

Single Electronics : One Electron, One Bit?

P. Hadley, C. J. P. M. Harmans and J. E. Mooij *Delft University of Technology*

1. Introduction

What will be the principle on which the electronics of the 21st century will be based? The present types of semiconductor devices can certainly be made even smaller than they are now. At some stage, however, on further reduction of the dimensions the underlying physics will stop to function. One cannot talk of a p-doped semiconductor with a sample that is made smaller than the average distance between acceptors. Future fabrication techniques will probably allow control on the nanometer scale. By that time, new device concepts are needed. One promising candidate is "single electronics," circuits in which single electrons are manipulated. They have the potential of providing fast, compact logic circuits where a bit of information is represented by a single electron. Clearly this is one limit to what is possible in terms of scaling down the size of electronic circuits.

Charge is quantized; it comes in units of e . The discreteness of charge plays a clear role on the level of an atom, where the energy needed to remove an electron is known as the ionization energy. Single electronics makes use of the analogous "ionization energy" of a conductor, called the charging energy $E_C = e^2/2C$, where C is the electrical capacitance. In conventional electronic circuits the charging energy is so small that it can be neglected. As dimensions decrease, the charging energy increases. The now available techniques allow fabrication of devices with dimensions below 100 nm. The capacitance of such structures is of order 10^{-16} to 10^{-15} farad; the associated charging energy is 0.1 to 1 meV. Charging effects are observed in these circuits when they are cooled to temperatures where the thermal energy $k_B T$ is lower than E_C . By cooling to below 1 K the principles of single electronics can now be demonstrated. In recent years remarkable progress has been made.

The basic charging effects have also been observed at room temperature on small metallic particles with a scanning tunneling microscope. However, such particles cannot yet be arranged in circuits. Single electron devices and circuits have been realized for low-temperature operation with metallic tunnel junctions and with high-mobility two-dimensional-electron-gas semiconductor systems. We will discuss their fabrication and properties and will show the results that have been obtained so far. We will restrict ourselves to the elementary aspects. The reader who is interested in more detailed information is referred to

references at the end.

Single electronics is clearly still in its infancy. As controlled fabrication becomes possible at smaller and smaller dimensions, the useful range of temperatures, frequencies and voltages will expand. Single electronics may become practically feasible in an acceptable range of conditions by the time that conventional semiconducting electronics starts to lose its basis for functioning.

2. Basic Elements

Small-capacitance conducting "islands", with well-defined charge but accessible for transport of electrons, can be formed between the barriers of two tunnel junctions in series. When the junction resistance is high, conduction through these tunnel junctions is discrete; the electrons are clearly on one side of the tunnel barrier before the tunneling event and on the other side afterwards. The important capacitance, C_Σ , is the sum of all the capacitances of the island to nearby electrodes and to ground. In reproducibly fabricated samples the capacitance between an island and an electrode that are just separated by a tunnel barrier ranges from 10^{-16} to 10^{-15} F. The capacitance with electrodes that are too far away for tunneling is typically 10^{-18} to 10^{-15} F. The stray capacitance to ground of a micron-sized island is of order 10^{-18} F.

The resistance of the tunnel junctions has to be higher than the "quantum resistance" h/e^2 , of order 25 k Ω (h is Planck's constant). Why this is necessary can be explained from the Heisenberg uncertainty relation saying that the product of the uncertainties in energy and time is equal to h . An energy E_C loses its importance for time scales shorter than h/E_C . In circuits, the typical response time is $R_i C$, where R_i is the junction resistance. This response time has to be much longer than $h/E_C = hC/e^2$. The condition follows: $R_i \gg h/e^2$. The typical value for R_i in single electronics is 100 k Ω . This unavoidable limitation is important as it leads to a response time that is now 10^{-10} s for capacitances of 10^{-15} F, decreasing to 10^{-13} s for future capacitances of 10^{-18} F.

