

# The single electron transistor and artificial atoms

M. A. Kastner

Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139  
USA  
mkastner@mit.edu

Received 8 Aug. 2000, accepted 29 Sep. 2000 by C. Thomsen

**Abstract.** Modern techniques of lithography make it possible to confine electrons to sufficiently small dimensions that the quantization of both their charge and their energy are easily observable. When such confined electrons are allowed to tunnel to metallic leads a single electron transistor (SET) is created. This transistor turns on and off again every time one electron is added to the isolated region. Whereas we can understand conventional transistors using classical concepts, the SET is quantum mechanical in an essential way. In fact, there is a close analogy between the confined electrons inside an SET and an atom. In this review, the physics underlying the operation of SETs is explained, a brief history of its invention is presented, and issues of current interest are discussed.

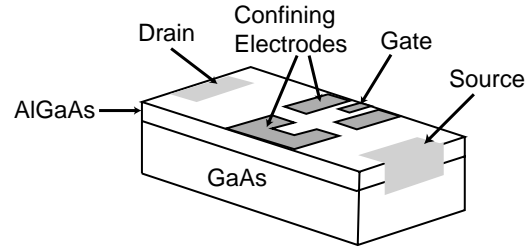
**Keywords:** single electron transistor, mesoscopic physics, nanostructure

**PACS:** 73.23.Hk, 72.15.Qm, 73.23.-b

## 1 The physics of single electron transistors

A conventional field-effect transistor, the kind that makes all modern electronics work, is a switch that turns on when electrons are added to a semiconductor and turns off when they are removed. These on and off states give the ones and zeros that digital computers need for calculation. Interestingly, these transistors are almost completely classical in their physics. Only a few numbers that characterize their behavior are affected by quantum mechanics. However, if one makes a new kind of transistor, in which the electrons are confined within a small volume and communicate with the electrical leads by tunneling, all this changes. One then has a transistor that turns on and off again every time one electron is added to it; we call it a single electron transistor (SET). Furthermore, the behavior of the device is entirely quantum mechanical. It is very appropriate, therefore, to discuss this device on the 100th anniversary of Planck's constant.

Various structures have been made in the past two decades, in which electrons are confined to small volumes in metals or semiconductors. Perhaps not surprisingly, there is a deep analogy between such confined electrons and atoms. Whereas natural atoms are studied by adding, removing or exciting electrons with light, these artificial atoms typically have such small energy scales that they are best studied by measuring the voltage and current resulting from tunneling between the artificial atom and nearby



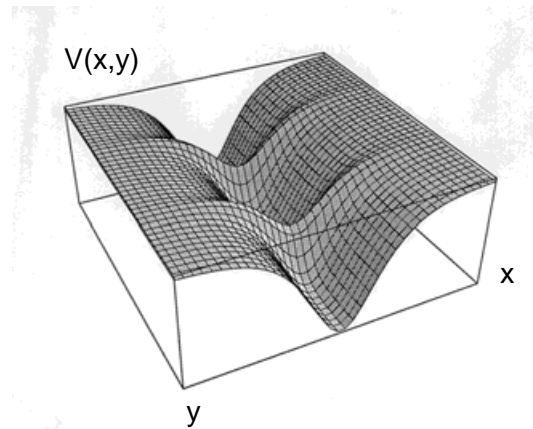
**Fig. 1** Schematic drawing of a SET. Wires are connected to source and drain contacts to pass current through the 2DEG at the GaAs/AlGaAs interface. Wires are also connected to the confining electrodes to bias them negatively and to the gate electrode that controls the electrostatic energy of the confined electrons.

electrodes. The reader who is interested in a more thorough discussion of these and related devices should read one of several review articles. [1–5]

