SIMULATION OF SEMICONDUCTOR DEVICES AND PROCESSES Vol. 5
Edited by S. Selberherr, H. Stippel, E. Strasser – September 1993

Modeling Nano-Structure Devices

K. Hess[‡†] and L. F. Register[‡]

[‡]Beckman Institute for Advanced Science and Technology and Coordinated Science Laboratory, University of Illinois at Urbana-Champaigne
Urbana, IL 61801, USA

[†]Department of Electrical and Computer Engineering,
University of Illinois at Urbana-Champaigne
Urbana, IL 61801, USA

Abstract

Fundamental problems of and approaches to modeling nanostructure devices are reviewed. First the requirements for modeling charge transport in classical and nanostructure devices are compared and contrasted. Then the quantum mechanical concepts of transmission probabilities and eigen energies in nanostructures are related back to the classical concepts of resistance and capacitance, respectively. Next a small illustrative sampling of numerical approaches to calculation of the quantum mechanical properties of nanostructures is presented. Finally examples are given of how such theoretical concepts and numerical methods can be applied to modeling existing and future devices.

1. Introduction

In conventional semiconductor devices, most quantum transport effects can be treated indirectly. The effects of the rapidly varying crystal potential on electron transport can be modeled via the concepts of effective masses, energy gaps, and the positively charged quasi-particle holes. On the macroscopic level, charge transport can then be modeled using the concepts of classical mechanics, aided sometimes by the Pauli exclusion principle and the Born approximation for scattering problems. Only the possibility of electron-hole recombination across the energy gap and the corresponding non-conservation of particle numbers bears witness to the deeper quantum mechanical character of the problem.

Of course, size quantization effects do appear in some conventional devices. The conduction channel of a MOSFET extends only a few nanometers from the Si-SiO_2 interface into the silicon and, thus, the charge carrier motion perpendicular to the interface is quantized. However, charge transport over the oxide barrier is usually negligible, and the carriers can be treated as quasi two-dimensional particles in the plane of the conduction channel and still be modeled semi-classically. How far the dimensions of the conduction channel can be shrunk in this latter plane before quantization leads to prominent effects is currently the subject of speculation.

In addition to quantization effects, the current mainstays of semiconductor device technology, electron and hole gases and the formation of homojunctions by doping, also present barriers to shrinkage as the desired step from the vacuum to the dense

solid has not been achieved entirely after the invention of the point contact transistor by Bardeen and Brattain in 1947 [1]. On the nano-scale, doping is inhomogeneous and majority carriers and particularly minority carriers form dilute gases in conventional devices [2]. If shrinkage is to substantially continue, the homojunction and these dilute gasses must yield to heterojunctions and Fermi-liquids. The increased use of heterojunctions in silicon technology and the interesting science surrounding superlattices and most recently nanostructures and mesoscopic systems indicate motion in this direction.

Nanostructures and mesoscopic systems currently provide an interesting playground for applications oriented science [3, 4]. From an engineering point of view, the mechanism that will make nanostructure devices tick and permit integration on a gigantic scale has yet to be found. In our opinion, however, nanostructure devices will involve regions of significantly unbroken carrier phase extending over many atomic distances, *mesoscopic systems or structures*, alternating with regions where phase coherence is broken and excess carrier energy is dissipated, *reservoirs*, such as contacts and interconnects. Within these mesoscopic structures current will be carried by charge carrier transport within, over and (via tunneling) through heterojunction barriers. Further, these devices will exhibit problems of multi-scale; nanostructure regions, governed by quantum mechanical properties, will be connected to much larger regions, where classical transport models are appropriate. These envisioned transport mechanisms already are evidenced in existing devices containing nanostructures such as quantum well lasers and the now familiar resonant tunneling diode and, therefore, set the stage for the modeling and simulation tools that must be developed.

2. Basic physical theories

Quantum transport theory encompasses a wide range of phenomena including superconductivity, quantum Hall effect, tunneling, non-parabolic and multi-band energy band structures, carrier-carrier interactions and carrier-phonon interactions. Nevertheless, it is often possible to translate the quantum theoretical problem of carrier transport in nanostructures into conventional engineering concepts such as of resistance, capacitance and inductance. Below, only resistance and capacitance are addressed due to spatial limitations. Quantum inductance is discussed extensively in Ref. [5].

