Interactions of metallic quantum dots on a semiconductor substrate

Wei Lu* and David Salac
Department of Mechanical Engineering, University of Michigan, Ann Arbor, Michigan 48109, USA
(Received 20 June 2006; published 11 August 2006)

Experiments have shown that uniform metallic quantum dots may self-assemble on a semiconductor substrate. The observation calls for a repulsive force when the dots are close. In a traditional quantum dot system, such as Ge dots on a Si substrate, such an action is achieved by elastic interaction. This paper proposes a mechanism for metallic dots without coherent lattice or lattice mismatch, so that elastic effect may not account for the phenomena. We show that electric double layers due to contact potential can lead to size-dependent repulsion, which counterbalances van der Waal attraction and determines feature sizes.

DOI: 10.1103/PhysRevB.74.073304 PACS number(s): 68.43.Hn, 68.35.Md

Nanometer scale metallic dots or clusters grown on a semiconductor substrate have wide applications in optical, electronic, and magnetic devices. Production of these structures over a large area using techniques such as lithography and etching can be expensive and difficult. Recent experiments showed that uniform metallic dots may form spontaneously. Examples include Cu on TiO_2,[1] Au on Si(111),[2] Fe on NaCl (100),[3] and $CoSi_2$ on Si(111).[4] Despite their large collective surface area, the densely packed dots did not coalesce, but maintained small distances from one another. The observations call for a repulsive force when they are close. In a traditional quantum dot system, such as Ge dots on a Si substrate, the repulsion is achieved by elastic interaction. Both Ge and Si have the same cubic lattice structure, but their lattice constants differ by about 4%. The Ge dots deform laterally to match the Si lattice. Each Ge dot, due to its larger lattice constant, induces below it a tensile stress region in the substrate. The repulsion between these tensile stress regions keeps the dots separated.

This paper proposes a repulsion mechanism for metallic dots. These systems may not involve coherent lattice or lattice mismatch, so that the elastic effect cannot explain the phenomena. A qualitative understanding is provided in the following. The metallic dots and the substrate have different Fermi levels. When they are brought in contact, charge transport occurs. Take *n*-doped semiconductor as an example. As a result of the charge transport, negative charges accumulate at the metal interface while a cloud of positive charges form in the substrate. This configuration is known as electric double layer. Thus, underneath a metallic dot, there is a charge cloud in the semiconductor. The exact shape and density of this cloud depends on factors such as the strength of contact potential and contact geometry. When the dot moves, the charge cloud moves with it. When two dots approach each other, the accompanying charge clouds overlap, leading to a repulsive force that prevents them to coalesce. This possibility is exciting since the repulsive force may help to assemble a lattice of dots such as colloidal crystals.

Electric double layers form the basis of *p*-*n* junctions and metal-semiconductor contacts.[5] However, the attention of existing work focuses on electronic properties, such as the recent study of nanoscale contacts.[6,7] Little work has been done to investigate the self-assembly phenomena. This paper aims to investigate the role of electric double layers in controlling morphology. While experiments have also shown that electrostatic interaction may lead to nicely ordered particle lattices in dusty plasma[8] and colloidal suspensions,[9] several differences distinguish the metallic dot system from them. First, in a semiconductor, say *n* doped, electrons are mobile, but positive ions do not. By contrast, in a liquid electrolyte, ions of both signs are mobile. Second, charge distribution leads to different degrees of screening in the two systems. Imagine two colloidal particles carrying negative charges. Each of these particles will be surrounded by a layer of positive charges in the solution, which significantly reduces the repulsion between them. The screening effect in metallic dots is much weaker since they are only partially surrounded by opposite charges in the substrate. Third and most importantly, colloidal particles and particles in dusty plasmas usually carry fixed charges while metallic dots keep constant contact potential, which leads to a different energy landscape.

