

Electronic energy spectrum and the concept of capacitance in quantum dots

M. Macucci

Dipartimento di Ingegneria dell'Informazione, Università di Pisa, Via Diotisalvi 2, I-56126 Pisa, Italy

Karl Hess

Beckman Institute, University of Illinois, 405 North Mathews Avenue, Urbana, Illinois 61801

G. J. Iafrate

U.S. Army Research Office, Research Triangle Park, North Carolina 27709-2211

(Received 11 May 1993; revised manuscript received 12 July 1993)

The chemical potential and the capacitance of a model quantum dot have been computed, including contributions of exchange and correlation in the limit of 0 K temperature. The Schrödinger equation has been solved self-consistently, taking into account the electron-electron Coulomb interaction and many-body effects within the framework of density-functional theory. We have also studied the effect of conducting backgates and of nearby electrodes using the method of images. Depending on the size of the dot, we derive a prevalence of either the quantization energy or the electrostatic energy: there is a smooth transition from predominant quantum effects for small dots to classical capacitance behavior for large dots. Our simulation reproduces characteristic effects that have been experimentally observed, such as the capacitance increase for increasing electron numbers and irregularities in the chemical potential values when randomly distributed charged impurities are present.

I. INTRODUCTION

Mesoscopic systems such as low-dimensional semiconductor nanostructures have been the subject of many recent theoretical¹⁻⁹ and experimental¹⁰⁻¹⁵ investigations. Of particular interest has been the transport of electrons on and off single or multiple quantum dots. Quantum dots are regions of electron confinement whose dimensions in all directions are smaller than the typical length L_ϕ of phase coherence. L_ϕ depends on many factors, but is particularly large for a quasi-two-dimensional electron gas generated by modulation doping¹⁶ at a semiconductor heterojunction. The extension of the wave function perpendicular to the heterointerface is typically much smaller than L_ϕ and smaller than the other dimensions of the quantum dot. Therefore this dimension can often be ignored and the electrons are regarded as two-dimensional quasiparticles, an approach we are taking here to reduce the needed numerical resources. The other dimensions of quantum dots as they are experimentally investigated are often not well known and range from hundreds¹¹ to tens¹⁷ of nanometers.

A most characteristic result observed in quantum dot electron transport experiments is the appearance of a "periodic" oscillation of the conductance between two leads loosely coupled through the dot as a function of electron density (which is typically varied with gate voltages).^{10,14} This effect is a consequence of the generalized Coulomb blockade effect,¹⁸⁻²⁰ and has been given theoretical foundations in a number of papers.^{6,14} A careful inspection of the experimental data reveals that these conductance oscillations are not exactly periodic. For example, the period diminishes with increasing electron

density¹⁴ and irregularities of level spacing¹⁷ have been reported. The magnitude of the period itself is often not calculated from first principles, but attributed to a dot-self, dot-gate, or dot-dot capacitance. We attempt here to give a detailed numerical solution to a "model" quantum dot which explains various important contributions to the period of the conductance oscillations. Our approach is based on local-density-functional theory for a two-dimensional electron gas. We have included the effect of gates and generally of neighboring conducting structures that cause image force effects. We also have included the contribution of spin and many-body effects within the given framework. Our major findings are that the concept of a constant capacitance cannot be pushed too far, particularly for small dot sizes and electron numbers. Then the form of the electron wave function and the effects of exchange and correlation cause changes of the capacitance as electrons are added to a dot. There are also straightforward effects such as changes of dot size with electron number and changes of dot shape due to the proximity of charged impurities that contribute to deviations from strict periodicity. At smaller electron densities and for small dot sizes, effects such as shell filling and level pairing may also become important. Overall, a picture that is considerably more involved than that of a constant electrostatic capacitance emerges. We discuss the implications of our results for the definition and concept of capacitance on the nanoscale and compare the magnitude of the various contributions to capacitance. We conclude that quantum effects will become prominent at dot sizes smaller than the ones currently investigated or in materials with higher dielectric constant than gallium arsenide.

II. MODEL DEFINITION AND NUMERICAL APPROACH

We have studied a model quantum dot with a quasiparabolic confining potential, including the Coulomb interaction between electrons and exchange and correlation effects within the framework of density-functional theory.

The choice of a two-dimensional model represents a compromise between realistic simulation, accuracy, and computational feasibility. Calculations including tens of electrons in the dot require a rather fine discretization grid, since wave functions with a significant number of maxima (minima) have to be considered. A three-dimensional approach leads to a prohibitively large eigenvalue problem and would probably not add significantly new features to the trends of the calculation, because the vertical dimension of the dot [the thickness of the quasi-two-dimensional (2D) electron gas layer from which the dot is obtained] is much smaller than the other two dimensions. In this situation the vertical degree of freedom gives a substantially constant contribution to the total energy of the system.

A typical lateral confinement structure¹² which can produce a quantum dot such as the one we have studied is shown in Fig. 1. The quasi-2D electron gas is obtained by modulation doping and the dot is defined by the top metal gate, characterized by a rectangular aperture (the dashed lines indicate the geometry of the gate actually used in the experiments, in which the dot is connected to quantum wires via tunneling barriers, the coupling by tunneling to neighboring leads is weak and can be treated as a small perturbation). When a suitable negative potential is applied to the metal gate with respect to the n^+ GaAs substrate, the 2D electron gas is depleted everywhere with the exception of the central area under the aperture. The resulting shape of the potential confining the electrons in the dot is known to be approximately parabolic³ and our calculations have been performed with a quasiparabolic potential produced by a positive background charge (introduced for the purpose of computational convenience) uniformly distributed on

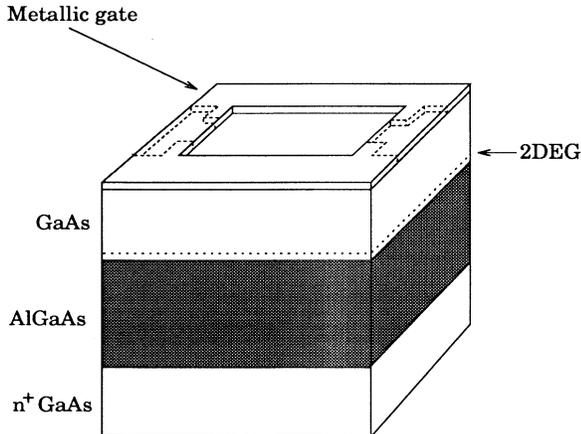


FIG. 1. Schematic representation of a typical heterostructure with electron confinement to a quantum dot.

the surface of the box. For all our calculations the box has a rectangular shape with an aspect ratio of 4:3 and the total background charge generating the confining potential has a value of $100q$ (q being the charge of the electron), unless otherwise specified.

Along the perimeter of the box we enforce Dirichlet boundary conditions, equivalent to hard walls. In almost all the cases we are considering, however, the boundary conditions act where the wave functions have already vanished because of the action of the quasiparabolic potential.