The basic element of single electronics is the single electron transistor, that consists of a conducting island between two tunnel barriers, capacitively coupled to a gate electrode. The overall transport through the two tunnel junctions can be controlled with the gate, in a way that will be explained in the following paragraphs.

We consider junctions with resistances of 100 k Ω or more, at temperatures where $k_B T \ll E_c$. Tunneling of an electron in a specific junction in the circuit is then only possible when the energy of the circuit is lower after the tunneling. The circuit energy contains the electrostatic energy of all the capacitors in the circuit as well as the free energy of the voltage sources that changes when charge moves in or out of them. This simple rule follows from the "orthodox" theory that Likharev and Averin have developed for these circuits. It is not enough to consider only the junction where tunneling takes place. In fact, a single junction that is voltage biased does not exhibit the charging effects. Before and after tunneling the electrostatic energy of the junction-capacitor is the same; the source energy has decreased by eV when one electron has passed. So the overall energy change is always negative, even for small voltages and in very small junctions, and tunneling is always allowed in the easy direction. Current biasing of junctions is difficult, as the lead capacitance usually shunts the high impedance source.

In Fig. 1 the voltage-biased double junction is illustrated. In 1a the voltage distribution is shown when the charge on the middle island is zero. The energies of the source and the two junction-capacitors in this situation are also indicated. The source energy is negative because a charge $CV/2$ has moved to the capacitor. The question now is whether this zero-charge state is stable against the tunneling of an electron from the middle island to the top electrode (equivalent to a positive charge e moving from the source to the island). In 1b the energy is calculated after tunneling. The energy difference ΔE associated with

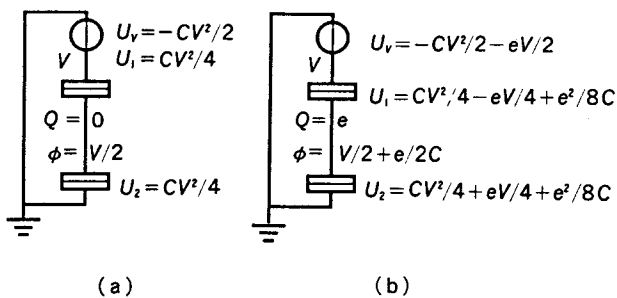


Fig. 1 Quantitative illustration of the basic principle of single electronics: Double junction system with voltage bias. To determine whether a particular tunneling event will occur, the energies before and after the event have to be calculated. In (a) the charge on the middle island between the junctions is zero, the total energy of capacitors and source is zero. In (b) an electron has tunneled from the island to the positive electrode. The source energy has now changed by $-eV$ because of the entering electron and by $+eV/2$ because the charge on the top capacitor is less positive. The total energy change for this tunneling process is $-eV/2 + e^2/4C$; tunneling will only occur for $V > e/2$. Possible other tunneling events can be analyzed in the same way.

the tunneling event turns out to be $\Delta E = -eV/2 + e^2/4C$. For $V > e/2C$ this ΔE is negative. So, for $V < e/2C$ this tunnel event will not take place. On analysis one finds that similarly the tunneling of an electron into the island from the source will not take place unless $V < -e/2C$. The region $-e/2C < V < e/2C$ is the Coulomb gap, where tunneling is forbidden. For $V > e/2C$ an electron tunnels from the island to the positive electrode. Immediately after, an electron will tunnel from the negative electrode to the island, resulting in a net transport of e through the circuit. The charge on the island is zero again. Obviously this cycle repeats and current keeps flowing. The current is lower than in the "normal" state because each electron has to wait until the previous electron has tunneled out of the island. The solid line of Fig. 2 gives the typical I - V characteristic, as measured. The suppression of the tunnel current around zero voltage is known as the "Coulomb blockade."

