considered. Energy conservation in the tunneling process requires that $E_{nk}=\varepsilon_p$. It is convenient to assume that $\mathcal{V}_{\rm eff}$ is constant outside a larger range to the left hand side of the barrier $x<x_{\rm L}$, i. e. $\epsilon_n(x)=\mathcal{V}_{\rm eff,L}=$ const for $x<x_{\rm L}\leqq a$ (see Fig. 1). Since the constriction is formed electrostatically, its boundaries are smooth and scattered components ($\exp(-ikx)$) of the wavefunction can be neglected. The *classical turning point* a in the reservoir is the point in x where $E_{nk}=\varepsilon_p$, i. e. where the energies of the electron before and after the tunneling process are matched and the *classically forbidden* region starts. Classically, the wave is reflected at this point. Quantum mechanically, it can tunnel through the barrier. The WKB approximation is used in the forbidden region $x>a$. The wave function component in x is

$$\xi_\kappa(x)=\exp[w(x)] \quad \text{with} \quad w(x)=\int_a^x {\rm d}x'\,\kappa(x') \quad (14)$$

and $\kappa(x)$ is the positive root

$$\frac{\hbar^2}{2m_\parallel^*}\kappa^2(x)=\epsilon_n(x)-E_{nk} \quad (15)$$

with the assumption of $|{\rm d}\kappa/{\rm d}x|\ll\kappa^2$. The constriction is assumed narrow enough that only a single transverse state is below the Fermi level. As the electron moves away from the constriction, the channel becomes wider and the number of transverse states grows. However, within the constriction a single mode is present i. e. the matrix element is calculated only for the transition from the ground state ($n=0$). Both the channel region and the quantum dot region overlap and the matrix element is calculated by integrating over a surface $\partial\Omega$ in the yz-plane at some point $x_{\rm b}$ (usually taken as the mid-point of the barrier). This leads to

$$M_p \approx -\frac{\hbar^2}{2m_\perp^*}\xi_\kappa(x_{\rm b})\iint\limits_{\partial\Omega(y,z)}{\rm d}y\,{\rm d}z\,\phi_0(x_{\rm b};y,z) \times\left[\kappa(x_{\rm b})\psi_p(x_{\rm b};y,z)-\frac{\partial}{\partial x}\psi_p(x_{\rm b};y,z)\right], \quad (16)$$

where $\kappa(x_{\rm b})$ is defined by Eq. (15) with $E_{0k}=\varepsilon_p$. The transition rate for the tunneling from the lowest reservoir subband to the pth quantum dot state, Γ_p, is obtained from Fermi's golden rule, i. e.

$$\Gamma_p=\frac{2\pi}{\hbar}|M_p|^2. \quad (17)$$

We calculate the eigenvalue spectrum of the quantum dot self-consistently by solving a nonlinear Schrödinger/Poisson equation on a 3D mesh comprising the dot area as well as the barrier regions [6]. The wavefunction for the transverse mode ϕ_0 in the constriction is obtained as the solution of the Schrödinger/Poisson equation solving the Schrödinger equation in slices along the 1D-channel in the leads. The wavefunctions are then used to evaluate the matrix element, Eq. (16), and the tunneling rate, Eq. (17). Finally, the conductance is obtained from Eq. (1).

4. Results

We present selfconsistent calculations for a SiGe based heterostructure single-electron transistor. The quantum dot is electrostatically confined by means of a

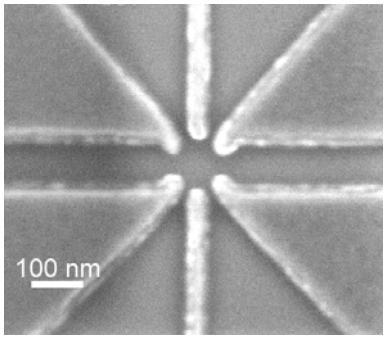

Fig. 2 Top view of the Ti/Al gate structure of the SiGe heterostructure SET. Courtesy PSI Villigen.