We assume the effective mass approximation to be valid and use a value typical for gallium arsenide: $m^* = 61.03 \times 10^{-33}$ kg. For the discretization of the Schrödinger equation we apply a standard five point formula over a grid with a number of mesh points varying between 60×45 and 120×90 .

In the following, the term “box” will be used to denote the region defined by the Dirichlet boundary conditions and by the extension of the background charge defining the potential for confinement, while the term “dot” will refer to the area where the electron density is nonzero.

The Schrödinger equation is solved self-consistently with the Hartree (Coulombic), exchange, and correlation potentials. Self-consistency is reached through a fixed-point iteration, which in the following will be termed “outer iteration.” The “inner iteration,” a Ritz iteration, solves for the eigenvalues and eigenvectors of the banded matrix which is obtained from the discretization of the Schrödinger equation. The eigenvectors resulting from each outer iteration are used as initial guesses for the eigenvalue computation in the following outer iteration. This is useful to speed up the convergence of the eigenvalue solver, since the potential (and therefore the matrix elements) of each new outer iteration differs from the previous iteration only by a small amount.

The Schrödinger equation we are solving is the following:

$$-\frac{\hbar^2}{2m^*} \nabla^2 \psi_i(\mathbf{r}_i) + [V_c(\mathbf{r}_i) + V_{\text{ex}}(\mathbf{r}_i) + V_{\text{corr}}(\mathbf{r}_i) + V_b(\mathbf{r}_i)] \psi_i(\mathbf{r}_i) = \varepsilon_i \psi_i(\mathbf{r}_i). \quad (1)$$

Here \hbar is the reduced Planck constant, m^* is the effective mass of the electron, $\psi_i(\mathbf{r}_i)$ is the wave function for the i th electron, and ε_i is its energy eigenvalue. V_c is the Coulomb interaction term given by

$$V_c(\mathbf{r}_i) = \sum_{j=1}^N \frac{1}{4\pi\epsilon_0\epsilon_r} \int_S \frac{q^2 |\psi_j(\mathbf{r}_j)|^2}{|\mathbf{r}_i - \mathbf{r}_j|} d\mathbf{r}_j, \quad (2)$$

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. The relative dielectric constant is the one for gallium arsenide $\epsilon_r = 12.9$. The exchange $V_{\text{ex}}(\mathbf{r}_i)$ and correlation $V_{\text{corr}}(\mathbf{r}_i)$ terms have received extensive attention in the recent literature. From the numerical viewpoint, easiest to include are polynomial representations of the theory of Tanatar and Ceperley.²¹ The theory is particularly simple if spin polarization is neglected and this is the approach we have used in most of our simulations.

We refer to it as the TC theory. We have also verified that inclusion of spin polarization following the approach of Perdew and Zunger²² leads to very slight modifications. A different, in principle more rigorous, approach to the determination of $V_{\text{ex}}(\mathbf{r}_i)$ has been described by Krieger, Li, and Iafrate.²³ This method requires considerable numerical resources and its applications are currently being tested.

The results for the exchange [$V_{\text{ex}}(\mathbf{r}_i)$] and correlation [$V_{\text{corr}}(\mathbf{r}_i)$] terms of the TC theory are given in the Appendix with and without the inclusion of spin polarization. The term $V_b(\mathbf{r}_i)$ in the Schrödinger equation represents the background, confining potential.

We are interested in the limit of zero temperature. Due to the stepwise shape of the Fermi function, standard methods such as the Newton scheme are not successfully applicable for this case. We have therefore used a simple relaxation technique, with the total potential actually used at each iteration given by $V(i) = V(i-1)(1-\alpha) + V^{\text{ED}}(i)\alpha$, where $V(i-1)$ is the potential which has been used for the previous iteration and $V^{\text{ED}}(i)$ is the potential directly obtained from the electron density resulting from the previous iteration. α is the relaxation parameter and has, in general, to be adjusted “manually.”

Our convergence criterion consists in checking whether the sum of the mean square differences between the normalized eigenfunctions at two consecutive iterations has become less than a given threshold, which is typically set to a value between 10^{-5} and 10^{-8} , in order to obtain a reasonably precise result for the quantities of interest (chemical potential and differential capacitance). These quantities are obtained through successive differentiations of the total electron energy and therefore are very sensitive to even extremely small errors in the eigenvalues. For dot sizes around 100 nm convergence is easily achieved up to 10–12 electrons and, with a larger but still acceptable residual, up to 25–30 electrons.

We have performed calculations for isolated quantum dots, for dots in the presence of a conducting backgate parallel to the plane which contains the 2D electron gas, and for dots with a conducting plane orthogonal to the 2D gas, representing an upper limit for the effect of the edge of a lead. The method of images has been used for the computation of the electrostatic potentials.

III. COMPUTATION OF THE CHEMICAL POTENTIAL

One of the most important quantities in the study of highly confined systems in semiconductors is the chemical potential as a function of the number of electrons contained in the structure. The conductance peaks, observed in several experiments between leads connected to a quantum dot by tunneling barriers,^{10,11} correspond to the alignment of the chemical potential of the quantum dot with the chemical potential in the leads.¹⁴ The chemical potential is defined as

$$\mu(N) = E(N) - E(N-1), \quad (3)$$

where E indicates the total energy of the electrons in the dot. Therefore, the chemical potential represents the en-

ergy necessary to add the N th electron to a system of $N-1$ electrons. In the noninteracting electron picture, this energy corresponds to the state occupied by the N th electron. The same would be true for a system obeying Koopman’s theorem²⁴ (i.e., the lower orbitals are not significantly perturbed by the addition of one electron). For the quantum dot model which we are considering this is almost never precise, as we found by computing the total energies for a few values of N . Two approaches are possible. One consists in the computation of $E(N)$ and of $E(N-1)$ by summing the energy eigenvalues for all the electrons and subtracting the Coulomb interaction energy, which would otherwise be counted twice.²⁵ A further correction is to be included in order to take exchange and correlation contributions into proper account:²⁵

$$E(N) = \sum_{i=1}^N \varepsilon_i - \frac{1}{2} \int \int \frac{q^2}{4\pi\epsilon_0\epsilon_r} \frac{n(\mathbf{r})n(\boldsymbol{\rho})}{|\mathbf{r}-\boldsymbol{\rho}|} d\mathbf{r}d\boldsymbol{\rho} + \int n(\mathbf{r})\{E_{\text{ex}}[n(\mathbf{r})] + E_{\text{corr}}[n(\mathbf{r})] - V_{\text{ex}}[n(\mathbf{r})] - V_{\text{corr}}[n(\mathbf{r})]\} d\mathbf{r}, \quad (4)$$

where N is the total number of electrons, ε_i are the energy eigenvalues for each electron, $n(r)$ is the total electron density [normalized so that $\int n(\mathbf{r})d\mathbf{r} = N$], and $E_{\text{ex}}, E_{\text{corr}}, V_{\text{ex}}, V_{\text{corr}}$ are the exchange and correlation energies and potentials, respectively.