When a capacitive gate is coupled to the island, the Coulomb gap can be suppressed. For symmetric biasing, a charge $C_g V_g$ is drawn to the gate capacitor C_g for a gate voltage V_g . Now tunneling of the first electron occurs when V exceeds a value $(e + 2C_g V_g)/C_\Sigma$, where $C_\Sigma = 2C + C_g$ is the sum capacitance of the island. For $C_g V_g = e/2$ there is no Coulomb gap. As the center island charge can vary with e by tunneling, it turns out that the gap is suppressed for $C_g V_g$ equal to any odd multiple of $e/2$, and maximum for even multiples. The Coulomb gap is a periodic function of V_g with period e/C_g . In this way C_g can be measured. Figure 2 shows the current-voltage characteristic for the transistor with maximum (solid line) and minimum (dashed line) Coulomb gap, as measured with metallic tunnel junctions. The gap gives the value of C , the junction resistance follows from the slope. So all parameters can accurately be measured and used for modelling. The transistor can be switched from the Cou-

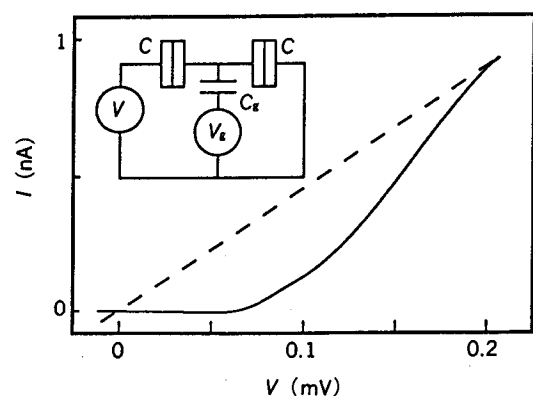


Fig. 2 Current-voltage characteristic for a single electron transistor at low temperature. Solid line: gate adjusted for maximum Coulomb gap. Dotted line: gate adjusted for minimum gap. The junctions are metallic, their capacitance is 1 fF. The Coulomb gap is suppressed when the charge on the gate capacitor is a half-integer electron charge.

lomb-blocked to the conducting state with the gate voltage. The gate draws no continuous current under any condition.

The previous discussion was intended to explain the basic principle on which single electronics circuits are based. Simple electrostatics is enough to determine whether charge will move, given the right parameters. Unwanted processes have to be considered when the parameters are marginal. If the temperature is not low enough, compared with the charging energy, there is a chance that tunneling will occur even if the energy increases. This chance is proportional to $\exp(-\Delta E/k_B T)$. If the junction resistances are not high enough, tunneling may occur through two junctions at the same time, through a second order quantum tunneling process called co-tunneling.

3. Metal Tunnel Junctions

Many charging effect devices are presently fabricated from normal metal tunnel junctions. A tunnel junction consists of two metallic regions separated by a thin insulating barrier. The insulator must be thin enough that electrons can tunnel quantum mechanically through it. A tunnel junction can be thought of as a capacitor in parallel with a tunnel resistance. The most commonly used materials for the tunnel junctions are aluminum/aluminum-oxide/aluminum. The preference for aluminum is due to the uniform, pinhole-free barriers that can be made with aluminum oxide. In the conventional process, the junctions are made by shadow evaporation. The aluminum is evaporated in two sequential evaporations from two angles using a free-standing bridge as a shadow mask. After the first Al evaporation the barrier is formed by oxidizing the aluminum at room temperature in about 1 mbar of oxygen. The oxidation pressure determines the oxide thickness and, thus, the resulting junction resistance. An oxygen pressure of about 1 mbar produces a junction resistance of $1 \text{ k}\Omega\text{-}\mu\text{m}^2$. The second Al evaporation then forms the upper electrode of the junction. A schematic drawing of the resulting junctions is shown in Fig. 3. The junction is formed where the two Al evaporations overlap. Both evaporations and the oxidation are done in the same vacuum cycle to minimize impurities.

