Electronic structure and many-body effects in self-assembled quantum dots

1999 J. Phys.: Condens. Matter 11 5953

(http://iopscience.iop.org/0953-8984/11/31/305)

J. Phys.: Condens. Matter **11** (1999) 5953–5967. Printed in the UK PII: S0953-8984(99)01252-7

Electronic structure and many-body effects in self-assembled quantum dots

Jean-Pierre Leburton†‡, Leonardo R C Fonseca‡, Satyadev Nagaraja†‡, John Shumway‡§, David Ceperley‡§ and Richard M Martin‡§

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

‡ Beckman Institute of Advanced Science & Technology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

§ Department of Physics, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

Received 29 January 1999

Abstract. A detailed model for the electronic properties of self-assembled InAs/GaAs quantum dots (SADs) is presented, with emphasis on inter-level transitions and many-body effects. The model is based on the self-consistent solution of three-dimensional Poisson and Schrödinger equations within the local (spin-) density approximation. Nonparabolicity of the band structure and a continuum model for the strain between GaAs and InAs results in position- and energy-dependent effective mass. The electronic spectra of SADs of various shapes have been determined with intraband level transitions and mid-infrared optical matrix elements. Shell structures obeying Hund's rule for various occupation numbers in pyramidal SADs agree well with recent capacitance measurements. It is shown that many-body interactions between orbital pairs of electrons are determined in a first approximation by classical Coulomb interaction.

1. Introduction

Self-organized islanding of InAs on a GaAs substrate induced by lattice mismatch between the two semiconductors is currently the most powerful and popular technique for the fabrication of high-density, high-quality, and relatively regular quantum dot structures. The high confinement present in these self-assembled structures make them ideal candidates for investigation in studying the physics of highly confined few-electron systems. From a practical standpoint, proposed applications of quantum dots such as in lasers [1], spectral detectors [2], and optical memories [3] have contributed to the proliferation of interest in these zero-dimensional systems.

Very recent studies have demonstrated the electronic shell structure and spin effects in lithographically defined quantum dots containing more than 40 extra electrons [4]. These 'artificial atoms' have shown shell filling with spin ordering predicted by Hund's rule. In other experiments using small self-assembled quantum dots (SADs) containing up to six extra electrons [5], the shell structure in the dots displayed an energy spectrum very different from the simple Coulomb blockade picture observed in metallic and mesoscopic structures. The combination of capacitance and far-infrared spectroscopy in these experiments [5] has provided experimental information on level separation, as well as on the electron–electron interaction energy of the s and p states within a single InAs–GaAs quantum dot.

In this study, we investigate the quantum mechanical properties and electron–electron interaction within single quantum dots [6–8], using a realistic structure with electron interactions treated within the density functional theory (DFT) [9]. The structures are complete

0953-8984/99/315953+15$30.00

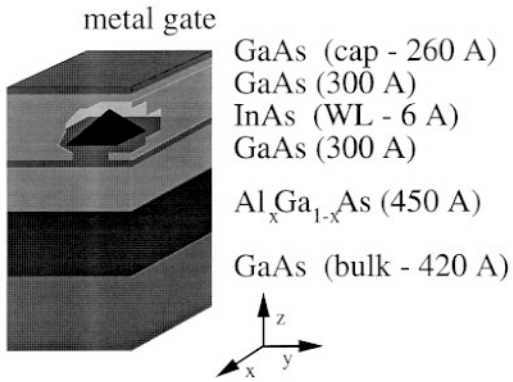

Figure 1. A schematic representation of the SAD used in the present work.

multilayer devices containing an InAs SAD embedded in a GaAs matrix (figure 1). The number of electrons in the dots is controlled by applying a voltage to a metal gate on top of the device. The strain in the SAD, wetting layer, and surrounding GaAs matrix is calculated using a continuum model described previously [9, 10]. The bulk electron effective mass and the band diagram are considerably modified by the strain in the region of the dots, becoming position dependent. In order to accurately determine the bias voltage at which charging occurs, we have used the concept of transition states. We have used the local spin-density approximation (LSDA) to calculate the many-body interaction [11] and the energies of possible spin configurations in the dot, showing that the dot filling indeed follows Hund's rule.

2. The quantum dot model and electronic structure

In the DFT framework, the three-dimensional (3D) Schrödinger equation, written in the effective-mass approximation (EMA), reads [9]

$$\left\{-\frac{\hbar^2}{2}\nabla\left[M^{-1}\nabla\right]+V(\boldsymbol{r})\right\}\psi_n(\boldsymbol{r})=E_n\psi_n(\boldsymbol{r}) \tag{1}$$

where M is the electron effective-mass tensor and the potential energy V is given by

$$V(\boldsymbol{r})=V_{\rm ext}(\boldsymbol{r})+V_{\rm off}(\boldsymbol{r})+V_{\rm c}(\boldsymbol{r})+V_{\rm P}(\boldsymbol{r}) \tag{2}$$

where $V_{\rm ext}(\boldsymbol{r})$ is the potential due to an externally applied voltage, $V_{\rm off}(\boldsymbol{r})$ is the conduction band offset, $V_{\rm c}(\boldsymbol{r})$ is the conduction band strain potential, and $V_{\rm P}(\boldsymbol{r})$ is the piezoelectric potential. Due to strain, the electron effective mass becomes anisotropic, leading to a mass tensor given by $\mathrm{diag}(M)=(m_{xx}m_{yy}m_{zz})$ and zero off-diagonal terms [12]. In the usual case of sample growth along the crystal direction (001), the electron masses along the plane perpendicular to the growth direction are equal, i.e. $m_{xx}=m_{yy}$.

Figure 1 shows a single-quantum-dot device which consists of a highly doped (10^{18} cm^{-3}) 420 Å GaAs substrate, followed by a 450 Å $Al_{0.3}Ga_{0.7}As$ barrier layer. The active region consists of two layers of undoped GaAs, 300 Å wide, surrounding a 6 Å InAs wetting layer and a InAs pyramid. A highly doped (10^{18} cm^{-3}) 260 Å GaAs cap and a metallic gate complete the device. We have assumed a conduction band offset $\Delta E_{\rm c}=770$ meV between bulk GaAs and bulk InAs (ratio $\Delta E_{\rm c}/\Delta E_{\rm g}=70\%$), and bulk electron effective masses in GaAs and InAs of 0.067 $m_{\rm e}$ and 0.023 $m_{\rm e}$ [13], respectively, where $m_{\rm e}$ is the bare electron mass. As described below, these effective masses and the conduction band offset change considerably as

strain is considered in the calculation (the calculation of strain is indispensable for an accurate simulation of the InAs–GaAs heterostructure due to its rather large lattice mismatch, of the order of 7%). All of the calculations correspond to 4.2 K.