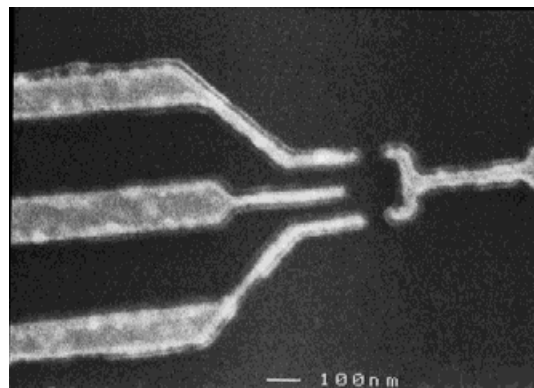
A schematic of one kind of SET is shown in Fig. 1. It consists of a semiconductor, in this case GaAs, separated from metal electrodes by an insulator, in this case AlGaAs. The AlGaAs is doped with Si, which donates electrons. These fall into the GaAs, because their energy is lower in the latter material. The resulting positive charge on the Si atoms creates a potential that holds the electrons at the GaAs/AlGaAs interface, creating a two dimensional electron gas (2DEG). The source and drain contacts allow one to drive electrons from an external circuit through the 2DEG. The 2DEG is confined perpendicular to the GaAs/AlGaAs interface, and the confinement in the other two directions is accomplished with electric fields imposed by very small confinement electrodes. A negative voltage on these electrodes creates a potential similar to the one sketched in Fig. 2; the negative voltage repels electrons from underneath the confinement electrodes and creates saddle point potential barriers under the constrictions. For the remainder of our discussion, we assume that the voltage on these constriction electrodes is fixed, resulting in a fixed confinement potential. However, the voltage on an additional electrode, the gate, is varied to adjust the potential of the electrons confined in the potential well.

Figure 3 shows an electron micrograph of such constriction and gate electrodes for one of the smallest SETs made in this way so far. [6, 7] The region surrounded by electrodes appears to be a few hundred nanometers in diameter. However, the droplet of electrons confined in it is considerably smaller. We estimate that these SETs have about 50 electrons confined to a droplet about 100 nm in diameter. The GaAs/AlGaAs structure is grown with molecular beam epitaxy. The electrodes are fabricated using electron beam lithography.

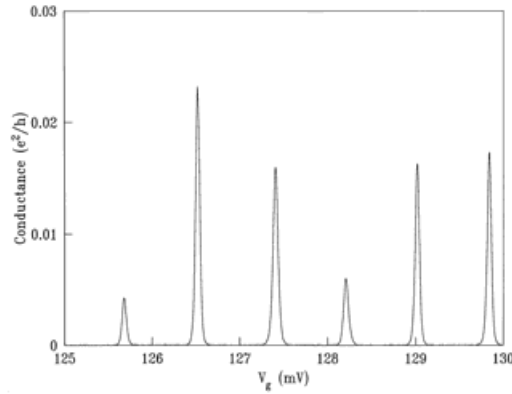
When the voltage on the gate electrode is increased, the potential minimum, in which the electrons are trapped, becomes deeper. This causes the number of trapped electrons to increase. However, unlike a conventional transistor, in which the charge increases continuously, the charge in the trap increases in discrete steps, and this is reflected in the conductance between source and drain.



**Fig. 2** Sketch of the electrostatic potential energy experienced by an electron moving at the interface between GaAs and AlGaAs in Fig. 1.



**Fig. 3** Electronmicrograph of the top surface of the SET used in the experiments of Goldhaber-Gordon *et al.* [6,7]. See also Fig. 1.



**Fig. 4** Conductance of a SET as a function of the gate voltage. The spacing between the peaks is the voltage necessary to add one electron to the artificial atom. These results are for a device of the kind studied by Meirav *et al.* [8]

Figure 4 shows the conductance as a function of gate voltage  $V_g$  for a SET made by U. Meirav *et al.* [8], larger than the one in Fig. 3. The conductance is measured by applying a very small voltage  $V_{ds}$  between drain and source, small enough that the current is proportional to  $V_{ds}$ . As seen in the figure, the conductance increases and decreases by several orders of magnitude almost periodically in  $V_g$ . A calculation of the capacitance between the gate electrode and the droplet of confined electrons shows that the voltage between two peaks or two valleys is just that necessary to add one electron to the droplet. The name “single electron transistor” comes from the observation that the transistor turns on and off again every time a single electron is added to it.