In a metal-semiconductor junction, the charge density decays with the distance from the interface. The traditional one-dimensional full-depletion approximation for metal-semiconductor interfaces can only be applied to describe the charge cloud in a large diode.[5] As the contact size scales down, truly three-dimensional models must be adopted. We find that when the contact is smaller than a characteristic length, the size of the charge cloud is no longer determined by the doping level or the free carrier concentration, but instead by the size and shape of the interface, which strongly affects the interaction among dots. The electric double layers form an energy barrier that keeps two dots separated. If the barrier is passed, van der Waal attraction will bring the dots all the way together. Surface energy will then drive them to combine into a large dot.

In the following we consider metallic dots on a semi-infinite substrate. A coordinate is attached so that the substrate occupies the half space below the x_1-x_2 plane ($x_3<0$). Inside the semiconductor, the electric potential ϕ and the charge density ρ vary with the position. The electrostatic field obeys the Poisson equation, i.e.,

$$-\varepsilon_s \nabla^2 \phi = \rho, \tag{1}$$

where ε_s is the permittivity of the substrate. The charge density ρ is a function of the electron density, hole density and donor and acceptor densities. The electrons obey the Fermi-Dirac statistics. The general form of the function, ρ, can be

rather cumbersome. For a n-doped nondegenerate semiconductor with fully ionized donors, which is typical in many situations, the charge density is given by[5]

$$\rho = qN_d[1 - \exp(q\phi/k_BT)], \tag{2}$$

where N_d is the donor density in the substrate, q the charge of one electron, k_B Boltzmann's constant, and T the temperature. The right hand side of Eq. (2) implies that the charge density at a spatial point depends on the potential at that point.

Equations (1) and (2) give a characteristic length $l_D = \sqrt{\varepsilon_s k_B T/(q^2 N_d)}$, which is known as the Debye length. A silicon substrate has $\varepsilon_s = 11.7\varepsilon_0$, where ε_0 is the permittivity of vacuum. Using $\varepsilon_0 = 8.85 \times 10^{-12}$ F/m, $q = 1.60 \times 10^{-19}$ C, $k_B = 1.38 \times 10^{-23}$ J/K, $T = 300$ K, and a typical donor density of $N_d = 10^{15}$ cm^{-3}, we have $l_D = 129.4$ nm. A characteristic voltage is defined by $\phi_0 = k_BT/q$, which has a value of $\phi_0 = 2.59 \times 10^{-2}$ V with the above parameters. Normalize with the length l_D and denote normalized potential by $\psi = \phi/\phi_0$. Equations (1) and (2) reduce to

$$-\nabla^2\psi = 1 - e^{\psi}. \tag{3}$$

To calculate the electrostatic energy in the system, we propose an approach similar to treating the double layers of colloidal particles.[10,11] Consider an incremental charge build-up process on the metal interface. Denote the total charge accumulated on the interface by Q_s. The energy stored in the system, U, is the work to move charges from the substrate to the metal interface. One may imagine that the system is loaded by gradually increasing Q_s from zero to its current value. The energy is thus given by $U = \int_0^{Q_s} \phi_s(Q_s')dQ_s'$. The electric potential of the dot, $\phi_s(Q_s')$, depends on charge accumulation. Unlike colloidal particles holding constant charges, the metallic dots hold constant contact potential ϕ_s when they move on the substrate. The Gibbs free energy of the system, G_e, can be obtained by the Legendre differential transformation, giving $G_e = -\int_0^{\phi_s} Q_s(\phi_s')d\phi_s'$. Note that the function $Q_s(\phi_s')$ is nonlinear since the charges in the substrate redistribute, which parallels charging a capacitor with a voltage-dependent capacitance. Direct application of this energy expression requires solving Eq. (3) for a series of surface potentials so that the $Q_s(\phi_s')$ curve can be constructed. This incremental computation has to be repeated for any new configuration since $Q_s(\phi_s')$ depends on the spatial distribution of the dots. Solving the nonlinear equation (3) repeatedly in three dimensional is computational intensive. Our analysis showed that the free energy can be expressed by a volume integration in the current electrostatic state without the need to calculate the history of $Q_s(\phi_s')$. This approach significantly reduces the computational effort since Eq. (3) only needs to be solved once for each configuration. The expression was obtained by energy variation, which gives

$$G_e = \int_V \left[\int_0^{\phi} \rho(\phi')d\phi' - \frac{1}{2}\varepsilon_s|\nabla\phi|^2 \right] dV. \tag{4}$$