The strain tensor is obtained from the minimization of the elastic energy of the system [14]. This procedure provides the strain tensor components ϵ_{xx}, ϵ_{yy}, and ϵ_{zz}, as well as the shear components ϵ_{xy}, ϵ_{xz}, and ϵ_{yz}. The hydrostatic and biaxial components of the strain, defined as

$$\epsilon_{\mathrm{h}}(\boldsymbol{r}) = \epsilon_{xx}(\boldsymbol{r}) + \epsilon_{yy}(\boldsymbol{r}) + \epsilon_{zz}(\boldsymbol{r}) \tag{3}$$

$$\epsilon_{\mathrm{b}}(\boldsymbol{r}) = \epsilon_{xx}(\boldsymbol{r}) + \epsilon_{yy}(\boldsymbol{r}) - 2\epsilon_{zz}(\boldsymbol{r}) \tag{4}$$

respectively play a major role in the electronic structure of the dot. Ignoring the split-off bands, one can derive the band-edge energies at the Brillouin zone centre ($\boldsymbol{k} = 0$) [13]:

$$\begin{aligned} V_{\mathrm{c}}(\boldsymbol{r}) &= E_{\mathrm{g}} + a_{\mathrm{c}}\epsilon_{\mathrm{h}}(\boldsymbol{r}) \\ V_{\mathrm{hh}}(\boldsymbol{r}) &= a_{\mathrm{v}}\epsilon_{\mathrm{h}}(\boldsymbol{r}) + \frac{b}{2}\epsilon_{\mathrm{b}}(\boldsymbol{r}) \\ V_{\mathrm{lh}}(\boldsymbol{r}) &= a_{\mathrm{v}}\epsilon_{\mathrm{h}}(\boldsymbol{r}) - \frac{b}{2}\epsilon_{\mathrm{b}}(\boldsymbol{r}) \end{aligned} \tag{5}$$

where E_{g} is the unstrained band-gap energy, and V_{hh} and V_{lh} are the heavy-hole and light-hole bands, respectively. The deformation potentials a_{c}, a_{v}, and b for InAs and GaAs are given in reference [9].

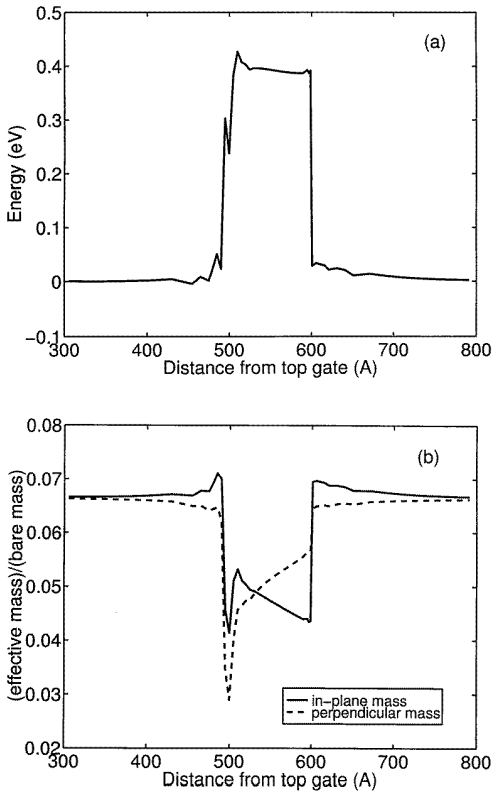

Figure 2. (a) The strain potential; (b) perpendicular and in-plane electron effective masses along the z-direction through the tip of the pyramid (in units of the bare electron mass). The pyramid base size is 200 Å and the height is 100 Å.

Figures 2(a) and 2(b) show the strain potential and electron effective masses along the z-direction, through the tip of the pyramid. Notice the large value of the strain potential which reduces the conduction band offset between the two materials, and the considerable modification of the bulk effective mass, by a factor of two on average, which is in good agreement with the work of Cusack *et al* [15]. The spikes near the tip of the pyramid have a numerical origin, and are caused by the difficulty in calculating the strain components around the edges of the pyramid.

The presence of the shear strains in the InAs–GaAs interfaces leads to the appearance of a polarization charge and its associated piezoelectric potential, which reduces the symmetry of the system, lifting some of the degeneracies calculated for unstrained pyramidal quantum dot systems. Piezoelectric effects in single dots of realistic sizes are very small, changing the eigenvalues of the system by less than 1 meV, and can be neglected [16]. However, they may be considerably larger in systems of closely spaced dots since their amplitude is different in each dot (see reference [12]).

Using time-independent perturbation theory up to the second order, we obtain the following expressions for the diagonal components of electron effective-mass tensor [12]:

$$\begin{aligned} m_z^*(E_i, \boldsymbol{r}) &= \frac{m^*(E_i - V_{\mathrm{lh}}(\boldsymbol{r}))}{E_{\mathrm{g}}} \\ m_{x,y}^*(E_i, \boldsymbol{r}) &= \frac{m^*(E_i - V_{\mathrm{hh}}(\boldsymbol{r}))(E_i - V_{\mathrm{lh}}(\boldsymbol{r}))}{E_{\mathrm{g}}(E_i - 0.75V_{\mathrm{lh}}(\boldsymbol{r}) - 0.25V_{\mathrm{hh}}(\boldsymbol{r}))} \end{aligned} \tag{6}$$

where E_i is the ith eigenvalue, and $m^*(E_i, \boldsymbol{r})$, $m_z^*(E_i, \boldsymbol{r})$, and $m_{x,y}^*(E_i, \boldsymbol{r})$ denote bulk, perpendicular, and in-plane electron effective masses (the wetting layer lies in the xy-plane) of an electron with associated eigenvalue E_i. The remaining components of the effective-mass tensor are zero. The presence of the eigenenergy E_i in equation (6) is for the correction of a large separation between the eigenstates and the conduction band edge. A detailed derivation of equation (6) including a comparison with an eight-band $\boldsymbol{k} \cdot \boldsymbol{p}$ calculation can be found elsewhere [17].

The Schrödinger equation is solved using the iterative extraction–orthogonalization method (IEOM) [18]. In this method a functional of the Hamiltonian is applied to a guess state iteratively to extract the ground state of the operator. Higher states are extracted by repeating the procedure for several initial guess states and orthogonalizing all the resulting states after each iteration of the Schrödinger equation. The major advantage of this method is its efficiency resulting from its ability to generate an arbitrarily small number of eigenstates N_E. As a result, the method scales as $N_E^2 N_G$, i.e., is linear in N_G, where N_G is the number of grid points. The Poisson equation is solved by combining Newton's technique with a successive over-relaxation method.