Another possible approach is based on Slater’s transition rule,²² which, with a good approximation, yields

$$\mu(N) = E(N) - E(N-1) = E(N-0.5). \quad (5)$$

We have followed this approach whenever applicable, because the chemical potential can then be calculated without the need for differentiation.

In Fig. 2 we show the results for the chemical potential versus the number of electrons in a 180×135 nm box. The inset shows the result of a hypothetical experiment in which the conductance would be measured^{10,11} between two leads weakly coupled to the dot. The conductance peaks, represented by the vertical lines, are placed in correspondence with the values of the electrostatic dot

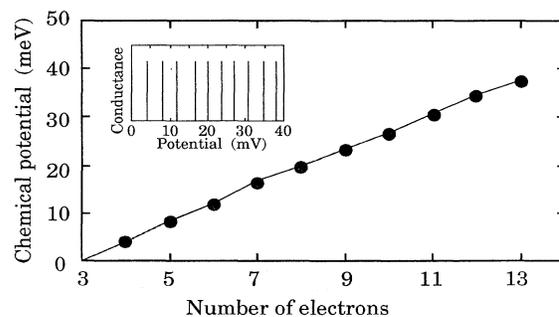


FIG. 2. Chemical potential vs electron number in a 180×135 nm box with hard walls and quasiparabolic confinement. Conductance peaks which would be obtained between two leads loosely coupled to the dot are plotted vs the dot potential in the inset.

potential (controlled by means of an electrode or gate) for which the chemical potentials of dot and leads are aligned for each integer number of electrons. In this hypothetical experiment the chemical potential in the dot is not perturbed by the presence of the leads because the coupling is assumed to be sufficiently weak. We stress the fact that this graph is for representation purposes only and the height of the conductance peaks in our figures has no physical meaning. The origin for the dot potential scale is set to be coincident with the position of the conductance peak for the third electron. Notice that the average spacing between the conductance peaks tends to decrease for an increasing number of electrons. This is due to essentially two effects: the increasing effective size of the dot and the variations in the electron density. For few electrons the electron density is peaked in the center of the box and the effective size of the dot (defined as the region over which the electron density differs appreciably from zero) is slightly smaller than in the case of many electrons. In this latter case the potential seen by each electron is shallow and the electron distribution fairly smooth. This leads to a differential capacitance C_d which increases with the number of electrons, therefore decreasing the width of the intervals between two consecutive values of the chemical potential, since $\mu(N+1) - \mu(N) = q^2/C_d(N)$ (see Sec. V). The reduction of the spacing between the peaks is clearly present in experimental results¹⁴ and was also previously numerically simulated by Stopa.⁹ The concept of the effective size can be better understood from Fig. 3, where the electron density is shown for $N = 9$ for the same parameters as used in Fig. 2. The density is significantly different from zero only over an area which is about 70% of the total area of the box. Changes in the number of electrons between one and a few tens cause variations between 60% and 90% for most of the dots we have studied (depending also on the shape of the confining potential).

Except for the effects of this phenomenon, the chemical potential grows almost linearly with the number of electrons, which implies that the spacing between the peaks is almost even for this box size and background

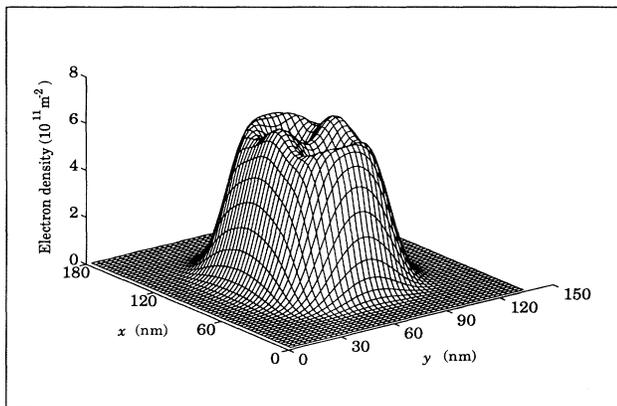


FIG. 3. Electron density in a 180×135 nm box with hard walls and quasiparabolic confinement for a total of nine electrons present in the box.

charge. This is a clear indication that the behavior of the dot is close to the one of a classical capacitor, i.e., that the Coulomb energy associated with the addition of an electron is much larger than the spacing between the energy levels originating from the energy quantization in the dot. We therefore expect much less regular spacing of the peaks for smaller dots, since the energy quantization scales with the reciprocal of the square of the box size, while the Coulomb interaction energy scales with the reciprocal of the box size.¹ The transition is actually fairly smooth, as can be seen from Fig. 4, where results for the chemical potential for various box sizes are reported in the same conductance peak format as the one of the inset of Fig. 2. The sizes are expressed in nanometers and the potential reference for each box is scaled multiplying it by a factor proportional to the linear dimension of the box itself (the aspect ratio, i.e., the ratio of the length to the width of the box, is constant). If the total energy of the electron system were due only to the Coulomb interaction (classical limit), then the graphs would be all identical after the scaling procedure. This is clearly not the case, and significant kinetic energy contributions appear for smaller dots, which give rise to irregularities in the spacings and to an increase in the value of the chemical potential. The increase is larger for smaller dots because the kinetic energy term scales, as mentioned before, with the square of the reciprocal dot size. The behavior becomes almost classical for dot sizes larger than 160 nm.

It is important to notice that for the smallest dots the peaks appear grouped in pairs. There is a simple explanation for this: for decreasing box size, the Coulomb contribution becomes less dominant and each of the chemical potential peaks corresponding to odd occupancy approaches the value for even occupancy. In the limit of kinetic contribution only, the chemical potential for each level of an even-odd pair is the same,¹⁴ since the energy needed to add each of the two electrons is the same and depends only on the energy of the corresponding orbital. A splitting appears as soon as the Coulomb interaction becomes significant, and for large dots it becomes so important that the effects of energy quantization disappear, which results in evenly spaced peaks. This trend can be examined quantitatively in the results shown in Fig. 5, which shows $R \equiv [\mu(5) - \mu(4)]/[\mu(6) - \mu(5)]$ versus the

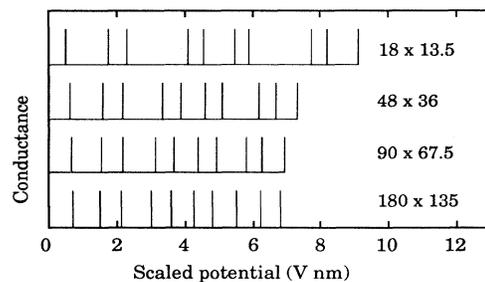


FIG. 4. Conductance peaks vs scaled potential for boxes of different sizes (expressed in nanometers) with 4:3 aspect ratio. The potential is scaled by multiplying its value in volts by the length of the longer side of the box in nanometers.