The junction capacitances are the product of the specific capacitance and the area of the junction. The specific capacitance is determined by the thickness and the dielectric constant of the insulating barrier. Typically, the specific capacitance is about $100 \text{ fF}/\mu\text{m}^2$. The junction areas are limited by lithography. The smallest junctions that have been made reproducibly have areas about $0.005 \mu\text{m}^2$. This corresponds to a capacitance of 0.5 fF . It is expected that using the same technique, junctions with dimensions of $20 \text{ nm} \times 20 \text{ nm}$ will be fabricated. These junctions will have a capacitance about $5 \times 10^{-17} \text{ F}$ and should exhibit charging effects to temperatures of 10 K . Other lithographic techniques, such as scanning-tunneling-

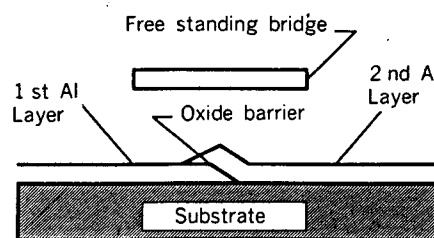


Fig. 3 The two aluminum layers are evaporated in two sequential evaporations using the bridge as a shadow mask. After the first Al evaporation the tunnel barrier is formed by oxidizing the aluminum. The bridge is supported by regions not shown in this cross section.

microscope based lithography, should allow for even smaller metallic junctions. The physics of charging effects in metal junctions will not change until the junctions are only a few atoms wide.

Small capacitance islands necessary for single-electronics can be formed by putting two metallic tunnel junctions in series. It is also important to be able to couple true capacitors to the small islands so that the Coulomb blockade can be modulated. Capacitors with no leakage both larger and smaller than the tunnel junction capacitances are needed. Small capacitances in the range 10^{-18} – 10^{-16} F are made simply by placing a metal electrode near one of the conducting islands. Somewhat larger capacitances can be made using interdigitated fingers. For capacitances of the order of the tunnel junction capacitance or greater, overlap capacitors are used. These are essentially the same as junctions where a very thick oxide layer is used to suppress tunneling.

4. Semiconductor Structures

The main difference between metallic and semiconductor systems is the density of electrons. A typical metallic island contains of order 10^{10} electrons, in contrast with 100 to 1000 electrons for its semiconductor counterpart. The strongly reduced screening of the electric fields in a semiconductor allows for in-situ control of the density on a very local scale. We will first discuss some of the typical elements employed in single-electron semiconductor device structures. Next some recent experimental results will be given that demonstrate the correspondence to metallic systems as well as the key differences.

The majority of single electron studies have been performed on GaAs-AlGaAs heterostructures containing a high-mobility two-dimensional electron gas (2DEG) at the interface between the two materials. The GaAs-AlGaAs heterostructures are grown by molecular beam epitaxy or chemical vapour deposition; the two-dimensional electron gas or 2DEG is introduced by modulation doping of donors in the AlGaAs layer. The Fermi wavelength of the electrons is 20 to 50 nm. At low temperatures the mean

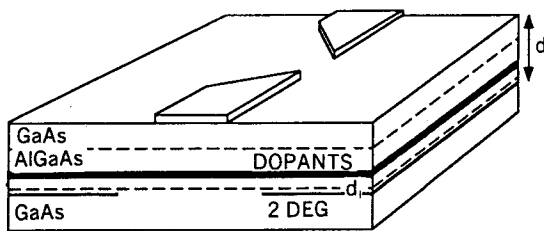


Fig. 4 The two-dimensional-electron gas (2DEG) lies at the interface between GaAs and AlGaAs at a distance, $d \approx 100$ nm below the surface. The two metallic gates on the surface are used to expel the electrons beneath them and thereby structure the 2DEG.

free path is extremely large. Electronic states are coherent over distances that exceed the fabricated dimensions.