The integration extends over the volume of the substrate plus the air region above the substrate surface. Note that the air region has $\rho = 0$, where the free energy reduces to the ordinary form of $-\int_V \varepsilon_a|\nabla\phi|^2 dV/2$ with ε_a being the permittivity of air. Calculation of the potential gradient in the volume can be avoided by applying the divergence theorem. After normalization by the length l_D and voltage ϕ_0, we have

$$\frac{G_e}{N_d k_B T l_D^3} = \int_{V_{Sub}} \left[1 + \frac{\psi}{2} + \left(\frac{\psi}{2} - 1 \right) \exp(\psi) \right] dV + \frac{1}{2} \int_A \psi \left[\frac{\varepsilon_a}{\varepsilon_s} \left(\frac{\partial\psi}{\partial x_3} \right)_{x_3=0^+} - \left(\frac{\partial\psi}{\partial x_3} \right)_{x_3=0^-} \right] dA. \tag{5}$$

Note that here the volume integration extends only in the substrate V_{Sub} and the area integration extends on the contact interface. The terms $(\partial\psi/\partial x_3)_{x_3=0^+}$ and $(\partial\psi/\partial x_3)_{x_3=0^-}$ are the potential gradients at the contact interface calculated from the air and the substrate region, respectively. In fact, the integrand of the area integration in Eq. (5) vanishes in the noncontact substrate surface due to the continuity of the electric displacement. The area integral in Eq. (5), an adapted form for thin metal disks, comes from the surface integration around the metals. This term in its dimensional form is essentially $-Q_s\phi_s/2$, the free energy of ordinary linear dielectrics with fixed boundary potentials. Thus the volume integral in Eq. (5) can be viewed as a nonlinear correction due to the charge cloud in the substrate.

An analytical solution to Eq. (3) cannot be obtained due to the nonlinear charge density. While finite element and finite difference methods have been used, accurate numerical approach for large and inhomogeneous systems remains to be a challenge. We propose a spectral and interface relaxation method for both high accuracy and efficiency. The electric fields satisfy the Laplace equation in the air region and Eq. (3) in the substrate. The two fields have prescribed contact potential at the contact interface, and continuous normal electric displacement across the noncontact area. We temporally relax the second requirement by assuming a potential distribution $\psi_b(x_1, x_2)$ on the entire substrate surface. This function satisfies the contact potential at the contact interface and has arbitrary values in the noncontact area. Then $\psi_b(x_1, x_2)$ serves as a given potential boundary condition so that the electric fields in the air region and the substrate can be solved independently. In the substrate ($x_3 < 0$) we apply two-dimensional Fourier transform within the x_1-x_2 plane to convert the three dimensional partial differential equation into a set of ordinary differential equations in the x_3 direction. These tridiagonal systems of equations can be solved efficiently. The field in the air ($x_3 > 0$) can be solved analytically in Fourier space.

Here we focus on a class of low-profile metallic dots and treat them as zero-thickness objects. Thus the system can be viewed as two half-spaces. This configuration is an approximation for other three-dimensional shapes of quantum dots, as long as the dot-dot distance is relative large comparing to the height. Rigorous calculation of high-profile dots requires consideration of their three-dimensional geometries above the substrate. Generally speaking, the two half-spaces solved separately will not satisfy the continuity of the electric dis-