Ti/Al top gate (Fig. 2). The 2DEG is situated in a quantum well consisting of a 10 nm thick (100)-strained lattice matched $Si/Si_{1-y}Ge_y$-layer ($y = 0.25$) situated 50 nm below the top gate. The band gap in the quantum well is set to $E_g(\Delta) = 0.9$ eV and the effective masses of the electrons were set to $m_t^*/m_0 = 0.2$ and $m_l^*/m_0 = 0.9$ [7]. The values for m_t^* and m_l^* were also used in the tunneling formalism to approximate $m_\perp^*$ (Eq. 10)) and $m_\parallel^*$ (Eq. 12)). All calculations were performed for a temperature of $T = 1.0$ K. Figure 3 shows the dependence of the conductance peaks on the tunneling rates. In the upper panel, the tunneling rates in the conductance formula, Eq. (1), were kept at a constant value of $\Gamma_k = 1s^{-1}$ for all k. The non-uniform peak

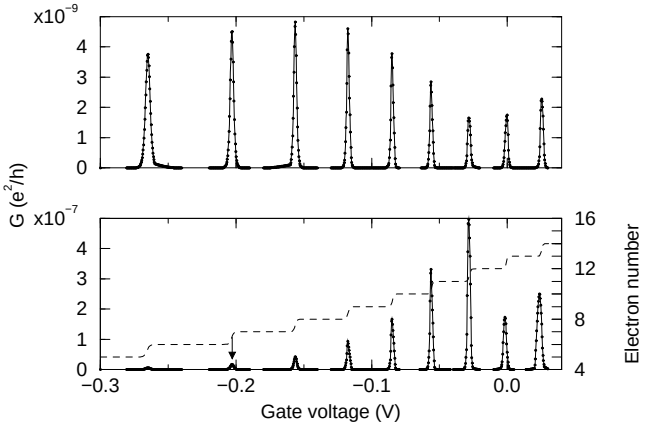

Fig. 3 Upper panel: conductance peaks for a constant tunneling rate $\Gamma_k = 1s^{-1}$ for all dot levels. Lower panel: conductance peaks calculated with tunneling rates according to Eqs. (16, 17). Dashed line: charging of the quantum dot.

height is solely due to the non-uniform spacing of the quantum-dot eigenvalues which contribute to the total energy of the quantum dot, Eq. (9), and the Helmholtz free energy Eq. (3). In the lower panel, the tunneling rates were explicitly calculated using the formalism described above. We see a strong suppression of the first up to the fifth peak. Figure 4 shows one particular conductance peak from the lower panel (marked with an arrow) in comparison to the corresponding peak from the upper panel in Fig. 3. Both peaks are scaled to the same peak height. It can be seen that even though the peak height is different, the shape of the peak remains nearly the same. The slight changes in the peak shape are due to a certain overlap with the neighboring peaks which are dominated by different tunnelin rates. This effect should vanish for $T \to 0$. It can be concluded that the principal shape of the conductance peaks experiences only negligible alterations through the inclusion of realistic tunneling rates i.e. the shape of the conductance peaks is mainly determined by the temperature dependent distribution of the electrons in the quantum dot and not by the tunneling rates.

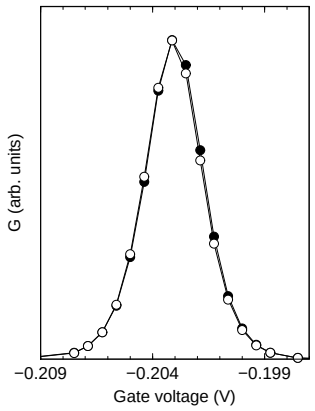

Fig. 4 The marked conductance peak from the lower panel in Fig. 3 overlaid by the corresponding peak from the upper panel.