The Landauer-Büttiker theory allows direct calculation of charge currents and interpretation of resistance in terms of quantum mechanical transmission coefficients [6, 7, 8]. The Landauer-Büttiker theory, as well as the Bardeen transfer Hamiltonian approach [9], view the mesoscopic (phase coherent) region on the basis of scattering theory, i.e. as simply an object that scatters the charge carrier's wave function. A simple classical analogy to this approach is the use of thermionic emission theory within a drift-diffusion simulator where transport over a heterojunction barrier from one reservoir to another is characterized by the classical transmission coefficient, a step function in energy. Extending this latter method to include quantum mechanical transmission coefficients, as for a tunneling diode, gives for the electron current density perpendicular to the barrier(s) J_{12} [10, 11],

$$J_{12} = 2\frac{e}{(2\pi)^3\hbar}\int_{-\infty}^{\infty} dk_{\|}\int_{E_{c,\max}}^{\infty} dE_{\perp}\; T(E_{\perp}, k_{\|})[f_1(E) - f_2(E)]\,. \tag{1}$$

Here, e is the (negative) electron charge, the leading factor of 2 accounts for spin degeneracy, T is the transmission probability, $f_1(E)$ and $f_2(E)$ are the occupation

probabilities for the electrons as a function of total energy in their respective reservoirs, and $E_{c,max}$ is the greater of the two minimum conduction band energies of the reservoirs. The Landauer-Büttiker theory generalizes this approach to arbitrary mesoscopic geometries and multiple reservoirs. Specifically, in the linear response regime, the current I_{12} flowing from one reservoir to a second via some mesoscopic structure (perhaps in addition to the current flow to and from other reservoirs) is given by [7, 8]

$$I_{12} = 2\frac{e^2}{h} \sum_{n_1,n_2} T_{n_1 n_2} \frac{\mu_1 - \mu_2}{e} . \tag{2}$$

Here, μ_1 and μ_2 are the respective chemical potentials of the two reservoirs, and the $T_{n_1 n_2}$ are the transmission probabilities between the states of the two reservoirs at the Fermi-surface energy. Because $(\mu_1 - \mu_2)/e$ represents the voltage drop V between the reservoirs, the conductance can be identified in terms of the transmission coefficients as $2(e^2/h)\sum_{n_1,n_2} T_{n_1 n_2}$. This formula, Eq. (2), is readily generalized beyond the linear response regime and for high temperatures; however, in either, case the transmission coefficients do not incorporate dissipative processes within the mesoscopic region. Single electron transmission coefficients can be generalized to include dissipative processes, but the significance of this generalized concept becomes unclear when Pauli exclusion effects must be considered simultaneously [12, 13]. Examples of familiar device geometries where the Landauer-Büttiker approach may be applicable include the transition region between the source contact and the quasi-two-dimensional channel of a MOSFET, and the region about an undoped quantum well within the heavily doped classical transport regions in a semiconductor laser. This approach has also been extremely successful in explaining a large amount of experimental results for nanostructures including conductance steps for transport through ultrasmall constrictions [3].

As transmission coefficients are converted to resistance via Landauer-Büttiker theory, the energy spectra of nanostructure systems can be converted to capacitance. The capacitance C of an isolated island or dot of conducting material is defined as the ratio of the stored charge Q in the dot to the voltage V of the dot [14], or, in a form more conducive to nanostructure applications, $C \equiv e^2(N/\mu)$ where N is the total number of electrons within the dot and μ is the chemical potential of the dot. Similarly, the differential capacitance, $C_d \equiv \Delta Q/\Delta V$, is given by [15],

$$C_d(N) \equiv \frac{e^2}{\mu(N+1) - \mu(N)} , \tag{3}$$

where the chemical potential $\mu(N)$ is given by,

$$\mu(N) = E(N) - E(N-1) . \tag{4}$$

Here, $E(N)$ is the total energy of the N-electron system. In metallic systems, capacitance and differential capacitance are effectively independent of N [14], even in many structures exhibiting coulomb blockade effects where μ changes in discrete steps on the scale of the thermal energy with the addition of single electrons to the dot [3, 4, 16]. However, in semiconductor nanostructures with few electrons, capacitance does depend on N due to the electrostatic interactions among the electrons, size quantization contributions to the energies $E(N)$ and Pauli exclusion effects [15, 17, 18].

The interpretation of conductance in terms of electron transmission probabilities, via Landauer-Büttiker theory, and capacitance in terms of multi-electron energies, reduces the problem of calculating nanostructure device parameters to one of solving the Schrödinger equation for one or more charge carriers within regions of phase coherence. Numerical methods for accomplishing this latter goal are considered next.