Figure 3 shows the first ten eigenvalues of empty dots as functions of the dot base length (also called the base diameter), keeping the wetting layer width fixed. The eigenvalues were calculated with respect to the average conduction band edge position in the barrier. Figure 3 also shows the approximate depth of the well. The notation $(n_i n_j n_k)$ corresponds to the number of nodes of the eigenfunction in the x-, y-, and z-directions, respectively. The eigenvalues were only calculated while the states remained bound or quasibound. Notice the slight splitting between the (100) and (010) states which is due to the piezoelectric effect. As the eigenvalues approach the top of the well, the slopes of the curves tend to decrease, as a result of deeper wavefunction penetration in the barrier region. For energies above the well height, the corresponding eigenfunctions spread over the wetting layer, becoming those of a two-dimensional electron gas.

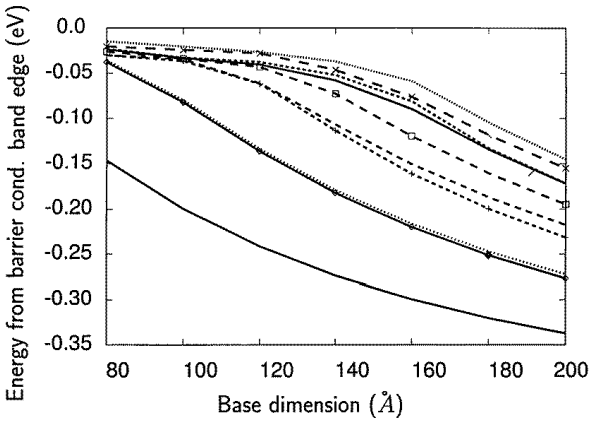

Figure 3. The approximate well depth (thick solid curve) and eigenvalues of the Hamiltonian as functions of the dot diameter d. The eigenvalues were calculated with respect to the average conduction band edge in the barrier. The first few states represented, from low to high energy, are (000), (100), (010), (110), (200) + (020), (200) − (020). The notation $(n_1n_2n_3)$ denotes the number of wavefunction nodes in the x-, y-, and z-directions. Pyramid height: $h = d/2$. The curves are guides for the eye.

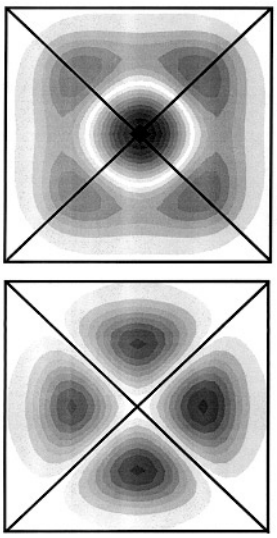

Figure 4. Projections of the states (200) + (020) (top) and (200) − (020) (bottom). Notice the lobes of the top state along the diagonals (crossed lines) of the base of the pyramid (square box), lowering its energy with respect to the more confined bottom state.

Ignoring the piezoelectric effect, the Hamiltonian of the quantum dot is invariant under the symmetry operations of the group C_{4v} [19], which allows the wavefunctions to be symmetrized according to the irreducible representations of this group. It can be easily verified that only p-like states (such as (100)) can be degenerate, while all other degeneracies are accidental. That explains why the states (200) and (020) are not degenerate in a finite-barrier pyramid, even though they are degenerate in a square well. In the latter case, degeneracy occurs because the Hamiltonian allows the separation of variables. One can also show that the irreducible

representations of those two states are the linear combinations (200) + (020) and (200) − (020), shown as the fifth and sixth curves going from the bottom to the top of figure 3 (the solid curve with stars and the dashed curve with crosses, respectively). Perhaps this counterintuitive result can be more easily understood with the help of figure 4. It shows that the projections of the states (200) + (020) and (200) − (020) are very different, and thus are affected differently by the pyramidal confining potential [9]. As a result, the energies of the two states do not need to be the same. Our calculations agree well with those of reference [15] for the range of sizes investigated in that work. As discussed previously, these eigenvalues may change considerably if a full eight-band $\boldsymbol{k} \cdot \boldsymbol{p}$ calculation is performed [17].

The quantum dot shape can sometimes be controlled during the experiment but, particularly in the case of SADs, the shape is a result of the growth process and must be measured. Scanning tunnelling microscopy (STM) and atomic force microscopy (AFM) [20], have shown that dots can grow as faceted pyramids or with rounded lens shapes. The shape is dependent on the growth process, and even shape transitions have been observed [20]. The reduced symmetry of square dots causes degeneracy splitting, so infrared absorption and photoconductivity experiments may distinguish between lens-like and pyramidal dots.

The 3D rotational symmetry of atoms leads to the quantization of angular momentum and description in terms of the quantum numbers l and m_l, or equivalently the notation s, p, d, f, ... Quantum dots have reduced symmetry and should not be labelled in this manner. Nevertheless, we will adopt the notation 1s, $2p_x$, $2p_y$, $3d_{xy}$, $3p_x$, $3d_{x^2-y^2}$, and 2s to describe the six lowest states so that our results may be compared to those in existing literature and as an aid in comparisons between lens-like and pyramidal symmetries. In effect, the states in a lens should be labelled by the irreducible representations of the C_∞ group [21]: $1a_1$ (1s), $1e_1$ ($2p_x$, $2p_y$),

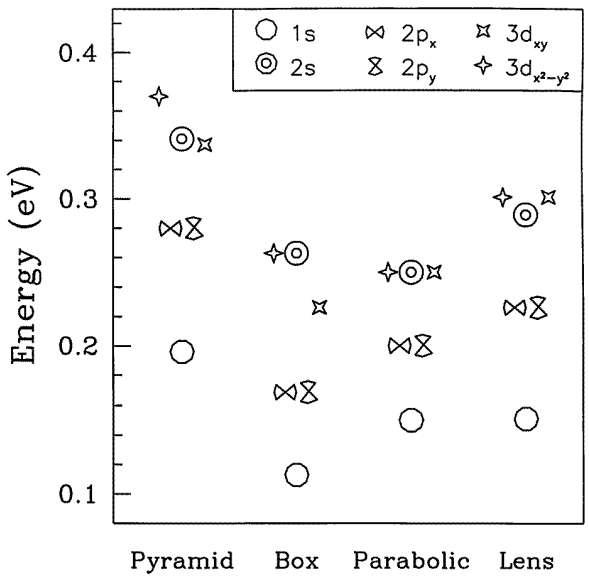

Figure 5. The single-particle energies of different-shaped potentials for 200 × 200 × 100 Å^3 dots with $m_e = 0.05$. The round lens ($D = 200$ Å, $h = 100$ Å) and parabolic dots ($w_x = w_y = 0.25 w_z$ meV) have degenerate d states, and the reduced symmetry of the pyramid and box cause the $3d_{xy}$ state to lie below the $3d_{x^2-y^2}$ state. The separable potentials for the box and parabolic potentials have an additional degeneracy between the 2s and $3d_{x^2-y^2}$ states, which is lifted by interactions.