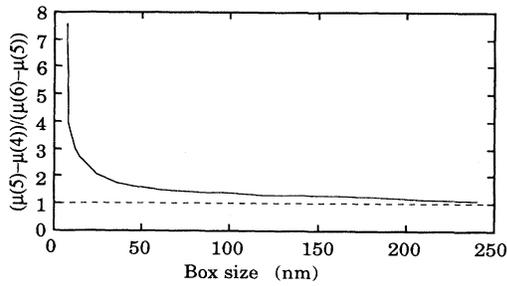


FIG. 5. Ratio of $[\mu(5) - \mu(4)]$ to $[\mu(6) - \mu(5)]$ vs the length of the longer side for a box with quasiparabolic confining potential and a 4:3 aspect ratio.

length of the box. The aspect ratio of the box is constant and again equal to 4:3. The confinement is obtained with hard walls and with a uniform charge distribution chosen for the different box sizes in such a way that the total positive charge in the dot is constant and equal to 1.602×10^{-17} C. As expected, R approaches 1 for dots of large size.

We remark that the results presented in Figs. 4 and 5 are rigorous for all dot sizes for our model, but that the data for the smallest dots (below 30 nm) represent the behavior of real dots only qualitatively. For very high confinement the third dimension of quantization may start playing a role, and energies high enough to populate higher valleys are involved.

The transition between the two limiting cases can be observed not only as a function of the dot size, but also as a function of the electron charge: for zero charge, and therefore no Coulomb interaction, the energy levels will be the single-electron levels and the peaks will merge in pairs. This is shown in Fig. 6, where the conductance peaks for a box identical to the one of Fig. 2 are reported for three values of a scaling factor β by which the electron charge is assumed to be multiplied. The hypothetical reduction of the electron charge is equivalent to a real situation: an increase, by a factor $1/\beta^2$ of the dielectric constant. This is a consequence of the fact that the Coulomb interaction is proportional to the square of the electron charge and to the reciprocal of the dielectric constant.

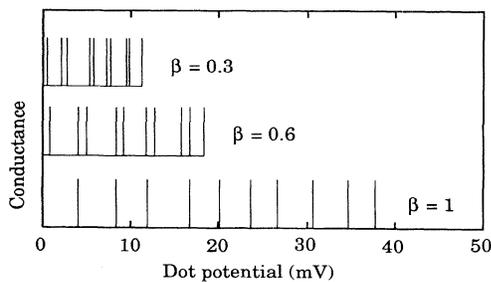


FIG. 6. Conductance peaks vs dot potential for a 180×135 nm box with hard walls and quasiparabolic confinement. The different curves are obtained reducing the electron charge by a factor β .

We have also investigated the effect of changes in the shape of the confining potential. In the limit of large dot size, only the fluctuations of the spacing between peaks, but not its average value, are influenced significantly as long as the dot area is kept constant. This, of course, is to be taken as a rule of thumb only. Complex effects of dot shape related to chaos theory have recently been discussed by Stone and Bruus.²⁶ Figure 7 shows results for the chemical potential (a) for a 60×45 nm box with flat potential and hard walls, (b) for a 60×45 nm hard wall box with additional quasiparabolic confinement, and (c) for a 42×31.5 nm flat potential and hard wall box. When the potential within the box is flat, the effective size of the dot coincides with the physical size of the box, since the wave functions are confined only by the hard walls. The effective sizes (areas where the electron density is significantly different from zero) of dots (b) and (c) are very close to each other and therefore the average spacing between peaks is almost the same, consistent with the fact that the average Coulomb energy depends on the average mutual distance between charges. The relative spacing between peaks still depends on the actual shape of the potential, which is different for cases (b) and (c).

Another important effect which may in part be caused by the specific shape of the potential has recently been reported by Ashoori *et al.*¹⁷ The spacing between peaks corresponding to the values of the chemical potential is uneven for the first few electrons and becomes uniform only for higher electron numbers. This may be due to the smaller effective dot size for small electron numbers; it also may be due to the influence of impurities as the following model calculation shows.

A model dot with quasiparabolic confinement and a 180×135 nm hard wall perimeter has been considered, to which two randomly placed impurity clusters have been added, with four and six positive charges, respectively. The additional charges create “holes” in the potential, which cause peaks in the wave functions of the first few electrons. The charge of the electrons in the lowest states screens the impurities and the potential appears more uniform to the electrons in higher states. In Fig. 8 we show results for the chemical potential of this dot for the

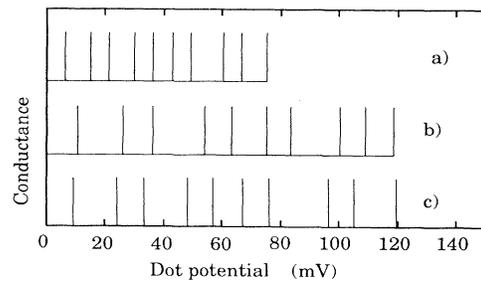


FIG. 7. Conductance peaks vs dot potential (a) for a 60×45 nm box with hard walls and flat potential, (b) for a 60×45 nm box with hard walls and quasiparabolic confinement, and (c) for a 42×31.5 nm box with hard walls and flat potential. The effective sizes of the dots for (b) and (c) are approximately equal.

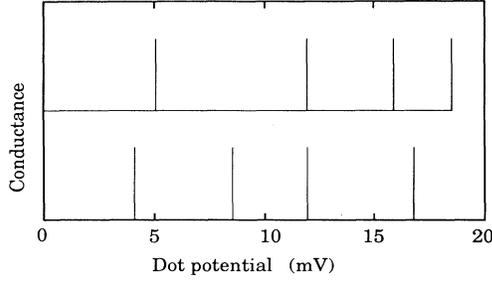


FIG. 8. Conductance peaks vs dot potential for a 180×135 nm box with hard walls and quasiparabolic confinement. The upper curve has been obtained with two randomly placed clusters of impurities, with four and six positive charges, respectively. The lower curve is for the same box without impurities. The reference level for the potential is assumed to be the one corresponding to the conductance peak for three electrons.

first few electrons: the upper graph is for the case just described, the lower one for the same dot without randomly distributed positive charges. The reference level for the dot potential is set to the position of the conductance peak for the third electron, as in the previous plots in this format. There are clear irregularities introduced in the upper plot by the presence of the impurities, irregularities similar to the ones observed in the experimental results of Ref. 17.

IV. EXCHANGE AND CORRELATION EFFECTS

As mentioned, we have included in our calculations the contributions of exchange and correlation effects within density-functional theory using the results of the TC theory with and without spin polarization. We now discuss the relative importance of these contributions to the chemical potential.