Typically the 2DEG lies 100 nm below the surface. It can be locally modulated by nanostructured gates (*Fig. 4*). For the so called split-gate configuration, shown in the figure, a negative voltage of typically -0.5 V leads to depletion of the 2DEG below the gates, confining electron transport to a path of variable width between the two gates. These gates are fabricated with electron beam lithography and evaporation of Ti and Au followed by lift-off. Gates are defined with feature sizes down to 20 nm. The transmission through this gate-defined structure is strongly dependent on the gate voltage. This structure is known as a quantum point contact, because its conductance varies discontinuously in quantum steps. For single electronics, the point contacts are used in the lowest-transmission mode where the effective resistance exceeds the quantum resistance. This means that electronic states are no longer extended through the contact and transport occurs by tunneling only.

There are some important differences between these semiconductor tunnel junctions and metallic ones. One is very important for the device applications: the barrier height can be varied by varying the split gate voltage. So, in the operation of the device electrons can either be confined or given the chance to tunnel. This is an extremely interesting possibility that has no equivalent in metallic systems where, once fabricated, the barrier is fixed. Metal oxide barriers are very thin and very high, so that their transmission is independent of the electronic energies. The sensitivity of semiconductor barriers has as a less favorable side: the transmission is not only influenced by the split-gate voltage, but also by the bias voltage and by voltages on neighboring gates.

The isolated island that is needed for single electronics is obtained by placing two point contacts in series, leaving a small area in between. *Figure 5* shows a typical example, with the two point contacts complemented by two center gates to confine the electrons. The size and shape of the island, and so the number of conduction electrons residing in it, can be controlled by the two center gates. The electronic states are coherent quantum states that extend over the island. The confined electron system is called a

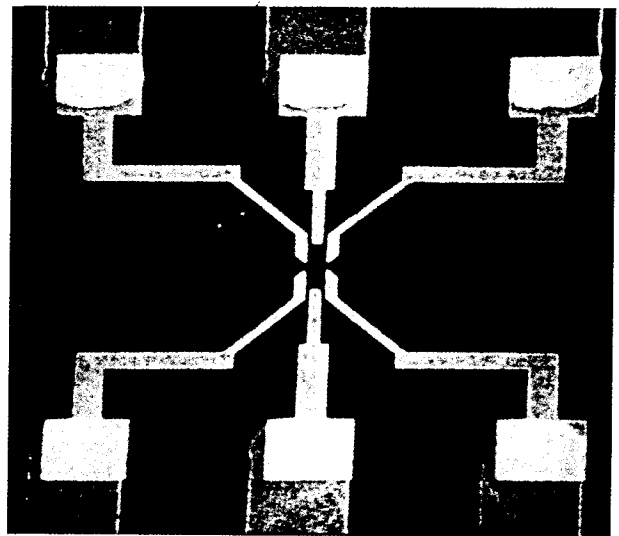


Fig. 5 An electron microscope image of a quantum dot. Shown are the two center gates that are used to adjust the size of the quantum dot and the split gates which are used to control the tunnel barriers between the quantum dot and the wide 2DEG areas.

quantum dot.

In a typical lateral device the diameter of the dot is about 500 nm; it contains approximately 1000 electrons and has a capacitance of 0.2 fF. The charging energy is larger than the average distance between the energy levels of the confined electrons. The smaller varieties approach 100 nm diameter with a capacitance of 0.05 fF, equivalent to a Coulomb energy of approximately 1.5 meV. While in the former the size is large compared to the depletion length, in the latter the effect of the limited screening results in an approximately parabolic confining potential with a reduced Fermi energy in the centre of the dot. This leads to a stronger reduction of the number of electrons in the dot, with the smaller dots containing only some 20 electrons. Here the level separation is typically larger than the charging energy. The size and shape of all quantum dots of this kind can be changed with the center gates. Application of a magnetic field also has a pronounced influence on the electronic states.

This is not the only possible configuration for a quantum dot. Based on multi-quantum well materials, with the bandgap discontinuity employed to form the barriers, so-called vertical double barrier resonant tunnel devices (DBRT's) have been made. These structures are fabricated by e-beam lithography followed by (reactive ion beam) dry etching to form the vertical pillars, with one contact made at the top of the pillar. These device types have a stronger confinement. Although the electrostatics of the quantum dot can be controlled by the addition of a gate surrounding the pillar, they are less flexible given the fixed as-grown barriers. More important, the vertical construction is maybe less ideal for realising integration of larger and more complex multi-dot systems.