$1e_2$ ($3d_{xy}$, $3d_{x^2-y^2}$), and $2a_1$ (2s), where the expressions in parentheses are the notation used by us. As mentioned before, for a pyramid the symmetry group is C_4, and proper labelling is $1a_1$ (1s), $1e_1$ ($2p_x$, $2p_y$) $1b_2$ ($3d_{xy}$, $3d_{x^2-y^2}$), and $2a_1$ (2s). The effect of symmetry reduction from a lens to a pyramid is to break the degeneracy of the d states, leading to a lower-energy state $3d_{xy}$ with lobes in the corners of the confining potential, and a higher-energy state $3d_{x^2-y^2}$ with nodes in the corners of the potential. The degeneracy in states with odd values of angular momentum is unaffected by a reduction to pyramidal symmetry. One weakness of our adopted notation is that the 2s state actually contains a linear combination of atomic-like 2s and $2p_z$ states, an effect which shows up in the z-polarized dipole selection rules.

Figure 5 shows the comparison of the lowest single-particle eigenstates for lens and pyramidal dots, as well as a couple of simple theoretical models: a hard-wall rectangular box and a parabolic potential. The box and parabolic potentials exhibit an accidental degeneracy, due to the separability of the degrees of freedom in the x-, y-, and z-directions in the single-particle Schrödinger equation for these potentials. The box and parabolic wavefunctions are often represented by the three quantum numbers (n_x, n_y, n_z), which refer to the numbers of nodes in the x-, y-, and z-directions, respectively. States (1, 0, 0) and (0, 1, 0) are degenerate, as are states (2, 0, 0) and (0, 2, 0). The (1, 0, 0), (0, 1, 0) pair corresponds to the ($2p_x$, $2p_x$) degenerate pair present in both the square and circular symmetry, but the (2, 0, 0), (0, 2, 0) pair must be recombined to get the $3d_{x^2-y^2} = [(2, 0, 0) + (0, 2, 0)]/\sqrt{2}$ and $2s = [(2, 0, 0) - (0, 2, 0)]/\sqrt{2}$ states. Thus these separable potentials contain an accidental degeneracy between the $3d_{x^2-y^2}$ and 2s states that is lifted by a slight change in the confining potential or the introduction of interactions [9].

The selection rules for dipole-allowed transitions are determined by the requirement that $\langle \Psi_i | \boldsymbol{r} \cdot \hat{\boldsymbol{e}} | \Psi_f \rangle$ be nonzero, where $\hat{\boldsymbol{e}}$ is the electron polarization operator, and Ψ_i and Ψ_f are the many-body initial and final states, respectively. For circular symmetry, the allowed transitions are $\Delta l_z = \pm 1$ for x-, y-polarizations and $\Delta l_z = 0$ for z-polarization. In the

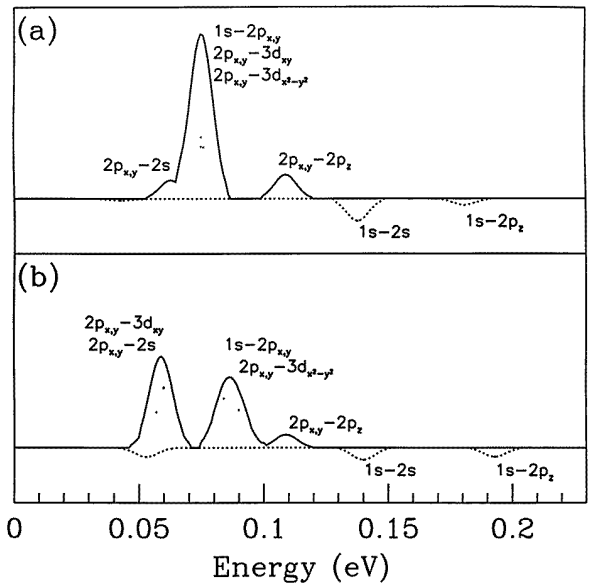

Figure 6. Computed dipole matrix elements for transitions between different states for lens- and pyramid-shaped quantum dots. Transitions from 3d states up have not been considered.

case of square symmetry, group theory gives the rules as e $\leftrightarrow$ $\{a_1, a_2, b_1, b_2\}$ because x-, y-polarizations and z-dipole transitions only occur between states of the same symmetry. In the case of charged quantum dots, the same transition rules for both shapes: x-, y-transitions, $2p_{x,y} \leftrightarrow \{1s, 3d_{xy}, 3d_{x^2-y^2}, 2s\}$, and one allowed z-transition, $1s \leftrightarrow 2s$, exist.

We show calculated IR absorption spectra for a pyramidal dot in figure 6. The Hartree energy shifts the energy of the 1s state more than those of other states, since this state is sited at the centre of the dot. The result is that the $1s \leftrightarrow 2p_{x,y}$ transition overlaps the energy of the transitions from the $2p_{x,y}$ to higher states [22]. This can give the appearance of a single transition, when, in fact, several transitions contribute to a spectral line. Also, it would be possible to incorrectly associate these evenly spaced transitions in this pyramidal dot as indicative of a parabolic confining potential. The prominent split of the energy for the transitions to the $3d_{xy}$ and $3d_{x^2-y^2}$ states in our calculations show the breaking of the axial symmetry in the dot.

3. Local spin-density approximation and many-body effects

The quantum mechanical correction of many-electron interactions in the context of device physics is calculated using the LSDA of the Kohn–Sham DFT [23]. In this approximation, the classical electrostatic repulsion among electrons (Hartree potential), and the quantum mechanical correction due to exchange and correlation are added to the total potential $V(\boldsymbol{r})$ in equation (2). The Hartree potential is obtained by solving self-consistently Poisson's equation (9), while the exchange–correlation energy $E_{\rm xc}[n^\alpha, n^\beta]$ is a function of the α-electron and β-electron densities $n^\alpha(\boldsymbol{r})$ and $n^\beta(\boldsymbol{r})$, where α and β denote up and down spins, respectively. The exchange term becomes not only a function of the total charge density $n = n^\uparrow + n^\downarrow$ but also a function of the polarization parameter ζ,

$$\zeta = \frac{n^\uparrow - n^\downarrow}{n}. \tag{7}$$

The Kohn–Sham equations (1) and (2) including many-body corrections only provide the ground state of the system and its total energy $E_{\rm T}$. A rigorous way of determining the number of electrons N in a quantum dot with the electron charge as a good quantum number is to minimize $E_{\rm T}(N)$, for $N = 1, 2, \ldots, N_{\max}$. This minimization should be repeated whenever the external voltage bias $V_{\rm g}$ is changed. The use of only eigenvalues to determine the charge in the dot, where charging occurs whenever an eigenvalue crosses the Fermi level, is only correct in the limit of weakly interacting electron systems.