We perform a comparison between the results of a computation with the TC exchange and correlation potentials without spin polarization, which we have used for all the calculations described so far because of the numerical simplicity, and the results obtained from the Hartree approximation, in which the electron-electron interaction is restricted to the average Coulomb potential of all the other electrons. The pure Hartree approach is computationally more expensive, even though it is less accurate. A different Schrödinger equation needs to be solved for each orbital, since the Coulomb term reads

$$V_{ci}(\mathbf{r}_i) = \frac{1}{4\pi\epsilon_0\epsilon_r} \sum_{\substack{j=1 \\ j \neq i}}^N \int_S \frac{q^2 |\psi_j(\mathbf{r}_j)|^2}{|\mathbf{r}_i - \mathbf{r}_j|} d\mathbf{r}_j, \quad (6)$$

where j is the index of the electron whose wave function we are calculating.

In Fig. 9 we compare the results for the chemical potential obtained by including exchange and correlation contributions as given by the TC theory (upper curve) and by the Hartree approximation (lower curve). These results are for a box with a 60×45 nm hard wall perimeter,

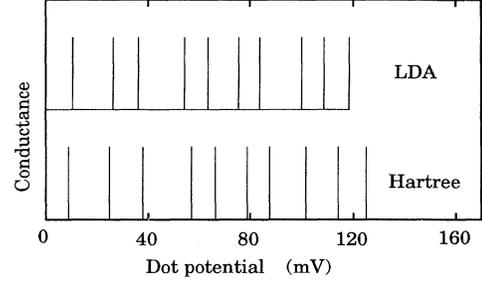


FIG. 9. Conductance peaks vs dot potential for a 60×45 nm box with hard walls and quasiparabolic confinement. The upper curve has been obtained by the TC local density functional (LDA) approximation, the lower one is the result with a purely Hartree approximation. The reference level for the potential is assumed to be the one corresponding to the conductance peak for three electrons.

in the presence of a quasiparabolic confining potential. The overall effect of the exchange and correlation contributions appears to be a uniform reduction in the values of the chemical potential with the ratios of the interpeak spacings substantially unaltered. Thus the TC corrections “subtract” from the total potential slightly more than what is subtracted by the removal of the electron self-interaction term in the Hartree approximation. This is consistent with the fact that the sign of the correlation term is negative and that the exchange term coincides, in the proximity of the electron, with the Coulomb self-interaction.²⁷ Numerical convergence is slightly improved by the presence of the exchange and correlation terms, because the overall potential term is reduced compared to the kinetic term.

A further improvement in the accuracy of our approach can be obtained from considerations of the effects of spin polarization. Even in the absence of magnetic fields, a slight spin polarization is present if the number of electrons on the dot is odd: the spin of the electron in the highest occupied orbital is not balanced. The importance of the unpaired electron becomes clearly smaller as the total number of electrons increases. We have used the spin polarized version of the Tanatar-Ceperley exchange potential (for the correlation potential we have included the nonpolarized expression because data are available only for no or full polarization). The spin polarized exchange term has been derived following a procedure analogous to the one described by Perdew and Zunger²² for the 3D Ceperley exchange.

Applying the expressions given in the Appendix has resulted in extremely small perturbations of the values for the chemical potential, smaller than a few percent for all previously reported dot sizes, even for just two or three electrons. Such effects would be important in the presence of magnetic fields, when significant spin polarization appears.

V. THE CAPACITANCE OF QUANTUM DOTS

The theoretical explanations of the periodic conductance oscillations of quantum dots have invariably in-

volved a capacitance of the dot which gives rise to an energy change $e^2/2C$ for the addition of one electron to the dot. The value of the capacitance C is usually estimated from simple approximations such as the self-capacitance of a conducting disk¹⁴ or the parallel plate capacitor formula.¹⁷ It is clear from the preceding sections that there exists a considerable self-capacitance of the dot, i.e., a finite change in chemical potential per electron addition even without the presence of the gate. Our investigations show that this indeed represents frequently the largest capacitance. It is then also clear that the capacitance is not constant, as the period in chemical potential shows both fluctuations and trends of decrease with increasing electron number. A prominent effect comes from changes in dot size due to the self-consistency of charge redistribution. When electrons are added, the dot size often increases as also found by Stopa.⁹ The fluctuations arise from various effects, including the form of the wave functions (size quantization), the proximity of impurities and corresponding changes in dot shape, and possible exchange and correlation contributions. The presence of gates and metallic objects (contacts) further modifies all these effects. As a consequence, the use of a constant capacitance and even the concept of capacitance need to be augmented as described below.

It is natural to associate a capacitance or, more appropriately, a differential capacitance to a quantum dot, since each increment of the charge stored in the dot causes an increment of the chemical potential. The differential capacitance $C = \Delta Q/\Delta V$ can immediately be obtained from the results for the chemical potential. For a single charge we have $q = \Delta Q$ and $[\mu(N+1) - \mu(N)]/q = \Delta V$. Thus $C_d = q^2/[\mu(N+1) - \mu(N)]$. In Sec. III we have seen that for relatively large box sizes the separation between conductance peaks (i.e., the increment in the chemical potential) does not fluctuate much, but shows a clear trend of a decrease with increasing electrons number.

In Fig. 10 we report dot capacitance vs electron number for a 160×120 nm box with hard walls and quasiparabolic confinement. The inset contains a 3D representation of the self-consistent potential for ten electrons. The capacitance tends to increase with the number of electrons, mainly because of the expansion of the dot. There is a clear trend towards a saturation:

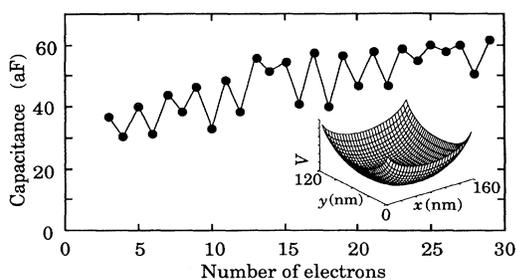


FIG. 10. Capacitance vs number of electrons for a 160×120 nm box with hard walls and quasiparabolic confinement. The self-consistent potential for ten electrons is shown in the inset.

for many electrons the capacitance should approach the one for a 2D metal disk having an area corresponding to the area of significant electron density. From the self-consistent wave functions, this area can be determined to be approximately circular, with a radius $R = 60$ nm. A metal disk of this size would have a self-capacitance of $8\epsilon_0\epsilon_r R = 55 \times 10^{-18}$ F,¹⁴ in agreement with the results. As the number of electrons increases, it becomes very difficult to reach convergence of the self-consistent iteration procedure, in particular in our case of zero temperature, when Newton-type schemes cannot advantageously be used. The particular sensitivity to numerical errors arrives from the fact that the capacitance is obtained through a differentiation of the chemical potential $\mu(N)$ of the system, therefore a small numerical noise in $\mu(N)$ leads to very large relative errors in $C(N)$.