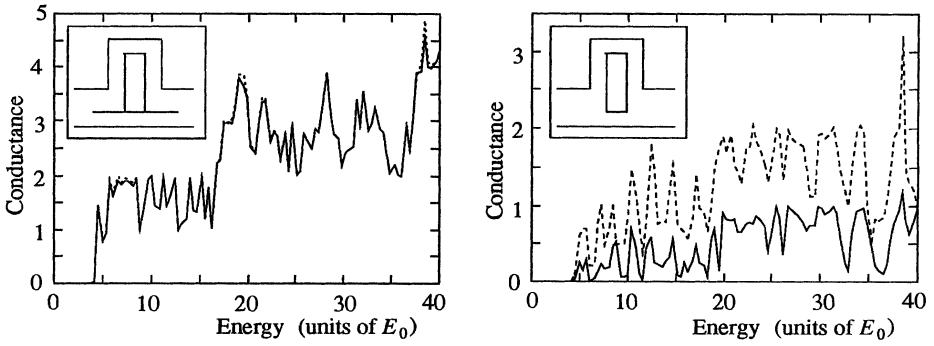

Figure 1: Conductance (in units of $2e^2/h$) as a function of energy (in units of the transverse ground-state energy of of the leads) for two similar mesoscopic structures, shown in the inserts. The exact conductances are indicated by solid lines, while the dashed lines are the simple (classical) sum of the conductances of the upper and lower channels in the absence of the other. (After Ref. [19], ©1992 The American Physical Society.)

3. Numerical methods

There are many numerical approaches to calculating the required transmission coefficients and energy spectra for charge carriers in nano-scale systems. Here, however, we have room to discuss only a few. A common feature of two of the methods discussed is that each attempts to model a many particle problem: one, electron transport in the presence of phonons and, the other, the effects of coulomb and exchange interactions within many electron systems.

One computational efficient method for simulating electron transport in multi-dimensional mesoscopic structures employs a mode matching technique [19]. The analyzed structures are assumed to be defined by hard walls (i.e., potential walls of infinite height) and separable into rectangular sections. Within each of these sections at a given energy, the wave function solution of the time-independent Schrödinger equation can be written as the sum of the products of propagating and evanescent modes in the longitudinal direction and standing waves in the transverse direction. At each interface, enforcing the continuity of the wave function and its normal derivative produces a linear system of equations with a sparse coefficient matrix that can be efficiently solved. Because of the efficiency of the method but restrictions on device geometries, this method is best suited for addressing fundamental questions of quantum transport such as how do conductances add in mesoscopic structures, as in the calculations of conductance, via Eq. (2), in the examples of Fig. 1.

A more computationally intensive but more flexible approach to simulating transmission through multidimensional mesoscopic structures is based on solution of the time-dependent Schödinger equation [20]. This method features robust open boundary conditions and an arbitrarily variable potential function that allows simulation of the transient through steady-state transport in a broad range of mesoscopic structures, and has recently been updated to allow calculation of week dissipative coupling to phonons when Pauli exclusion effects are not critical [21]. Specifically, if the Schrödinger equation for an electron–few-oscillator (few-phonon-mode) system,

$$i\hbar\frac{\partial}{\partial t}\Psi_{n_1,\ldots,n_K}(\vec{r},t) = \left[-\frac{\hbar^2}{2m^*}\nabla^2_{\vec{r}} + V(\vec{r}) + \sum_{k=1}^{K}\hbar\omega_k(n_k+\frac{1}{2})\right]\Psi_{n_1,\ldots,n_K}(\vec{r},t) \tag{5}$$

$$+\sum_{k=1}^{K}\left(\frac{\hbar}{2\omega_k K}\right)^{\frac{1}{2}} M_k(\vec{r})\left[\sqrt{n_k}\,\Psi_{n_1,\ldots,n_k-1,\ldots,n_K}(\vec{r},t) + \sqrt{n_k+1}\,\Psi_{n_1,\ldots,n_k+1,\ldots,n_K}(\vec{r},t)\right],$$

is solved numerically, quantum transport with first order accurate phonon scattering can be treated within the limits of the single electron picture. Here, $V(\vec{r})$ is the applied potential seen by the electron within the mesoscopic structure, and the n_k are the discretized coordinates of the few K oscillators corresponding to their uncoupled eigenstates. The coupling functions $M_k(\vec{r})$ are obtained stochastically as,