However, the Kohn–Sham theory is not restricted to integer numbers of electrons in the system. Differentiating $E_{\rm T}$ with respect to the noninteger occupation number n_i of level i one obtains

$$\frac{\partial E_{\rm T}}{\partial n_i} = \epsilon_i. \tag{8}$$

Equation (8), the Janak theorem [24], provides a meaning for the eigenvalues of the Kohn–Sham equation. Integrating equation (8) between N and $N+1$ one obtains the so-called Slater formula [25]:

$$E_{\rm T}(N+1) - E_{\rm T}(N) = \int_0^1 \epsilon_{\rm LAO}(n)\, {\rm d}n \approx \epsilon_{\rm LAO}\left(\frac{1}{2}\right) \tag{9}$$

where $\epsilon_{\rm LAO}$ corresponds the eigenvalue of the lowest available orbital. The last step in equation (9) is exact if $\epsilon_{\rm LAO}$ is a linear function of the occupation number. In order to determine whether there are N or $N+1$ electrons in the dot, one defines the transition state as the state

containing $N+0.5$ electrons. If $\epsilon(\frac{1}{2})$ is positive, then $E_{\mathrm{T}}(N+1) > E_{\mathrm{T}}(N)$, and the dot contains N electrons; otherwise it contains $N+1$ electrons [9].

Figure 7 shows the number of electrons in a pyramidal quantum dot with a 200 Å base diameter and a height of 70 Å as a function of the gate voltage V_{g}. Two curves are shown, corresponding to charging sequences which obey and disobey Hund's rule. When Hund's rule is obeyed it means that the charging of the fourfold-degenerate second level follows the spin sequence $2p_x^{\uparrow}2p_y^{\uparrow}2p_x^{\downarrow}2p_y^{\downarrow}$. The curve that does not follow Hund's rule was obtained by charging the second level according to the spin sequence $2p_x^{\uparrow}2p_y^{\downarrow}2p_x^{\uparrow}2p_y^{\downarrow}$. The third possibility, namely the spin sequence $2p_x^{\uparrow}2p_x^{\downarrow}2p_y^{\uparrow}2p_y^{\downarrow}$, was not considered because it will clearly be unfavourable due to the intense Coulomb repulsion between the $2p_x^{\uparrow}$ and the $2p_x^{\downarrow}$ electrons resulting from their large wavefunction overlap. The step size obtained with the LSDA corresponding to the charging of the fourth electron ($N = 3$) is longer for the charging of the dot according to the spin sequence $1s^{\uparrow}1s^{\downarrow}2p_x^{\uparrow}2p_y^{\downarrow}$ than for $1s^{\uparrow}1s^{\downarrow}2p_x^{\uparrow}2p_y^{\uparrow}$, indicating that indeed Hund's rule is followed by this system. The electron–electron interaction energy difference between the two spin configurations for four electrons in the dot is $\sim$3 meV.

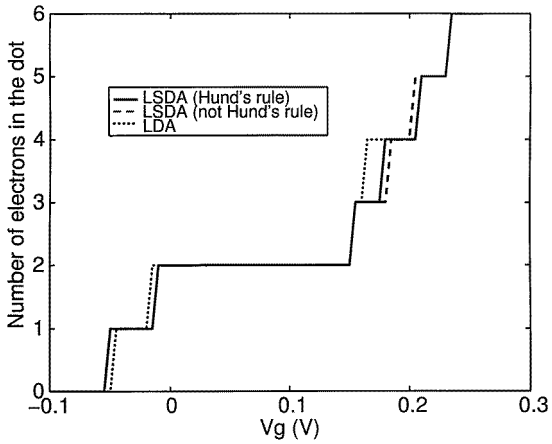

Figure 7. The number of electrons in the dot as a function of the gate voltage V_{g} in the case of obeying Hund's rule (the second-level population following the spin sequence $2p_x^{\uparrow}2p_y^{\uparrow}2p_x^{\downarrow}2p_y^{\downarrow}$), and disobeying Hund's rule (the second-level spin sequence is $2p_x^{\uparrow}2p_y^{\downarrow}2p_x^{\uparrow}2p_y^{\downarrow}$). The dotted curve coincides with the solid curve whenever the dots are not visible. Pyramid base size = 200 Å and height = 70 Å.

The addition of the fourth electron following Hund's rule is less costly because of the presence of exchange interaction (attractive) in this case, but not if the spin of the fourth electron is different from the spin of the third. Analogously, the step corresponding to the charging of the fifth electron ($N = 4$) is shorter for the sequence $2p_x^{\uparrow}2p_y^{\downarrow}2p_x^{\uparrow}2p_y^{\downarrow}$ because the fifth electron, either in the $2p_x^{\uparrow}$ or $2p_y^{\downarrow}$ state, interacts by exchange with one of the two electrons already in the second level, while, according to Hund's rule, the fifth electron does not interact by exchange with either of the other two because of their different spin states.

The local density approximation DFT has been widely used to calculate the electron–electron interaction in dots [9, 12, 18, 26–29] because of its simple implementation and negligible demand for computer time. However, the LSDA theory is an approximate theory, well known for high-accuracy calculations as well as erroneous predictions of the physical properties of some systems [30]. In particular, the experience with the LSDA for atoms and

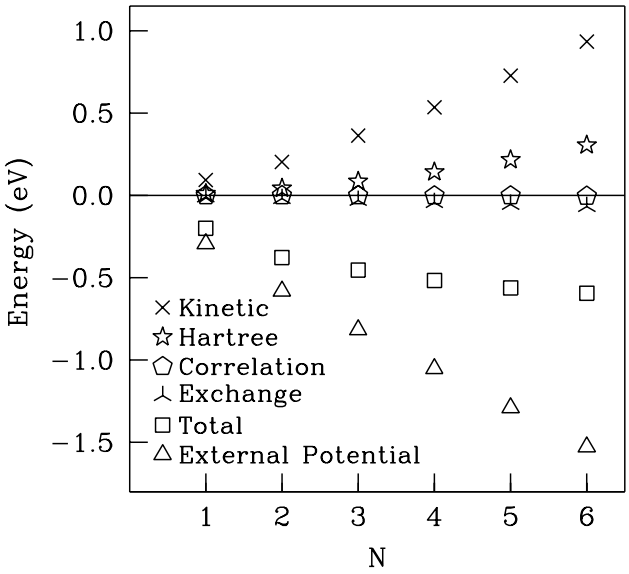

Figure 8. Contributions to the energy of a 200×100 Å pyramidal dot as a function of the occupation N, calculated both in the LSDA and using the quantum Monte Carlo (QMC) technique. Differences between the LSDA and QMC results are not apparent on the energy scale considered in this figure. This clearly shows the small effect of interactions beyond the Hartree (mean-field) level.