The self-capacitance scales almost linearly with the effective radius of the dot, as in the case of a metal disk. In Fig. 11 we report the capacitance values for four boxes of different sizes vs the electron number. The solid squares represent the results for a 60×45 nm box, the solid circles for an 84×63 nm box, the empty circles for a 120×90 nm box, and the empty squares for a 160×120 nm box. The boxes have perimetral hard walls and there is further confinement provided by a quasiparabolic potential resulting from a uniformly distributed 1.602×10^{-17} C positive background charge.

Up to this point, we have considered only the self-capacitance of the dot, as if it were isolated in space, without other electrodes in its proximity. In the actual experimental situation, however, a conducting backgate and the electrodes (generating the confinement) contribute to the capacitance. An exact solution for a complex structure such as the ones experimentally investigated^{10,11,15,17} is beyond the scope of the present paper as we only wish to determine the relative importance of the various contributions and to give a qualitative understanding of the effects of external electrodes on the dot spectrum. Complete self-consistent solutions have been reported for particular cases as mentioned.⁹ Our approach to the inclusion of gates/contacts is based on the application of the method of images,²⁸ which provides a good estimate of the capacitance between the dot and a conducting backgate and an upper limit for the capacitance between the dot and a lateral lead.

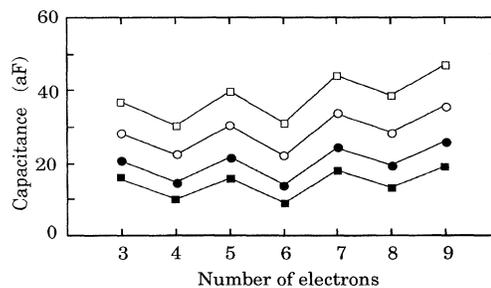


FIG. 11. Capacitance vs number of electrons for boxes with quasiparabolic confinement of different sizes: 60×45 nm (solid squares), 84×63 nm (solid circles), 120×90 nm (empty circles), and 160×120 nm (empty squares).

The backgate is usually much larger than the dot size, and we approximate it by an infinite conducting plane. The effect of this plane of constant potential is then equivalent to the introduction of an image dot placed at twice the distance. The image dot contains opposite charges and interacts with the real dot via the Coulomb force. From the analogy with the classical case we expect the capacitance to increase from the self-capacitance value with increasing gate proximity. We must remark that for a gate at a very short distance from the dot (relative to the dot size) and, therefore, for strong interaction with the image charges, the exchange and correlation terms in the Appendix are not valid anymore. For such cases we have used a simple Hartree approximation, and $\Delta\mu$ has been obtained through a double differentiation of the total energy $E(N)$ of the electron system, which is given by

$$E(N) = \sum_{i=1}^N \varepsilon_i - \frac{1}{2} \iint \frac{q^2}{4\pi\epsilon_0\epsilon_r} \frac{n(\mathbf{r})n(\boldsymbol{\rho})}{|\mathbf{r}-\boldsymbol{\rho}|} d\mathbf{r}d\boldsymbol{\rho} + \frac{1}{2} \iint \frac{q^2}{4\pi\epsilon_0\epsilon_r} \frac{n(\mathbf{r})n^*(\boldsymbol{\rho})}{|\mathbf{r}-\boldsymbol{\rho}|} d\mathbf{r}d\boldsymbol{\rho}, \quad (7)$$

where the second term on the right hand side corrects for the double counting of the electron-electron interaction, as in Eq.(4), and the third term for the double counting of the interaction between the electrons and their images, whose density is indicated with $n^*(\mathbf{r})$.

In Fig. 12 we report the results for the capacitance vs the number of electrons for a 120×90 nm box with hard walls and quasiparabolic confinement. The squares indicate the self-capacitance of the dot and the circles represent the capacitance values with a backgate at 60 nm (solid circles) or at 30 nm (empty circles). Overall, the expected increase occurs for all electron numbers. Moving the gate even closer to the dot causes a further increase of capacitance and an interesting additional effect appears, as shown in Fig. 13. Here the values of the chemical potential are shown (for a 160×120 nm box) in the usual conductance peak format. The dot to gate distance is now (a) 6 nm, (b) 8 nm, (c) 16 nm, and (d) infinity. For infinite distance the spacing between conductance peaks is almost even. However, for the very short gate distances, pairing between peaks becomes apparent as it does in Fig. 6 for reduced Coulomb interaction. This is,

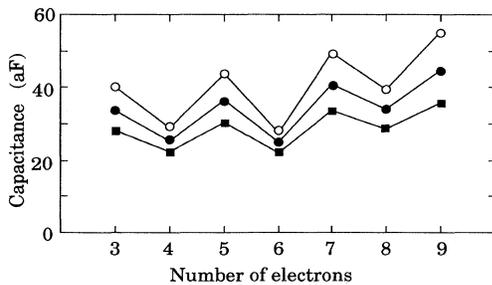


FIG. 12. Capacitance vs number of electrons for a 120×90 nm box with hard walls and quasiparabolic confinement with no gate (solid squares), with a backgate at 60 nm (solid circles), and with a backgate at 30 nm (empty circles).

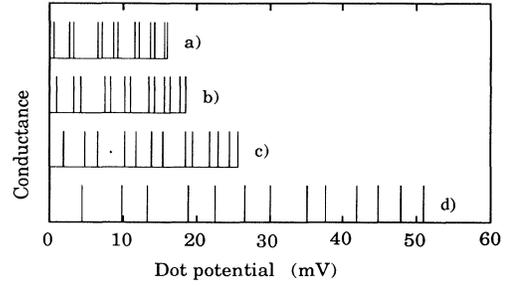


FIG. 13. Conductance peaks vs dot potential for a 160×120 nm box with hard walls and quasiparabolic confinement. The different curves are for a backgate at a distance of (a) 6 nm, (b) 8 nm, (c) 16 nm, and (d) for no backgate. The reference level for the potential is assumed to be the one corresponding to the conductance peak for three electrons.

of course, the consequence of the reduction of the relative importance of the Coulomb energy compared to the kinetic component, due to the presence of the image charge in close proximity of each electron. If the image charge, which has opposite sign, is at a distance from the corresponding electron comparable to the average interelectronic separation in the dot, the electrostatic interaction will be reduced and, in the limit of zero dot-gate distance, totally suppressed. Therefore, for extremely close backgates, significant quantum effects would be noticeable even in large dots. This case is not completely hypothetical; something analogous should be realized in quantum wells containing large and equal numbers of electrons and holes as they are present in semiconductor lasers.