For current to flow the number of electrons on the droplet must fluctuate, say between  $N$  and  $N + 1$ . Thus the  $N$ th peak in the conductance occurs when the state of the droplet containing  $N$  electrons is in equilibrium with the state containing  $N + 1$  electrons. Were the gate the only electrode contributing to the electrostatic energy of the droplet, the gate voltage at which the  $N$ th peak occurred multiplied by the charge of the electron  $e$  would be the energy difference between the two states. Since there are several electrodes near the droplet, the energy change caused by  $V_g$  is  $\alpha e V_g$  where  $\alpha = C_g/C$  is the ratio of the gate capacitance to the total capacitance. Therefore, a conductance peak occurs when

$$\alpha e V_g(N) = E(N + 1) - E(N), \quad (1)$$

apart from a constant, where  $E(N)$  is the total energy of the droplet with  $N$  electrons. Thus, predicting the position of the  $N$ th peak in Fig. 4 requires a model for the total energy of an artificial atom with  $N$  electrons.

There are several levels of sophistication with which such models have been developed. The simplest is called the Coulomb blockade model, and treats the droplet of confined electrons as a metal particle. Think about how an electron tunnels from

one lead onto the metal particle and then onto the other lead. Suppose the particle is neutral to begin with. To add a charge  $Q$  to the particle requires energy  $Q^2/2C$ , where  $C$  is the total capacitance between the particle and the rest of the system. Setting  $Q = -Ne$  in Eq. 1 gives values of the peak positions that are equally spaced in gate voltage with separation  $\Delta V_g = e/C_g$ .

This result is a direct consequence of the charge quantization. Taking  $E(Q) = Q^2/2C$  means that the energy as a function of  $Q$  is a parabola with minimum at  $Q_0$ . By varying  $V_g$  we can choose any value of  $Q_0$ , the charge that would minimize the energy, were not the charge quantized. However, because the real charge *is* quantized, only discrete values of the energy  $E$  are possible. When  $Q_0 = -Ne$ , for which an integer number  $N$  of electrons minimizes  $E$ , the Coulomb interaction results in an energy difference  $U = e^2/2C$  for increasing or decreasing  $N$  by one. There is thus an energy gap that suppresses charge fluctuations. For all values of  $Q_0$  except  $Q_0 = -(N + 1/2)e$  there is a smaller, but non-zero, energy gap for adding or subtracting an electron. Under these circumstances no current can flow at temperature  $T = 0$ . However, when  $Q_0 = -(N + 1/2)e$  the state with  $Q = -Ne$  and that with  $Q = -(N + 1)e$  are degenerate; the charge fluctuates between the two values even at zero temperature. Consequently, the energy gap disappears and current can flow. The peaks in conductance are, therefore, periodic, occurring whenever the average charge on the artificial atom is  $Q_0 = -(N + 1/2)e$ .

Thus the approximately equally spaced peaks in conductance of the SET results from quantization of charge. But under what conditions is charge quantized? This may seem like a silly question. Since the discovery of the electron we have known that its charge is quantized. However, because the wavefunctions of electrons in conductors are extended over macroscopic distances, the charge in any small volume is not quantized. It is the localization of electrons to a small region of space that quantizes their charge.

However, the degree of localization depends on the transmission of the tunnel barriers. A very elegant argument tells us how resistive the tunnel barriers must be for charge quantization. One simply demands that the  $RC$  time constant for an electron to tunnel off the droplet into the leads be great enough that the energy uncertainty is less than the charging energy. If the tunneling resistance is  $R$ , this condition is  $RC > h/U$  or approximately  $R > h/e^2$ , the fundamental unit of resistance that enters, for example, in the quantum Hall effect. Thus, while the calculation of the charging energy is entirely classical, Planck's constant determines whether the charging energy is present or not. This condition is valid at  $T = 0$ , independent of  $C$  and therefore of the size of the artificial atom. Of course, thermal charge fluctuations can overcome this localization, so charge quantization is observable only at temperatures  $kT < U$ , which means that it is easier to see the effects in smaller artificial atoms, which have larger  $U$ .