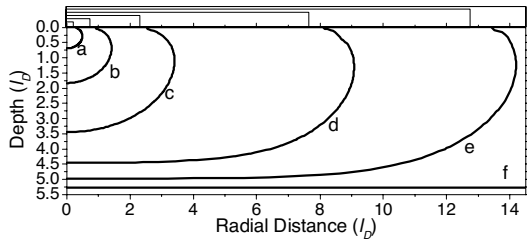

FIG. 1. Contour of 4.5% ψ_s for disks with radius (a) $0.232l_D$ (30 nm), (b) $0.773l_D$ (100 nm), (c) $2.32l_D$ (300 nm), (d) $7.73l_D$ (1 μm), (e) $12.36l_D$ (1.6 μm), (f) infinite contact.

placement. The discontinuity corresponds to a surface charge density of $\sigma=(\partial\psi/\partial x_3)_{x_3=0^-}-(\varepsilon_a/\varepsilon_s)(\partial\psi/\partial x_3)_{x_3=0^+}$, which increases the energy by $G_b=\int_{A_{Non\text{-}contact}}\sigma\psi_b dA/2$. Minimizing the functional with the Langevin steepest decent approach gives $\beta\partial\psi_b/\partial t=-\delta G_b/\delta\psi_b$, namely

$$\beta\frac{\partial\psi_b}{\partial t}=\frac{\varepsilon_a}{\varepsilon_s}\left(\frac{\partial\psi}{\partial x_3}\right)_{x_3=0^+}-\left(\frac{\partial\psi}{\partial x_3}\right)_{x_3=0^-}, \tag{6}$$

where β is an adjustable parameter to control the rate of convergence. The evolution of $\psi_b(x_1,x_2)$ reduces and eventually eliminates the discontinuity when it reaches the actual potential distribution.

In the simulation we prescribe ϕ_s at the contact interface and zero potential deep inside the substrate. Contact potential ϕ_s can be measured experimentally or be related to other physical quantities by $\phi_s=\phi_B-V_n$, where ϕ_B is the barrier height between the metal and semiconductor and $qV_n=E_C-E_F$ is the energy difference between the conduction band E_c and the Fermi level E_F in the semiconductor. For $CoSi_2$ on Si(111) the barrier height is $\phi_B=0.67$ eV.[12] The silicon substrate has $E_C-E_F=1.1$ eV. Then the contact potential is $\phi_s=-0.4$ V or $\psi_s=-15.4$.

The $CoSi_2$ metallic dots on n-doped silicon substrate were taken as an example system in our study. To better reflect the profile we refer them as disks. The donor density was taken to be $N_d=10^{15}$ cm^{-3}, which gave a Debye length of $l_D=129.4$ nm. The normalized contact potential was -15.4. Computations suggest that the size of the charge cloud in the semiconductor becomes highly dependent on the dimension of the contact interface when the contact falls into nanoscale. This trend is shown in Fig. 1, which gives the potential contour of 4.5% ψ_s for various disk radii. According to Eq. (2), this is also the charge density contour of half of the maximum charge density at the contact interface. The enclosure of the contour can be viewed as the size of the charge cloud since the charge density decays sharply outside. The curve for infinite contact was obtained from a one-dimensional model. The depth of the charge cloud increases with the contact size, and quickly approaches that of an infinite contact, i.e., curve f. Thus a large contact causes the charge cloud to expand most laterally, with little increase in the depth. This behavior has a bearing on the interaction of disks and suggests that the traditional one-dimensional analysis cannot be applied when the contact size falls into nanoscale.

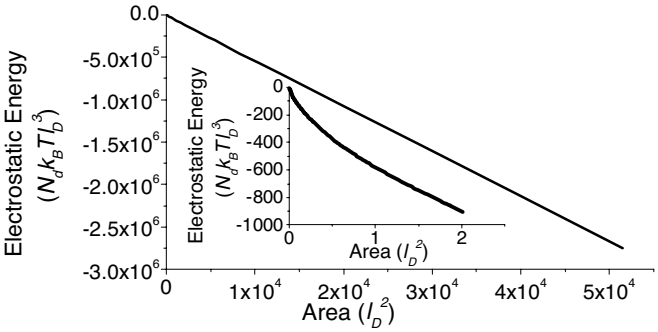

FIG. 2. Large disks show a linear relationship between electrostatic free energy and disk area. The relation is nonlinear for small disks.