5. Summary

We showed that the inclusion of realistic tunneling rates within the linear-response conductance formula of Beenakker [4] leads to additional modulations of the conductance-peaks heights. While the line shape remains almost the same, some of the peaks, especially at the lower end of the spectrum are almost completely suppressed. This effect is due to the low energy of those eigenvalues which contribute to the tunneling rate (at low temperatures this is only one level), and therefore the much wider barrier these electrons have to traverse in the tunneling process. This result is in qualitative agreement with many experimental findings (see for instance [8]).

References

[1] J. Bardeen, Tunneling from a many-particle point of view, Phys. Rev. Lett., vol. 6, pp. 57-59, January 1961.

[2] S. R. Patel, D. R. Stewart, C. M. Marcus, M. Gökcedag, Y. Alhassid, A. D. Stone, C. I. Duruöz, and J. S. Harris Jr., Charging the Electronic Spectrum of a Quantum Dot by Adding Electrons, Phys. Rev. Lett., vol. 81, pp. 5900-5903, December 1998.

[3] D. Berman, N. B. Zhitenev, R. C. Ashoori, and M. Shayegan, Observation of Quantum Fluctuations of Charge on a Quantum Dot, Phys. Rev. Lett., vol. 82, pp. 161-164, January 1999.

[4] C. W. J. Beenakker, Theory of Coulomb-blockade oscillations in the conductance of a quantum dot, Phys. Rev. B, vol. 44, pp. 1646-1656, July 1991.

[5] M. Stopa, Coulomb blockade amplitudes and semiconductor quantum-dot self-consistent level structure, Phys. Rev. B, vol. 48, pp. 18340-18343, December 1993.

[6] A. Scholze, A. Schenk, and W. Fichtner, TCAD oriented simulation of single-electron transistors at device level, Proc. of SISPAD'98, Leuven, Belgium, pp. 203-206, October 1998.

[7] M. M. Rieger and P. Vogel, Electronic-band parameters in strained $Si_{1-x}Ge_x$ alloys on $Si_{1-y}Ge_y$ substrates, Phys. Rev. B, vol. 48, pp. 14276-14287, November 1993.

[8] F. Simmel, T. Heinzel, and D. A. Wharam, Statistics of conductance oscillations of a quantum dot in the Coulomb-blockade regime, Europhys. Lett., vol. 38, pp. 123-128, April 1997.

Andreas Scholze was born in 1969. He received the Dipl. Phys. degree from the University of Jena in 1995. After graduation he joined the Integrated Systems Laboratory of ETH as a research assistant, where he now works with the group of Prof. W. Fichtner in the fields of nanodevice simulation and novel device principles. His main activities include the development of device simulation software for the simulation of nanodevices, especially of single-electron transitors.

Andreas Schenk was born in Berlin, Germany, in 1957. He received the Diploma in physics and the Ph.D. in theoretical physics from the Humboldt University Berlin (HUB) in 1981 and 1987, respectively. In 1987 he became a research assistant at the Department of Semiconductor Theory of HUB, and 1988 he joined the R&D division of WF Berlin. From 1987 till 1991 he was working on various aspects of the physics and simulation of optoelectronic devices, esp. infrared detector arrays, and the development and implementation of physical models for modeling infrared sensors. He is now with ETH, Zurich, Switzerland, as a senior lecturer at the Integrated Systems Laboratory. His main activities are in the development of physics-based models for simulation of submicron silicon devices. He is a member of the German Physical Society (DPG).

Wolfgang Fichtner received the Dipl.Ing. degree in physics and the Ph.D. degree in electrical engineering from the Technical University of Vienna, Austria, in 1974 and 1978, respectively. From 1975 to 1978, he was an Assistant Professor in the Department of Electrical Engineering, Technical University of Vienna. From 1979 through 1985, he worked at AT&T Bell Laboratories, Murray Hill, NJ. Since 1985 he is Professor and Head of the Integrated Systems Laboratory at the Swiss Federal Institute of Technology (ETH). In 1993, he founded ISE Integrated Systems Engineering AG, a company in the field of technology CAD. Wolfgang Fichtner is a Fellow of the IEEE and a member of the Swiss National Academy of Engineering.

List of Figures