In addition to charge quantization, energy quantization is important when electrons are confined to small volumes. Interestingly, the criterion for charge and energy quantization at  $T = 0$  are exactly the same. Whereas  $U$  is the energy to add an extra electron to the artificial atom, there is a typical level spacing  $\Delta\epsilon$  necessary to excite the artificial atom with fixed number of electrons. Furthermore, the levels of the artificial atom are not perfectly sharp, but rather have typical width  $\Gamma$ . The level width is caused by lifetime broadening, because an electron in a level on the artificial

atom can tunnel into the leads. Alternatively, one can say that the eigenstates of the system are mixtures of localized states on the artificial atom and extended states in the leads.

Clearly, energy quantization means that  $\Delta\epsilon > \Gamma$ . Following Thouless [9] the current through the SET for a single quantum level is the charge of the electron divided by the time  $t$  for an electron in a single quantum state to traverse the artificial atom while in that level. If  $(dN/d\epsilon)$  is the density of states in the artificial atom, then  $(dN/d\epsilon)eV_{ds}$  is the number of current-carrying channels between the Fermi energy in the source and that in the drain. Thus, the current is given by

$$I = \frac{e}{t} \frac{dN}{d\epsilon} eV_{ds}. \quad (2)$$

The width gives the traversal time,  $t = h/\Gamma$  and  $(dN/d\epsilon) = 1/\Delta\epsilon$ , so the condition for energy quantization is  $R = V_{ds}/I > h/e^2$ , the same as for charge quantization.

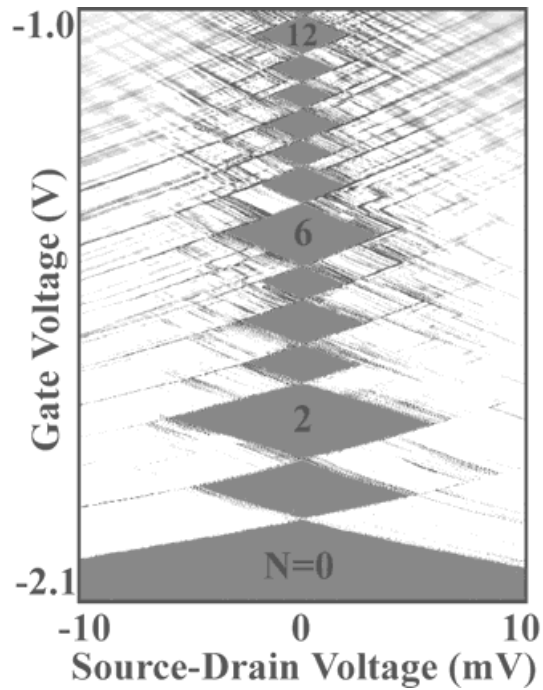
However, while the conditions for charge and energy quantization at zero temperature are the same, charge quantization often survives to higher temperatures. The charge quantization can be observed when  $kT < U$ , but energy quantization requires  $kT < \Delta\epsilon$ . Since  $U > \Delta\epsilon$  for most SETs made to date, energy quantization is more difficult to observe than charge quantization.

Energy quantization can be observed by measuring the variations between peak positions for data like those in Fig. 4. Alternatively, the energy level spectrum can be measured directly by observing the tunneling current at fixed  $V_g$  as a function of  $V_{ds}$ . Suppose we adjust  $V_g$  so that  $Q_0 = -Ne$  and then begin to increase  $V_{ds}$ . The Fermi level in the source rises in proportion to  $V_{ds}$  relative to the drain, so it also rises relative to the energy levels of the artificial atom. Current begins to flow when the Fermi energy of the source is raised just above the first quantized energy level of the atom. As the Fermi energy is raised further, higher energy levels in the atom fall below the Fermi energy, and more current flows because there are additional channels for the electron to use for tunneling onto the artificial atom. One measures the energies by measuring the voltage at which the current increases, or, equivalently, the voltage at which there is a peak in the derivative of the current  $dI/dV_{ds}$ .