molecules may not carry over to quantum dots, since the confining potentials and electron densities can be very different. In figure 8 we compare the results obtained with the LSDA in pyramidal SADs against those obtained with the diffusion quantum Monte Carlo (DMC) method, which provides an exact treatment of the many-body interactions [31]. We choose to investigate this particular dot shape because its geometry is particularly challenging from a computational viewpoint; therefore our conclusions will be more general. Figure 8 shows the values of the different components of the total energy as a function of dot occupation. Differences between the LSDA and QMC results are not apparent on the scale of the figure. This figure clearly shows that the external potential energy and kinetic energy are much larger than the interactions. In other words the interactions may enter as a perturbative effect from the noninteracting system. The reason for this can be seen from the scaled electron density. Then energy and length scales for the electron interaction are scaled by the dielectric and mass, so the effective Bohr radius is $a_0^* = \epsilon/m^* a_0 \approx 150$ Å and the effective Hartree energy is $\mathrm{Ha}^* = \epsilon^{-2} m^*$ Ha ≈ 7 meV. If we approximate the electrons as uniformly occupying the interior of the dot, we obtain an effective electron density of $r_s \approx 0.46 N^{-1/3}$, which is rather high. In fact, the electron gas has a well-known ground-state energy expansion for small r_s [32]:

$$E = 2.2099 r_s^{-2} - 0.9163 r_s^{-1} - 0.094 + 0.0622 \ln(r_s) + \cdots \tag{10}$$

where the first term is the kinetic energy, the second term is the exchange, and the remaining terms refer to correlation energy. For the case of six electrons in the dot, $r_s \approx 0.25$, and the expansion gives $E_k = 1440$ meV, $E_x = 151$ meV, and $E_c = 0.6$ meV. Although the comparison between these very different electronic systems cannot be stretched too far, this does show that our energy scales are reasonable for a highly effective electron density. The leading effect of the interaction is the Hartree energy, with small corrections for exchange and still smaller for correlation.

4. Orbital-pair interactions

The interaction between pairs of electrons occupying different or the same spatial or spin state can also be determined from the total electron–electron interaction energies. In the case of the individual electrons occupying the same spatial state, the spin states are constrained to be different, whereas for different spatial states the individual spin states may be the same or different.

The total electron–electron interaction energy for N electrons,

$$E_{ee}(N) = E_H(N) + E_{xc}(N)$$

is the sum of the repulsive Hartree energy, $E_H(N)$, and the attractive exchange–correlation energy, $E_{xc}(N)$. In the single-particle picture, $E_{ee}(N)$ involves the sum of interactions between pairs of electrons in different single-particle states—the so-called orbital-pair interactions. For example, $E_{ee}(2)$, the interaction between two electrons occupying the s-type ground state, hereafter referred to as the s–s interaction, is the Hartree energy for $N = 2$, and is designated as E_{ss}. Since the two electrons are of opposite spins, the exchange interaction is absent. For $N = 3$ the many-body ground-state configuration has two electrons in the s state, and one in the p_x or p_y state, represented as $1s^{\uparrow}$, $1s^{\downarrow}$, $1p_x^{\uparrow}$. $E_{ee}(3)$ can then be written as

$$E_{ee}(3) = E_{ss} + E_{ps}^{A} + E_{ps}^{P} \tag{11}$$

where E_{ps}^{A} and E_{ps}^{P} denote, respectively, the interaction between the p and s electrons (p–s interaction) with anti-parallel and parallel spins. From the above equation, the average p–s interaction, $E_{ps} = (E_{ps}^{A} + E_{ps}^{P})/2$, can be computed if we assume that the E_{ss}-interaction calculated previously does not change with N in a first approximation. This is quite reasonable for SADs, since the strong confining potential and the symmetry of the s state would prevent any appreciable deformation of the s wavefunctions. The average p–s interaction is then given by

$$E_{ps} = (E_{ee}(3) - E_{ss})/2. \tag{12}$$

Following the above procedure, E_{pp}^{P} and E_{pp}^{A} denoting the average interaction between two p orbitals (the p_x–p_x or the p_x–p_y interactions) of parallel and anti-parallel spins, respectively, are given by

$$E_{pp}^{P} = E_{ee}(4) - E_{ss} - 4E_{ps} \tag{13}$$

$$E_{pp}^{A} = (E_{ee}(5) - E_{ss} - 6E_{ps} - E_{pp}^{P})/2 \tag{14}$$

where, E_{pp}^{P} used in equation (14) is that computed from equation (13). Alternatively, E_{pp}^{P} and E_{pp}^{A} can be determined from $E_{ee}(6)$, and are given by

$$E_{pp}^{P} = (E_{ee}(6) - E_{ss} - 8E_{ps} - 4E_{pp}^{A})/2 \tag{15}$$

$$E_{pp}^{A} = (E_{ee}(6) - E_{ss} - 6E_{ps} - 2E_{pp}^{P})/4 \tag{16}$$

where E_{pp}^{A} used in equation (15) is computed from equation (13), and E_{pp}^{P} in equation (16) is computed from equation (15). Following the above approach, interactions involving the next-higher d state, i.e., d–d, d–p and d–s, can also be computed if the d state is triply degenerate. This, however, is not the case in SADs since the degeneracy of the d state is lifted by the pyramidal symmetry of the confining potential [33, 34]. Hence, in the present work we do not go beyond the p–p interaction.

We have estimated pairwise electron–electron interaction energies from the DFT using equations (12)–(16). Table 1 summarizes those results and compares them with the values

Table 1. Charging energy per electron pair (in meV) in single SADs.

Interaction type	Calculated	Measured
s–s	22	~ 23
s–p	~ 15	~ 7
p–p	~ 13	~ 18

inferred from the capacitance data of Fricke *et al* [5]. Differences between energies obtained from calculations and inferred from measurements may result from our assumption that the different types of interaction (s–s, p–s, or p–p) remain unchanged as more electrons are added to the dot. A second source of error in our calculation is the exclusion of inter-dot repulsion, which should push the electrons closer together inside the dot. However, as we have already pointed out, for dot densities of 10^{10} cm^{-2}, the inclusion of inter-dot effects should change our results by less than 1 meV by pushing the electrons in the dot closer together. Finally, the approximation used to calculate E_{H} may lead to some correction in our calculation. As far as the analysis of the experimental data is concerned [5], it included the image charge effect but excluded inter-dot repulsion. In fact, these two competing effects nearly cancel each other for dot densities in the range (1–10) $\times$ 10^{10} cm^{-2}. The analysis of the experimental data also excluded the presence of a charged interface between the gates (the layer of dots charged with one electron each). Indeed, the presence of the layer of charged dots between the gates *decreases* the electron–electron interaction energy by $\sim\gamma_{\mathrm{dot}} \times 3$ meV, where $\rho_{\mathrm{dot}} = \gamma_{\mathrm{dot}} \times 10^{10}$ cm^{-2} is the density of dots in the plane. Because this correction is considerable and is linear in γ_{dot}, it is clear that the extraction of the electron–electron energy from the experimental data requires precise knowledge of the dot density.