It is interesting to see how the capacitance for a given number of electrons varies as a function of the dot size in the absence and in the presence of a conducting gate. In Fig. 14 we show the capacitance at $N = 4$ vs the length of the longer side of a quantum box. The lowest curve is for an isolated dot, without any gate: the capacitance growth is almost exactly linear with the size, as in the classical case. The second lowest curve is for a gate at a distance d of 60 nm. The capacitance increase is still linear in this case. A superlinear behavior appears in the third curve and becomes more pronounced for the shorter distance of 5 nm, approaching a well known limit: the parallel plates capacitor, in which the capacitance is proportional to the surface area of the electrodes and thus

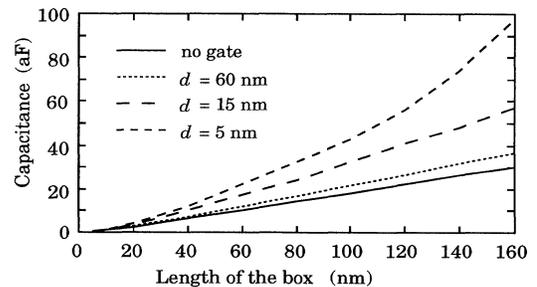


FIG. 14. Capacitance vs box length for a backgate placed at various distances. The confining potential is quasiparabolic and the number of electrons in the box is constant ($N = 4$).

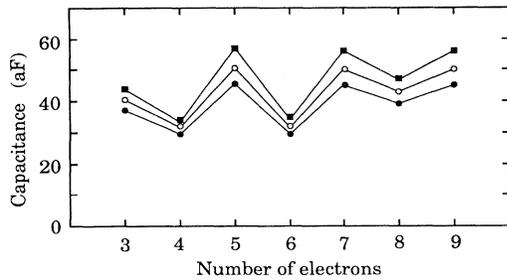


FIG. 15. Capacitance vs number of electrons for a 160×120 nm box with hard walls and quasiparabolic confinement with no lateral wall (solid circles) and with a conducting lateral wall at 80 nm (empty circles) and at 5 nm (solid squares).

to the square of the dot size (effective radius). This limit, however, is never reached in the known experimental situations, which deal with gate-to-dot distances that are of the same order of magnitude as the dot size, and therefore much larger than what would be needed to validate the parallel plate capacitor approximation.

Even though data have been reported for dot sizes down to zero, it is to be noted that they are exact for our model dot, but the assumptions behind the ability of our model to represent real quantum dots may break down for sizes below 20–30 nm.

Regarding the dot-to-lead capacitance, we have computed an upper limit corresponding to the capacitance between the dot and a vertical conducting wall located next to (and perpendicular to) the dot. Again we use the method of images, placing now an image box in the same plane as the original box, mirror symmetric with respect to the conducting plane. The overall effect of this coplanar image box is not very large, even for very short distances between dot and lead. Results are presented in Fig. 15, where the capacitance of a 160×120 nm quantum box is plotted vs the number of electrons for no lateral wall (solid circles), for a lateral wall at 80 nm (empty circles), and at 5 nm (solid squares). Even for the case with the wall at the shortest distance, the relative increase of the capacitance is not large, about 30% at most. We conclude that in most experimental situations the main contribution to capacitance comes from the self-capacitance, which grows almost linearly with the dot size (effective radius), just as the self-capacitance of a 2D metal plate. Additional contributions to capacitance are mainly due to the backgate and also to the leads. These additions are typically not prevalent and the usage of the parallel plate capacitor formula to compute total dot capacitance is not even qualitatively justified.

VI. CONCLUSIONS

The electronic structure of a model 2D quantum dot has been investigated, taking into account the electron-electron interaction and many-body effects based on density functional theory. We have been mainly interested in the numerical computation of the chemical potential and of the differential capacitance of the dot as a function of its size and of the number of electrons.

Our results show a gradual transition from a behavior

dominated by quantum effects in small dots to an almost classical, capacitorlike behavior of large dots. Quantum effects become important when the quantization energy is of significance compared to the Coulomb energy. This happens in very small dots, since the quantization energy scales with the reciprocal square of the dot size L , while the Coulomb energy is linear in $1/L$. The predominance of quantum effects also should occur in materials with high dielectric constant, because of electrostatic screening.

We have investigated the various contributions to the dot capacitance from nearby electrodes and concluded that they give only corrections to the self-capacitance which represents the dominant term in the geometries usually used in experiments. The parallel plate capacitor formula represents an inadequate approximation for these structures.

Our model dot reproduces typical features observed in the experiments, such as the capacitance increase associated with the addition of more electrons on the dot and the irregularities in the spacings between the values of the chemical potential when randomly distributed impurities are present.

Many-body effects, included by means of the exchange and correlation terms, affect the capacitance values and may play a more prominent role in more complex and smaller structures.

Overall, a rich variety of quantum effects emerges, which prompts further theoretical and experimental investigations. In particular, quantum dots with sizes below 30 nm or dots of materials with higher dielectric constant than the one of gallium arsenide may provide extremely interesting opportunities of testing the validity of the approach to the many-body problem.

Note added in proof. It has come to our attention that extensive work on mesoscopic capacitors has been performed by M. Büttiker and co-workers and will be published in Phys. Lett. A.

ACKNOWLEDGMENTS

We thank R. M. Martin, P. von Allmen, U. Ravaioli, D. M. Ceperley, and P. D. Yoder for valuable discussions, A. T. Galick for providing the Ritz iteration eigenvalue solver, and J. Larson for his help in the implementation of numerical algorithms on the Cray Y-MP. This work was supported by the U.S. Army Research Office and the Office of Naval Research. Cray Research is gratefully acknowledged for providing a grant for computer time on the Cray Y-MP at the National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign. One of the authors (M. M.) also acknowledges support from the National Research Council (CNR) of Italy (MADESS project) and from the Italian Ministry of University and Scientific and Technological Research.

APPENDIX

This appendix deals with the derivation of the exchange and correlation potential terms to be used in the

Schrödinger equation from the polynomial expressions of the exchange and correlation energies obtained for a two-dimensional electron gas by Tanatar and Ceperley.²¹ Energies and potentials are expressed in scaled rydbergs, therefore the MKS value can be obtained multiplying by the conversion factor $2.17989 \times 10^{-18} m^*/(mc_r^2)$, where m is the mass of the free electron. The expression given in Ref. 21 for the exchange energy in the absence of spin polarization is

$$E_{\text{ex}} = -\frac{8\sqrt{2}}{3\pi} \frac{1}{r_s} \quad \text{with } r_s = \frac{a}{a_0}, \quad a = \frac{1}{\sqrt{\pi\rho}}, \quad (\text{A1})$$

where a_0 is the Bohr radius and ρ is the local electron density.