One of the most beautiful experiments of this kind is shown in Fig. 5. [10] This is for an SET made in a very different way from that in Fig. 1, such that the confining potential is almost perfectly circular. The diamonds with very low differential conductance are regions where only one charge state is stable. Data like those in Fig. 4 would be obtained by moving along the vertical axis at  $V_{ds} = 0$ . The boundary of the diamonds corresponds to the threshold for changing the charge of the artificial atom. One can overcome the charging energy by changing the source-drain voltage as well as by changing  $V_g$ .

The diagonal lines outside the diamonds correspond to excited energy levels of the artificial atom. In this case, the artificial atom is so small that  $\Delta\epsilon \sim U$  so the peak spacings at  $V_{ds} = 0$  are far from constant and reflect the shell structure of the artificial atoms. One can even see the effects of exchange, that is, the filling follows Hund's rule, making certain values of  $N$  more stable than others.

So far we have identified three different energy scales that are important in understanding SETs:  $U$ ,  $\Delta\epsilon$  and  $\Gamma$ . These are the typical energies, respectively, to add an



**Fig. 5** Differential conductance on a gray scale as a function of both gate and drain-source voltage. The dark diamonds correspond to regions in which there is a gap to current flow. For a SET made of metal all diamonds would have identical size and there would be no variations of conductance outside the diamonds. Semiconductor SETs have diamonds of different sizes and peaks in differential conductance outside the diamonds, corresponding to excited states. These results are from Kouwenhoven *et al.* [10] who have made SETs so perfect that shell structure gives more stability (larger diamonds) for the electron numbers indicated.

electron to the artificial atom, to excite the artificial atom with a fixed number of electrons and the broadening of the artificial atom's energy levels by quantum mechanical tunneling to the leads. Although the Coulomb blockade model is often adequate to estimate  $U$ , we know that it, like  $\Delta\epsilon$  and  $\Gamma$  has a quantum mechanical origin. There is one last energy scale that is even smaller than these three. The tunneling results in a kind of chemical bond between the artificial atom and the leads when the artificial atom has a spin. This is the origin of the Kondo effect, usually associated with magnetic impurities in metals. The Kondo "bond" is very weak in SETs, and was not observed for nearly a decade after theorists predicted it. [11] However, it has now been identified in a number of experiments [6, 7, 12–14], and is the subject of intense experimental and theoretical research.

## 2 History of the SET

The effects of charge quantization were first observed in tunnel junctions containing metal particles as early as 1968 [15]. Later, the idea that the Coulomb blockade can be overcome with a gate electrode was proposed by a number of authors [16–19], and Kulik and Shekhter [20] developed the theory of Coulomb-blockade oscillations, the periodic variation of conductance as a function of gate voltage. Their theory was classical, including charge quantization but not energy quantization. However, it was not until 1987 that Fulton and Dolan [21] made the first SET, entirely out of metals, and observed the predicted oscillations. They made a metal particle connected to two metal leads by tunnel junctions, all on top of an insulator with a gate electrode underneath. Since then, the capacitances of such metal SETs have been reduced to produce very precise charge quantization.

The first semiconductor SET was fabricated accidentally in 1989 by Scott-Thomas *et al.* [22] in narrow Si field effect transistors. In this case the tunnel barriers were produced by interface charges. Shortly thereafter Meirav *et al.* [8] made controlled devices of the kind depicted in Fig. 1, albeit with an unusual heterostructure with AlGaAs on the bottom instead of the top. In these and similar devices the effects of energy quantization were easily observed. [23–25] Only in the past few years have metal SETs been made small enough to observe energy quantization. [26] Foxman *et al.* [24] also measured the level width  $\Gamma$  and showed how the energy and charge quantization are lost as the resistance decreases toward  $h/e^2$ .