Figure 2 shows the relation between the electrostatic free energy and disk area. A relative large disk demonstrates a linear relationship. Noticeable deviation occurs when the disk size falls below l_D, as shown in the inset. This size effect is due to the depth of the charge cloud not being fully developed. For a small disk, an increase in the radius results in an expansion of both depth and width of the charge cloud. In contrast, the depth of the charge cloud is fully developed for a relative large disk. Any increase in the disk area essentially pushes the charge density profile to expand laterally. In this situation the charge cloud can be visualized as a core and shell configuration. The core is a cylinder with fixed height. The shell is a transition region where the cloud depth tapers to zero. When a disk is large, the cylinder dominates the free energy, whose volume scales with the disk area. This leads to the linear relation in Fig. 2. Calculating the slope of the straight line shows that the energy density per normalized disk area is $-52\ N_d k_B T l_D^3$. Our one dimensional analysis of an infinite large disk gives the same value.

Figure 3 shows the interactive energy of two disks with various radii. The zero energy state is defined when two disks are infinitely separated. Significant repulsion appears when the separation reduces to a scale comparable to l_D. The curves quickly become flat when the separation is larger than several l_D, suggesting an effective cut-off distance for the disks to sense each other. At a given separation, larger disks demonstrate stronger repulsion. The increasing repulsion with disk growth may help to prevent their coalescence. The inset shows the distribution of charge cloud in the substrate.

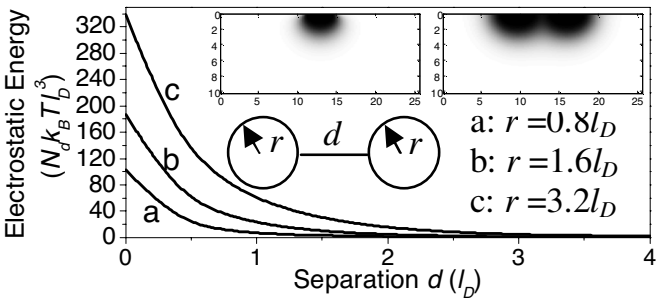

FIG. 3. Electrostatic interaction as a function of disk separation.

Darker regions correspond to higher charge density, which reaches the maximum of approximately qN_d at the contact interface and decays quickly in the substrate. When two disks are close, their charge clouds overlap and deform.

Is the repulsion strong enough to counterbalance van der Waals attraction? We calculated the latter by pairwise summation of intermolecular forces, namely $G_{vdW} = \Sigma_i \Sigma_{i>j}(-B/r_{ij}^6)$.[13] Here B is the London constant and i and j count over all atoms or molecules. The term r_{ij} is the distance between two atoms. The London constant relates to the Hamaker constant A by $A = \pi^2 \rho_1 \rho_2 B$, where ρ_1 and ρ_2 are the atomic density of two interacting bodies. The Hamaker constant can be determined from the index of refraction. For $CoSi_2$, the index value of 1.3 gives a Hamaker constant of 1.8×10^{-19} J or a London constant of 2.0×10^{-77} J m^6.[14] Figure 4(a) shows the total electrostatic and van der Waals energies for various radii. An energy barrier appears at a separation of approximately $0.02 l_D$, below which van der Waals interaction dominates and drives two disks to contact. The barrier depends on the disk size and the contact potential. Larger disks have higher barrier and are less likely to combine. The percentage of the substrate surface covered by metal is a linear function of r/d. For instance, a triangular lattice of disks with $r/d=5.3$ correspond to a coverage of 0.76. For a given coverage, the energy barrier determines the disk size. The dashed curve in Fig. 4(a), the $r/d=5.3$ contour, illustrates the idea. This curve passes point 3, the maximum of the $r=0.12 l_D$ curve. Smaller disks, such as point 4, tend to grow since there is no barrier for them to coalesce. On the other hand, larger disks such as point 1 or 2 are energetically unfavorable. These effects lead to disk radius and separation defined by point 3. Higher temperature leads to larger Debye length and thus larger physical disk size and separation. This prediction is consistent with experiments, which showed that increasing the annealing temperature in the $TiSi_2/Si(001)$ system produced disks with larger size and spacing.[15] Figure 4(b) shows the barrier for two small disks of radius $0.08 l_D$ with normalized contact potential varying from -1 to -15.4. When the magnitude of the contact potential is small, the repulsion cannot overcome van der Waals attraction and there is no barrier for coalescence.