For the correlation energy we have

$$E_{\text{corr}} = -C_0 \frac{1 + C_1 w}{1 + C_1 w + C_2 w^2 + C_3 w^3}, \quad w = \sqrt{r_s}$$

$$C_0 = -0.3578,$$

$$C_1 = 1.13, \quad C_2 = 0.9052, \quad C_3 = 0.4165. \quad (\text{A2})$$

V_{ex} and V_{corr} are obtained by taking the functional derivatives²⁵

$$V_{\text{ex}} = \frac{d(\rho E_{\text{ex}})}{d\rho}, \quad V_{\text{corr}} = \frac{d(\rho E_{\text{corr}})}{d\rho}. \quad (\text{A3})$$

Therefore we have

$$V_{\text{ex}} = -\frac{4\sqrt{2}}{\pi} \frac{1}{r_s} \quad (\text{A4})$$

and

$$V_{\text{corr}} = -C_0 \frac{1 + d_1 w + d_2 w^2 + d_3 w^3 + d_4 w^4}{(1 + C_1 w + C_2 w^2 + C_3 w^3)^2},$$

$$d_1 = 2.26, \quad d_2 = 2.635,$$

$$d_3 = 2.007, \quad d_4 = 0.70597. \quad (\text{A5})$$

When the topmost orbital is occupied by an unpaired electron, a further refinement in the calculation can be obtained considering the spin polarized version of the TC exchange energy:

$$E_{\text{ex}} = -\frac{4\sqrt{2}}{3\pi r_s} \left[(1 + \xi)^{3/2} + (1 - \xi)^{3/2} \right], \quad (\text{A6})$$

where ξ represents the degree of polarization, i.e., $\xi = (\rho_{\uparrow} - \rho_{\downarrow})/(\rho_{\uparrow} + \rho_{\downarrow})$, with ρ_{\uparrow} and ρ_{\downarrow} being the electron densities for spin up and spin down, respectively. By taking the proper functional derivatives we obtain the expressions for the spin-polarized exchange potential

$$V_{\text{ex}\uparrow} = -\frac{2\sqrt{2}}{\sqrt{\pi}} a_0 \left\{ \sqrt{\rho_{\uparrow} + \rho_{\downarrow}} \left[(1 + \xi)^{3/2} + (1 - \xi)^{3/2} \right] \right. \\ \left. + (\rho_{\uparrow} + \rho_{\downarrow})^{3/2} \frac{2\rho_{\downarrow}}{(\rho_{\uparrow} + \rho_{\downarrow})^2} \left[\sqrt{1 + \xi} - \sqrt{1 - \xi} \right] \right\}, \quad (\text{A7})$$

$$V_{\text{ex}\downarrow} = -\frac{2\sqrt{2}}{\sqrt{\pi}} a_0 \left\{ \sqrt{\rho_{\uparrow} + \rho_{\downarrow}} \left[(1 + \xi)^{3/2} + (1 - \xi)^{3/2} \right] \right. \\ \left. - (\rho_{\uparrow} + \rho_{\downarrow})^{3/2} \frac{2\rho_{\uparrow}}{(\rho_{\uparrow} + \rho_{\downarrow})^2} \left[\sqrt{1 + \xi} - \sqrt{1 - \xi} \right] \right\}. \quad (\text{A8})$$

- ¹G. W. Bryant, Phys. Rev. Lett. **59**, 1140 (1987).
²L. I. Glazman and R. I. Shekhter, J. Phys. Condens. Matter **1**, 5811 (1989).
³A. Kumar, S. E. Laux, and F. Stern, Phys. Rev. B **42**, 5166 (1990).
⁴Y. Meir, N. S. Wingreen, and P. A. Lee, Phys. Rev. Lett. **66**, 3048 (1991).
⁵A. Groshev, T. Ivanov, and V. Valtchinov, Phys. Rev. Lett. **66**, 1082 (1991).
⁶D. V. Averin, A. N. Korotkov, and K. K. Likharev, Phys. Rev. B **44**, 6199 (1991).
⁷P. A. Maksym and T. Chakraborty, Phys. Rev. Lett. **65**, 108 (1990).
⁸W. Häusler, B. Kramer, and J. Masek, Z. Phys. B **85**, 435 (1991).
⁹M. Stopa (unpublished).
¹⁰U. Meirav, M. A. Kastner, and S. J. Wind, Phys. Rev. Lett. **65**, 771 (1990).
¹¹L. P. Kouwenhoven, N. C. van der Vaart, A. T. Johnson, W. Kool, C. J. P. M. Harmans, J. G. Williamson, A. A. M. Staring, and C. T. Foxon, Z. Phys. B **85**, 367 (1991).
¹²U. Meirav, P. L. McEuen, M. A. Kastner, E. B. Foxman, A. Kumar, and S. J. Wind, Z. Phys. B **85**, 357 (1991).
¹³R. C. Ashoori, H. L. Störmer, J. S. Weiner, L. N. Pfeiffer, S. J. Pearton, K. W. Baldwin, and K. W. West, Phys. Rev. Lett. **68**, 3088 (1992).
¹⁴H. van Houten, C. W. J. Beenakker, and A. A. M. Staring, in *Single Charge Tunneling*, edited by H. Grabert and M. H. Devoret (Plenum Press, New York, 1992).
¹⁵Y. Hirayama, A. D. Wieck, T. Bever, K. von Klitzing, and K. Ploog, Phys. Rev. B **46**, 4035 (1992).
¹⁶R. Dingle, H. L. Störmer, H. C. Gossard, and W. Wiegman, Appl. Phys. Lett. **33**, 665 (1978).
¹⁷R. C. Ashoori, H. L. Störmer, J. S. Weiner, L. N. Pfeiffer, K. W. Baldwin, and K. W. West (unpublished).
¹⁸K. K. Likharev, IBM J. Res. Devel. **32**, 144 (1988).
¹⁹I. O. Kulik and R. I. Shekhter, Zh. Eksp. Teor. Fiz. **68**, 623 (1975) [Sov. Phys. JETP **41**, 308 (1975)].
²⁰Craig S. Lent, in *Proceedings of the International Symposium on Nanostructures and Mesoscopic Systems, Santa Fe, 1991*, edited by Wiley P. Kirk and Mark A. Reed (Academic, Boston, 1991); Leonard F. Register and Karl Hess, *ibid.* Note that the term Coulomb blockade is limited to a classical capacitance picture.
²¹B. Tanatar and D. M. Ceperley, Phys. Rev. B **39**, 5005 (1989).
²²J. P. Perdew and Alex Zunger, Phys. Rev. B **23**, 5048 (1981).
²³J. B. Krieger, Y. Li, and G. J. Iafrate, Phys. Rev. A **45**, 101 (1992).
²⁴N. A. Harrison, *Solid State Theory* (Dover, New York, 1980), p. 76.
²⁵W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
²⁶A. D. Stone and H. Bruus, Physica B **189**, 43 (1993).
²⁷J. C. Slater, Phys. Rev. **81**, 385 (1951).
²⁸J. D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1962), p. 26.