In most cases the potential confining the electrons in a SET is of sufficiently low symmetry that one is in the regime of quantum chaos: the only quantity that is quantized is the energy. In this case there is a very sophisticated approach, based in part on random matrix theory, for predicting the distributions of peak spacings and peak heights for data like those in Fig. 4. [27–29] There are challenging problems in this arena that are still unsolved. In particular, there is great interest in how the interplay of exchange and level spacing determines the spin of a small metal SET. [30]

As already mentioned, the data of Fig. 5 are for an SET of sufficiently high symmetry that angular momentum in the plane of the 2DEG is conserved, so shell structure is apparent. Another way to eliminate the scattering that destroys angular momentum conservation is to apply a magnetic field perpendicular to the 2DEG. At sufficiently



high fields elegant patterns are seen in the single-electron-peak positions as a function of field. [31]

The evolution of Coulomb charging peaks with magnetic field have been interpreted with various degrees of sophistication, imitating the development of the theory of atoms. First one tries the “constant interaction model” in which electrons are treated as independent except for a constant Coulomb charging energy. This gives only a qualitative picture of the physics. In order to be quantitative, one needs to at least treat the electron-electron interactions self-consistently (analogous to the Thomas-Fermi model) [32], and for some cases one needs to include exchange and correlations. In particular, it is found that electrons in an SET undergo a series of phase transitions at high magnetic field. [33]. One of these is well described by Hartree-Fock theory, but others appear to require additional correlations.

The future of research on SETs looks very bright. There are strong efforts around the world to make the artificial atoms in SETs smaller, in order to raise the temperature at which charge quantization can be observed. These involve self-assembly techniques [34] and novel lithographic and oxidation methods [35] whereby artificial atoms can be made nearly as small as natural ones. This is, of course, driven by an interest in using SETs for practical applications. However, as SETs get smaller, all of their energy scales can be larger, so it is very likely that new phenomena will emerge.

## Acknowledgments

The work at MIT is the result of the heroic efforts and brilliant insights of a wonderful group of students, postdocs and faculty who have collaborated with M. Kastner. It has been a great privilege to work with them and to be part of the mesoscopic physics community throughout the world. This work was supported by the US Army Research Office under contract DAAG 55-98-1-0138 and by the National Science Foundation under grant number DMR-9732579.

## References

- [1] M.A. Kastner, *Physics Today* **46** (1993) 24
- [2] M. A. Kastner, *Rev. Mod. Phys.* **64** (1992) 853
- [3] U. Meirav and E.B. Foxman, *Semiconductor Science and Technology* **10** (1995) 255
- [4] R.C. Ashoori, *Nature* **379** (1996) 413
- [5] C. W. J. Beenakker, H. van Houten and A. A. M. Staring, in *Single Charge Tunneling* edited by H. Grabert and M. H. Devoret, NATO ASI Series B, Plenum, New York, 1991
- [6] D. Goldhaber-Gordon, H. Shtrikman, D. Mahalu, D. Abusch-Magder, U. Meirav, and M. A. Kastner, *Nature* **391** (1998) 156
- [7] D. Goldhaber-Gordon, J. Göres, M. A. Kastner, H. Shtrikman, D. Mahalu, and U. Meirav, *Phys. Rev. Lett.* **81** (1998) 5225
- [8] U. Meirav, M.A. Kastner, and S.J. Wind, *Phys. Rev. Lett.* **65** (1990) 771
- [9] D. J. Thouless, *Phys. Rev. Lett.* **39** (1977) 1167
- [10] L.P. Kouwenhoven, T. H. Oosterkamp, M. W. S. Danoesastro, M. Eto, D. G. Austing, T. Honda, and S. Tarucha, *Science* **278** (1997) 1788
- [11] T. K. Ng and P. A. Lee, *Phys. Rev. Lett.* **61** (1988) 1768; L.I. Glazman and M.E. Raikh, *Phys. Rev. Lett.* **47** (1988) 452; S. Hershfield, J. H. Davies, and J. W. Wilkins, *Phys.*