$$M_k(\vec{r}) = \sum_{\vec{q}}\lambda_{\vec{q}}M_{\vec{q}}(\vec{r})\,, \tag{6}$$

where the $M_q(\vec{r})$ are the coupling functions for the true electron-phonon system for phonon modes $\vec{q}$, and the λ_s are random numbers ($\langle\lambda_{\vec{q}}\rangle = 0$, $\langle\lambda_{\vec{q}}\lambda_{\vec{q}'}\rangle = \delta_{\vec{q},\vec{q}'}$) such that the $M_k(\vec{r})$ have the spatial correlation function characteristic of the original system's coupling functions,

$$\langle M_k(\vec{r})M_k(\vec{r}')\rangle = \sum_{\vec{q}}M_{\vec{q}}(\vec{r})M_{\vec{q}}(\vec{r}')\,. \tag{7}$$

Because the correlation function on the right-hand-side of Eq. (7) is the common factor of all first order calculations in the electron-phonon coupling (as, for example, in "golden rule" calculations of scattering rates), the identity of Eq. (7) assures, on average, the equivalence of the true multiphonon system and the few-oscillator system of Eq. (5) to first order. As the number of samples K is increased, the sampling error is reduced and second and higher order artifacts of using a small oscillator system are reduced at least as 1/K, while some true higher order process are retained. Figure 2 shows simulation results for emission of 50meV polar optical phonons by an approximately 50meV electron incident through a quasi-one-dimensional quantum wire into a quasi-two-dimensional semiinfinite plane, each 50Å thick in the direction perpendicular to the plane of the simulation. The ground-state energy of an electron within the wire is approximately 40meV such that real (as opposed to virtual) phonon emission is only possible within the semiinfinite plane. Twenty samples of $M_k(\vec{r})$ were used, and much of the fluctuation in Fig 2(d) is due to sampling error. However, the calculated probabilities of reflection (inherently without phonon emission), transmission through the simulation region without phonon emission, and transmission with phonon emission are 14.6%, 64.5% and 20.9%, respectively, all to within approximately ±3% RMS deviation.

Calculation of the energy spectrum and, in turn, the capacitance for several electrons in a small island of material, requires balancing a desire for rigor with computational resources. One approach that appears well suited to the task is the local density functional formalism [15]. For this method, the Schrödinger equation that must be self-consistently solved for each electron i in the presence of the others j is,

$$-\frac{\hbar^2}{2m^*}\nabla^2\psi_i(\vec{r}_i) + [V_{\rm b}(\vec{r}_i) + V_{\rm c}(\vec{r}_i) + V_{\rm ex}(\vec{r}_i) + V_{\rm corr}(\vec{r}_i)]\psi_i(\vec{r}_i) = E\psi(\vec{r}_i)\,. \tag{8}$$

Here, $\psi_i(\vec{r}_i)$ is the wave function for the ith electron, $V_{\rm b}(\vec{r}_i)$ is the built in potential of the dot of material and any applied potential, and $V_{\rm c}$ is the coulomb interaction term among the electrons given by,