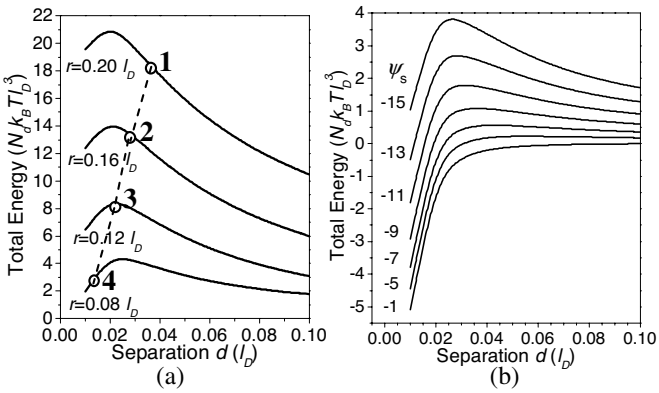

FIG. 4. The electrostatic and van der Waals energies for disks with (a) various radii (b) radius of $0.08 l_D$ for various contact potentials.

In summary, the electrostatic and van der Waals energies lead to two regimes separated by an energy barrier. These two energies determine the size and spatial ordering. The energy barrier depends on contact potential, which suggests a possibility of materials selection or application of a bias voltage to the substrate to change the contact potential and thus engineer feature sizes.

The authors acknowledge financial support from the National Science Foundation Career Award No. DMI-0348375.

*Electronic address: weilu@umich.edu

[1] D. L. Carroll, M. Wagner, M. Ruhle, and D. A. Bonnell, Phys. Rev. B **55**, 9792 (1997).
[2] M. Hugelmann and W. Schindler, Appl. Phys. Lett. **85**, 3608 (2004).
[3] Z. Gai, B. Wu, J. P. Pierce, G. A. Farnan, D. J. Shu, M. Wang, Z. Y. Zhang, and J. Shen, Phys. Rev. Lett. **89**, 235502 (2002).
[4] F. M. Ross, P. A. Bennett, R. M. Tromp, J. Tersoff, and M. Reuter, Micron **30**, 21 (1999).
[5] S. M. Sze, *Physics of Semiconductor Devices* (John Wiley & Sons, New York, 1981).
[6] T. Sato, S. Kasai, and H. Hasegawa, Appl. Surf. Sci. **175**, 181 (2001).
[7] G. D. J. Smit, S. Rogge, and T. M. Klapwijk, Appl. Phys. Lett. **80**, 2568 (2002).
[8] H. Thomas, G. E. Morfill, V. Demmel, J. Goree, B. Feuerbacher, and D. Möhlmann, Phys. Rev. Lett. **73**, 652 (1994).
[9] C. A. Murray and D. H. VanWinkle, Phys. Rev. Lett. **58**, 1200 (1987).
[10] D. Y. C. Chan and D. J. Mitchell, J. Colloid Interface Sci. **95**, 193 (1983).
[11] K. A. Sharp and B. Honig, J. Phys. Chem. **94**, 7684 (1990).
[12] R. T. Tung, Mater. Chem. Phys. **32**, 107 (1992).
[13] W. R. Bowen and F. Jenner, Adv. Colloid Interface Sci. **56**, 201 (1995).
[14] R. H. French, J. Am. Ceram. Soc. **83**, 2117 (2000).
[15] W. Yang, F. J. Jedema, H. Ade, and R. J. Nemanich, Thin Solid Films **308**, 627 (1997).