- Rev. Lett. **67** (1991) 3720; Y. Meir, N.S. Wingreen, and P.A. Lee, Phys. Rev. Lett. **70** (1993) 2601; N.S. Wingreen and Y. Meir, Phys. Rev. B **49** (1994) 11040
- [12] S. M. Cronenwett, T. H. Oosterkamp, and L. P. Kouwenhoven, Science **281** (1998) 540
- [13] J. Schmid, J. Weis, K. Eberl, and K. v. Klitzing, Physica B **256-258** (1998) 182
- [14] F. Simmel, R. H. Blick, J. P. Kotthaus, W. Wegscheider, and M. Bichler, Phys. Rev. Lett. **83** (1999) 804
- [15] H. R. Zeller, and I. Giaver, Phys. Rev. **181** (1969) 789
- [16] K. K. Likharev, IBM J. Res. Dev. **32** (1988) 144
- [17] D. V. Averin and K. K. Likharev, in *Mesoscopic Phenomena in Solids*, edited by B.L. Altshuler, P. A. Lee and R. A. Webb, Elsevier, Amsterdam 1991
- [18] M. Amman, K. Mullen and E. Ben-Jacob, J. Appl. Phys. **65** (1989) 339
- [19] L. I. Glazman and R. I. Shekhter, J. Phys. Condens. Matter **1** (1989) 5811
- [20] I. O. Kulik and R. I. Shekhter, Zh. Eksp. Teor. Fiz. **68** (1975) 623 [Sov. Phys. JEPT **41** (1975) 308]
- [21] T.A. Fulton and G.J. Dolan, Phys. Rev. Lett. **59** (1987) 109
- [22] J. H. F. Scott-Thomas, S. B. Field, M. A. Kastner, H. I. Smith, and D. A. Antoniadis, Phys. Rev. Lett. **62** (1989) 583
- [23] A.T. Johnson, L.P. Kouwenhoven, W. de Jong, N.C. van der Vaart, C.J.P.M. Harmans, and C.T. Foxon, Phys. Rev. Lett. **69** (1992) 1592
- [24] E. B. Foxman, P. L. McEuen, U. Meirav, N. S. Wingreen, Y. Meir, P. A. Belk, N. R. Belk, M. A. Kastner, and S. J. Wind, Phys. Rev. B **47** (1993) 10020
- [25] J. Weis, R.J. Haug, K. v. Klitzing, and K. Ploog, Phys. Rev. Lett. **71** (1993) 4019
- [26] D. C. Ralph, C. T. Black, and M. Tinkham, Phys. Rev. Lett. **78** (1997) 4087
- [27] Y. Alhassid, P. Jacquod, A. Wobst, Phys. Rev. B. **61** (2000) R13357, and references therein
- [28] A. M. Chang, H. U. Baranger, L. N. Pfeiffer, K. W. West, and T. Y. Chang, Phys. Rev. Lett. **76** (1996) 1695
- [29] J. A. Folk, S. R. Patel, S. F. Godijn, A. G. Huibers, S. M. Cronenwett, and C. M. Marcus, Phys. Rev. Lett. **76** (1996) 1699
- [30] P. W. Brouwer, X. Waintal, and B. I. Halperin, Phys. Rev. Lett. **85** (2000) 369
- [31] P.L. McEuen, E.B. Foxman, U. Meirav, M.A. Kastner, Y. Meir, N.S. Wingreen, and S.J. Wind, Phys. Rev. Lett. **66** (1991) 1926; P.L. McEuen, E.B. Foxman, J. Kinaret, U. Meirav, and M.A. Kastner, N.S. Wingreen, and S.J. Wind, Physical Review B **45** (1992) 11419 ; P.L. McEuen, N.S. Wingreen, E.B. Foxman, J. Kinaret, U. Meirav, M.A. Kastner, Y. Meir, and S.J. Wind, Physica B **189** (1993) 70
- [32] D. B. Chklovskii, B.I. Shklovskii, and L. I. Glazman, Phys. Rev. B **46**, (1992) 4026
- [33] O. Klein, C. de C. Chamon, D. Tang, D.M. Abusch-Magder, S.-G. Wen, M.A. Kastner and S.J. Wind, Phys. Rev. Lett. **74** (1994) 785
- [34] C. B. Murray, D. J. Norris, M. G. Bawendi, J. Am. Chem. Soc. **115** (1993) 8706
- [35] Y. Takahashi, M. Nagase, H. Namatsu, K. Kurihara, K. Iwdate, Y. Nakajima, S. Horiguchi, K. Murase, and M. Tabe, Electronics Letters **31** (1995) 136