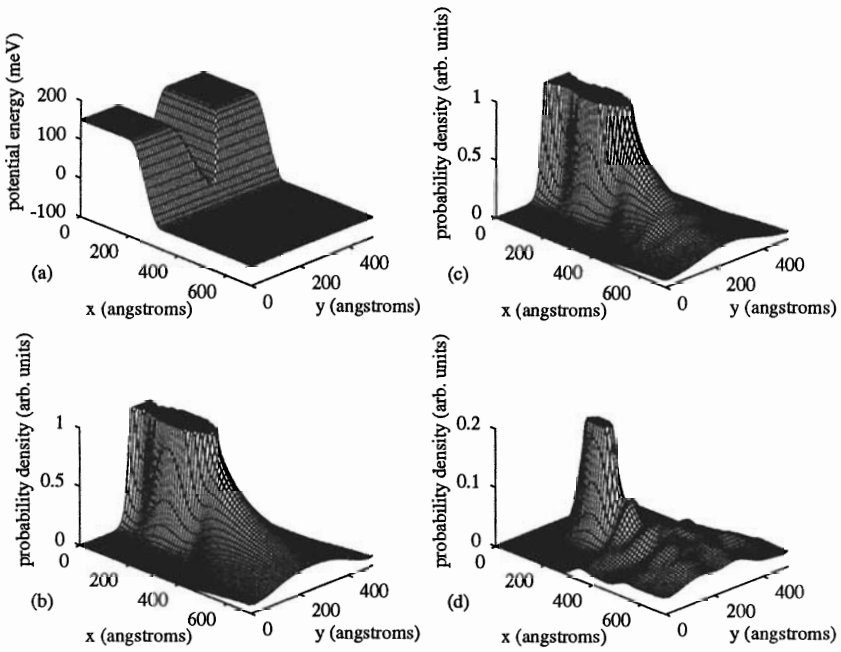

Figure 2: (a) Potential energy function and (b) probability density function for an uncoupled electron. (c) Probability density for the corresponding coupled electron-phonon system in the ground-state of the uncoupled phonon system (prior to phonon emission) and (d) the probability density in an excited state of the phonon system (after phonon emission). The peaks in the probability density have been cutoff to allow viewing of the regions of lower probability. (After Ref. [21].)

$$V_c(\vec{r}_i) = \sum_{j=1}^{N} \frac{1}{4\pi\epsilon_0\epsilon_r} \int_S \frac{e^2|\psi_i(\vec{r}_j)|^2}{|\vec{r}_i - \vec{r}_j|} d\vec{r}_j \,, \tag{9}$$

where N is the total number of electrons in the dot and ϵ_0 and ϵ_r are the absolute and relative dielectric constants, respectively. Nearby metallic electrodes (gates) can also be modeled via the method of images. $V_{\rm ex}(\vec{r}_i)$ and $V_{\rm corr}(\vec{r}_i)$ are the exchange and correlation terms, respectively. These latter terms have received extensive attention in the recent literature, and the easiest representations to include for a thin layer of charge in numerical calculations are the polynomial expressions from the theory of Tanatar and Ceperley [22]. Figure 3 shows the calculated differential capacitance, Eq. (3), vs. electron number for an isolated small quasi-two-dimensional box with hard walls and a quasi-parabolic confinement potential due to a uniformly distributed positive charge of $100|e|$.

4. Current and future device applications

The computational problems encountered in numerical simulation of quantum well lasers are illustrative of those that can be expected in general nanostructure simula-

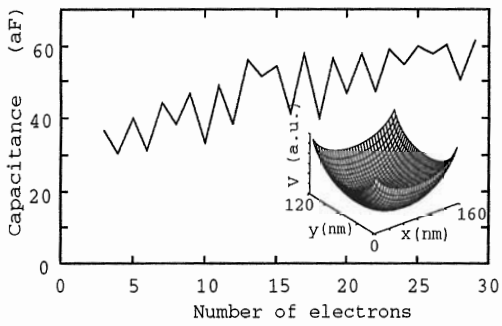

Figure 3: Differential capacitance vs. electron number for an isolated 160nm×120nm quasi-two-dimensional box with hard walls and a quasi-parabolic confinement potential due to a uniformly distributed charge of $100|e|$. The inset shows the self-consistent potential for 10 electrons. (After Ref. [15].)

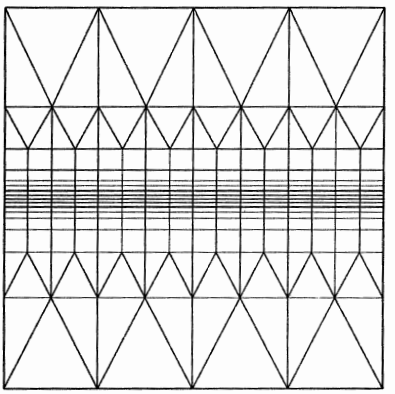

Figure 4: Grid used for numerical simulation of charge transport within a semiconductor quantum well laser. (After Ref. [23].)

tions. For a laser, the width of the quantum well region(s) relative to the entire laser thickness is as an inch to the height of the Sears Tower in Chicago and, thus, a grid system is required that provides suitable resolution both within and outside the quantum well region, as shown in the neighborhood of the quantum well in Fig. 4 [23]. The explosion of grid points near the well clearly demonstrates the multi-scale nature of the problem. Further, outside of the quantum well, simple classical drift-diffusion theory provides a good model of charge transport, while quantum mechanical calculations are required to obtain the energy spectrum and density of states for the charge carriers within the well to model photon emission. Thus, the question is one of how to connect these regions, and a straightforward solution answer lies in the spirit of the Landauer-Büttiker and Bardeen transfer Hamiltonian approaches. To a large extent, the entire quantum well region can be treated as a scattering (recombination) center within the drift diffusion simulator. Carriers within and above the well then can communicate with each other by a multipoint approach, as in the case of thermionic emission [23]. A rather complex understanding of laser switching, phonon emission and carrier-carrier interactions is still required to model carrier capture in the quantum well, but at least for quasi-two-dimensional quantum wells such information is available. For quantum dot lasers, if indeed carriers can be collected efficiently by quantum dots, an approach such as that of Eq. (5) will be required to model the process.