Because approaches like ours can be computationally expensive and time consuming, we propose a simple approximation that requires the computation of only E_{ss}, with the remaining interactions being deduced from the former ones. Such an approximation might be useful for quick estimates of electron–electron pair interactions. The proposed method is based on the observation that in the EMA/LSDA framework [9] the Hartree energy of a quantum dot accounts for most of the electron–electron interaction energy, while the exchange–correlation energies account for a small correction after the self-interaction energy part of the exchange is subtracted [34] because the Hartree energy scales as aN^2, where the constant a, evaluated at $N = 2$, is $E_{\mathrm{ee}}(2)/4$; we can write

$$E_{\mathrm{ee}} \simeq E_{\mathrm{ss}}\left(\frac{N}{2}\right)^2. \tag{17}$$

Substituting for $E(N)$ from equation (17) in equations (12), (13), (14), (15), and (16), the following approximate forms for E_{ps}, $E_{\mathrm{pp}}^{\mathrm{P}}$, and $E_{\mathrm{pp}}^{\mathrm{A}}$ are obtained:

$$E_{\mathrm{ps}} \simeq 0.625E_{\mathrm{ss}} \tag{18}$$

$$E_{\mathrm{pp}}^{\mathrm{P}} = E_{\mathrm{pp}}^{\mathrm{A}} \simeq 0.5E_{\mathrm{ss}}. \tag{19}$$

In order to investigate the accuracy of the estimate, we compared the pair interactions obtained from the LSDA with those obtained from our approximations for four different SADs: three lens-shaped dots (a truncated sphere) whose base $\times$ height values were 150×75 Å^2 (dot A), 200×100 Å^2 (dot B), 400×200 Å^2 (dot C), and a pyramidal dot with base $\times$ height equal to 200×100 Å^2.

Figure 9 shows the variation of the interactions for the four SADs, obtained from equations (17)–(19) as well as from EMA/LSDA calculations. It is seen that, in all of the cases, the s–s

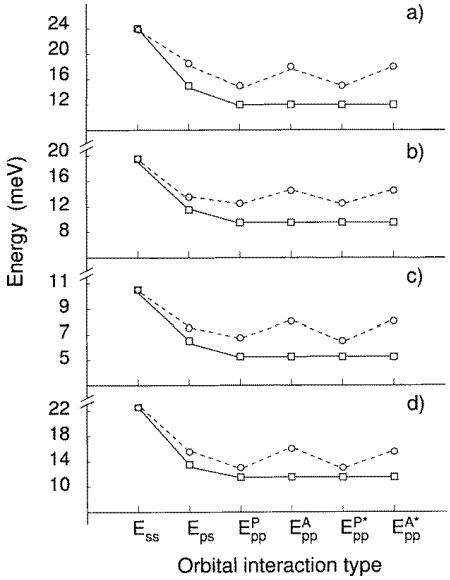

Figure 9. Pair interaction energies estimated from equations (18) and (19) (solid curves), and computed from equations (12)–(16) (dashed curves), for lens-shaped dots with diameter (D) × height (H) equal to (a) 150 Å × 75 Å, (b) 200 Å × 100 Å, and (c) 400 Å × 200 Å, and for (d) a pyramid-shaped dot with $D \times H = 150$ Å × 75 Å. E_{pp}^{P*} and E_{pp}^{A*} were computed using equations (15) and (16).

is the strongest of all of the interactions. This arises from the fact that the s wavefunction has a smaller spatial spread compared to, say, p or d wavefunctions, thus squeezing the s electrons. Our calculations show that the contribution of exchange–correlation effects to the orbital-pair interaction energies is small—less than 15% of E_{ss} for all of the dots. This is reasonable in high-electron-density systems such as SADs where the confinement is strong. It is easily seen that, for very small r_s, quantum many-body effects (the last three terms in equation (10)) constitute only a small fraction of the total energy which is dominated by the kinetic energy (the first term in equation (10)) of the order of 100 meV, the electron–electron repulsion being comparatively weak ($\simeq$20 meV) [9]. Nevertheless, the effect of (attractive) exchange interaction is seen clearly in the lower values of the p–p interaction for the parallel spin, while it is absent from the anti-parallel interactions.

As far as the interaction values estimated from our approximations (equations (17)–(19)) are concerned, two features are noticeable in figure 9: firstly, they are consistently lower than the interactions computed from LSDA calculations; and secondly, unlike the latter values, they are identical for both parallel and anti-parallel spins. The first feature arises from the nature of our approximation—the constant a in the expression $E_{ee}(N) = aN^2$ underestimates the interactions for $N > 2$. This can be explained noting the fact that a better approximation, accounting for the exchange and correlation effects, is

$$E(N) = (\alpha - \beta N^{-2/3})N^2 = \gamma(N)N^2$$

with $\gamma(N) \to \alpha$ for large N, and $\alpha, \beta > 0$. Hence from the substitution $a = \alpha - \beta/2^{2/3}$ with $a = E_{ee}(2)/4$ for $N = 2$, it is clear that

$$\alpha = \frac{E_{ee}(2) + \beta 2^{4/3}}{4} > a$$

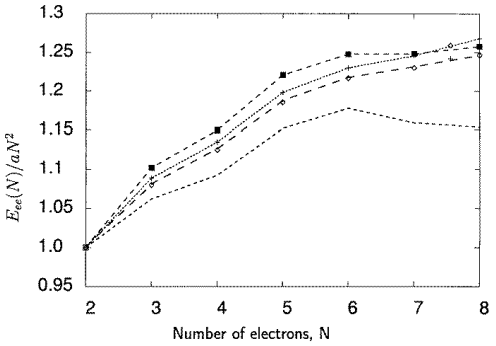

Figure 10. $E_{ee}(N)/aN^2$ for various values of N for a SAD. A value of 1 corresponds to the approximation of $E_{ee}(N)$ by a quadratic form.

thus the quadratic approximation results in an underestimation of $E_{ee}(N)$ and the orbital-pair interactions. This point is illustrated in figure 10 which shows the variation of $E_{ee}(N)/aN^2$ with $a = E_{ee}(2)/4$ as calculated with the LSDA for several kinds of dot. In this figure a quadratic dependence of the electron–electron interaction should be represented by a horizontal line at $E_{ee}(N)/aN^2 = 1$. It is clear that $E_{ee}(N)$ shows significant departure from the quadratic dependence on N for SADs, which can be accounted for by the exchange and correlation effect. The second feature is due to the fact that the quadratic Coulomb approximation to $E_{ee}(N)$ does not address explicitly spin-dependent quantum effects such as exchange and correlation. While these exchange and correlation effects make a greater contribution to the interaction energies at low N, they constitute a gradually decreasing percentage of the total electron energy with increasing N, thereby bringing our estimated values closer to the computed ones. Considering that our scheme is meant to give only an estimate of the interactions, the resulting 35% discrepancy should be acceptable in most cases.