One particularly intriguing device concept of the future is the single-electron transistor [3]. The basic component of such a device will be a dot of material in which the chemical potential changes discretely on the scale of the thermal energy with the number of electrons. For metallic dots, the total energies can be calculated from $E(N) = (Ne)^2/2C$ where the capacitance C is effectively a geometry dependent constant only. However, in semiconductor nanostructures with few electrons, calculation of the the multi-electron energies will require an approach such as that of Eqs. (8) and (9). These energies can then be translated into chemical potentials and capacitances via Eqs. (3) and (4). If the dot is loosely coupled to an external reservoir via a mesoscopic barrier or other structure, a current will flow according to the differences in chemical potentials as described by Eq. (2). However, because the chemical po-

tential of the dot changes discretely with the number of electrons, coulomb exclusion and blockade effects can prevent or control current flow as evidenced by the periodic conductance oscillations reported in many publications [3, 4, 16]. A gate (perhaps another dot) then can be used to control the energy spectrum of the dot and, thus, conduction through the dot to provide transistor action.

In conclusion, the development of theory and computational resources is now at a stage that much of nanostructure electronics can be understood and simulated. However, deep seated problems remain with regard to the multi-scale and many-body aspects of nanostructure systems.

Acknowledgement: This work was supported by the United States Office of Naval Research, Army Research Office and National Science Foundation (NCCE).

References

[1] J. Bardeen and W. H. Brattain, *Phys. Rev.* **74**, 230 (1948).

[2] A carrier density of $10^{18}/\mathrm{cm}^3$ translates to one carrier per 1,000 cubic nanometers.

[3] *Mesoscopic Phenomena in Solids*, ed. by B. L. Altshuler, P. A. Lee and R. A. Webb (North-Holland, Amsterdam, 1991). (Vol. 30 of *Modern Problems in Condensed Matter Sciences*, ed. by V. M. Agranovich and A. A. Maradudin.)

[4] *Nanostructures and Mesoscopic Systems*, ed. by W. P. Kirk and M. A. Reed (Academic Press, Boston, 1992).

[5] J. R. Tucker and M. J. Feldman, *Reviews of Modern Physics* **57**, 1055 (1985).

[6] R. Landauer, *IBM J. Res. and Develop.* **1**, 233 (1957).

[7] M. Büttiker, *Phys. Rev. Lett.* **57**, 1761 (1986).

[8] A. D. Stone and A. Szafer, *IBM J. Res. Develop.* **32**, 384 (1988).

[9] J. Bardeen , *Phys. Rev. Lett.* **6**, 57 (1961).

[10] C. B. Duke in *Solid State Physics*, ed. by F. Seitz, D. Turnbull and H. Ehrenreich (Academic Press, New York, 1969) Vol. 10, pp 24–32.

[11] Because electrons most easily exhibit quantum mechanical behavior in semiconductors, they are usually referred to here. However, the theoretical and numerical methods discussed here apply equally well to holes.

[12] F. Sols, *Annals of Physics* **214**, 386 (1992).

[13] F. Sols, unpublished.

[14] See, for example, P. Lorrain and D. R. Corson, *Electromagnetic Fields and Waves* (W. H. Freeman and Company, San Francisco, 1970), p76.

[15] M. Macucci, K. Hess and G. J. Iafrate, unpublished.

[16] *Condensed Matter: Special Issue on Single Charge Tunneling* **85**, ed. by H. Grabert, 319 (1991).

[17] D. V. Averin, A. N. Korotkov, and K. K. Likharev, *Phys. Rev. B* **44**, 6199 (1991).

[18] Y. Meir, N. S. Wingreen, and P. A. Lee, *Phys. Rev. Lett.* **66** 3048 (1991).

[19] M. Macucci and K. Hess, *Phys. Rev. B* bf 46, 15357 (1992).

[20] L. F. Register, U. Ravaioli and K. Hess, *J. Appl. Phys.* **69**, 7153 (1991). [Erratum: **71**, 1555 (1992).]

[21] L. F. Register and K. Hess, unpublished.

[22] B. Tanatar and D. M. Ceperley, *Phys. Rev. B* **39**, 5005 (1989).

[23] M. Grupen and K. Hess, unpublished.