Table 2. Values of the characteristic length scale r_0 and E_{ss} for various SADs.

Dot	Base, b (Å)	Height, h (Å)	α/r_0 (Å^{-1})	E_{ss} (meV)
A	150	75	8.4×10^{-3}	24
B	200	100	6.3×10^{-3}	19.1
C	400	200	2.7×10^{-3}	10.4
Pyramid	200	100	6.3×10^{-3}	22.3

Figure 9 also shows that the s–s interaction gets weaker as the dots get bigger. In fact, we observe that the s–s interaction scales approximately as $1/r_0$, where r_0 is a characteristic length for the dot given by $r_0 = \lambda\sqrt[3]{hb^2}$, where λ is a numerical constant, h the height, and b the radius of the dot. In table 2 we show values of $1/r_0$ for the different dots considered in this work. The $1/r_0$ behaviour of the s–s interaction can be seen by comparing the ratios of E_{ss} and $1/r_0$ for the three dots: E_{ss}^{A}:E_{ss}^{B}:E_{ss}^{C} = 1.3:1.0:0.5 whereas $1/r_0^{A}$:$1/r_0^{B}$:$1/r_0^{C}$ = 1.3:1.0:0.4. While the present data point to a $1/r_0$ scaling of E_{ss} in lens-shaped SADs only, there is no reason to believe that it should not hold for pyramidal dots as well.

Acknowledgments

This work was supported by CRI from the University of Illinois and NSF Grants No ECS 95-09751 and No DMR 94-22496. One of us (SN) would like to acknowledge support from the Beckman Institute Research Assistantship programme.

References

[1] Fafard S, Hinzer K, Raymond S, Dion M, McCaffrey J, Feng Y and Charbonneau S 1996 *Science* **274** 1350
[2] Jimenez J L, Fonseca L R C, Brady D J, Leburton J P, Wohlert D E and Cheng K Y 1997 *Appl. Phys. Lett.* **71** 3558
[3] Imamura K, Sugiyama Y, Nakata Y, Muto S and Yokoyama N 1995 *Japan. J. Appl. Phys.* **34** 1445
[4] Tarucha S, Austing D G, Honda T, van der Hage R J and Kouwenhoven L P 1996 *Phys. Rev. Lett.* **77** 3613
[5] Fricke M, Lorke A, Kotthaus J P, Medeiros-Ribeiro G and Petroff P M 1996 *Europhys. Lett.* **36** 197
[6] Solomon G S, Trezza J A, Marshall A F and Harris J S Jr 1996 *Phys. Rev. Lett.* **76** 952
[7] Miller M S, Malm J O, Pistol M E, Jeppesen S, Kowalski B, Georgsson K and Samuelson L 1996 *J. Appl. Phys.* **80** 3360
[8] Ledentsov N N *et al* 1996 *Phys. Rev.* B **54** 8743
[9] Fonseca L R C, Jimenez J L, Leburton J P and Martin R M 1998 *Phys. Rev.* B **57** 4017
[10] Ledentsov N N *et al* 1996 *Solid-State Electron.* **40** 785
[11] Perdew J P and Wang Y 1992 *Phys. Rev.* B **45** 13 244
[12] Fonseca L R C, Jimenez J L and Leburton J P 1998 *Phys. Rev.* B **58** 9955
[13] Chuang S L 1995 *Physics of Optoelectronic Devices* (New York: Wiley)
[14] Singh J 1993 *Physics of Semiconductors and their Heterostructures* (New York: McGraw-Hill)
[15] Cusack M A, Briddon P R and Jaros M 1996 *Phys. Rev.* B **54** R2300
[16] Grundmann M, Stier O and Bimberg D 1995 *Phys. Rev.* B **52** 969
[17] Jimenez J L, Fonseca L R C, Leburton J P and Brady D J 1999 unpublished
[18] Jovanovic D and Leburton J P 1994 *Phys. Rev.* B **49** 10 841
[19] Inui T, Tanabe Y and Onodera Y 1996 *Group Theory and its Applications in Physics* (New York: Springer)
[20] Garcia J M, Medeiros-Ribeiro G, Schmidt K, Ngo T, Feng J L, Lorke A, Kotthaus J and Petroff P M 1994 *Appl. Phys. Lett.* **71** 2252
[21] We use Mulliken notation for the irreducible representations, as described in
Inui T, Tanabe Y and Onodera Y 1996 *Group Theory and its Applications in Physics* (Berlin: Springer)
[22] Shumway J, Fonseca L R C, Leburton J P and Ceperley D 1999 unpublished
[23] Parr R G and Yang W 1989 *Density-Functional Theory of Atoms and Molecules* (New York: Oxford University Press)
[24] Janak J F 1978 *Phys. Rev.* B **18** 7165
Janak J F 1978 *Phys. Rev.* B **18** 10 841
[25] Slater J C 1972 *Adv. Quantum Chem.* **6** 1
[26] Nagaraja S, Matagne P, Thean V-Y, Leburton J P, Kim Y-H and Martin R M 1997 *Phys. Rev.* B **56** 15 752
[27] Lee I-H, Rao V, Martin R M and Leburton J P 1998 *Phys. Rev.* B **56** 9035
[28] Macucci M, Hess K and Iafrate G J 1997 *Phys. Rev.* B **55** R4879
[29] Koskinen M, Manninen M and Reimann S M 1997 *Phys. Rev. Lett.* **79** 1389
[30] Jones R O 1987 *Ab Initio Methods in Quantum Chemistry* ed K P Lawley (New York: Wiley) p 413
[31] Schmidt K E and Ceperley D M 1992 *Monte Carlo Methods in Condensed Matter Physics (Springer Topics in Applied Physics 71)* ed K Binder (Heidelberg: Springer) p 205
[32] Mahan G D 1990 *Many-Particle Physics* (New York: Plenum)
[33] Nagaraja S, Matagne P, Leburton J P, Kim Y H and Martin R M 1997 *Phys. Rev.* B **56** 15 572
[34] Shumway J, L. Fonseca R C, Leburton J P and Ceperley D 1999 unpublished