Coulomb oscillations in semiconductor systems were

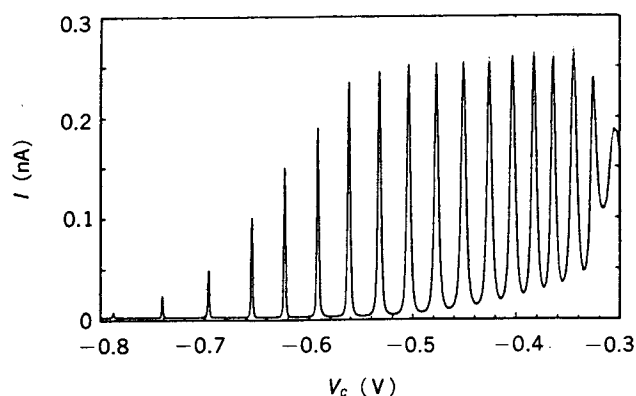


Fig. 6 Current through a quantum dot as a function of the gate voltage. The current peaks imply the existence of zero-dimensional states in the quantum dot.

discovered in lateral silicon devices in 1989 by Scott-Thomas and co-workers from MIT, USA. The overall shape of the characteristics strongly resembles those obtained in metallic junction systems. More recent results on a 600 nm diameter quantum dot in a GaAs-AlGaAs 2DEG are shown in *Fig. 6*. In these experiments the effect of the zero-dimensional states on the transport properties is studied. If the center gate voltage is adjusted such that an energy state in the dot lines up with the external Fermi energy of the large 2DEG areas connected to the dot via the two barriers, electrons can resonantly tunnel through the dot. This leads to a current if a small bias voltage is applied over the dot. It should be stressed that this process is a single quantum event, in contrast to the metallic case. This coherent resonant tunnelling is shown in *Fig. 6* for the dot at 50 mK in an applied magnetic field of 7 T. For the classical or sequential tunnelling process the total transmittance is given simply by the product of the transmittances of the two point contacts, indicated by the dashed line. In contrast the actual Coulomb oscillation amplitude is considerably larger and is well described by an expression containing multiple phase coherent reflection at the entrance and exit point contact barriers of the dot, much the same way as in an optical Fabry-Perot interferometer. This shows the coherent nature of the transport process, confirming the existence of the zero-dimensional eigenstates.

The results so far are mostly with GaAs-GaAlAs 2D systems. However, the very small size of the quantum dot elements that can be obtained with split gates should allow comparable phenomena in other material systems. In particular the recent rapid improvements in mobility obtained in the strained-lattice SiGe-Si system indicates it as an attractive alternative, even for coherent applications.

5. Possible Applications

Useful single electronic circuits have already been demonstrated. The simplest circuits, which consist of just

a few low capacitance islands and gates, are able to measure charge with unprecedented precision. As the technology progresses it is expected that single-electron devices will provide us with a range of quantum-noise-limited detectors. A photodetector with single photon sensitivity has been demonstrated. Slightly more complicated single-electron circuits are being used to make a fundamentally new current standard. The current is determined by counting the number of electrons that go by. Logic circuits have also been proposed. In single-electron logic, one bit of information is represented by a single electron. Extremely large integration scales, low dissipation, and fast switching will be possible with this logic. Much development work needs to be done before single-electron logic can be operated at a reasonable temperature. The work is now proceeding on improving the single-electron devices that measure charge which will be necessary to monitor the operation of single electron logic.

The classic example of a single-electron circuit is the single electron transistor that was introduced above. The input voltage is applied to the gate and the output voltage is measured across the two junctions in series. When the double junction is current biased with a small current, the voltage across the double junction is periodic in the gate voltage with a periodicity, e/C_g . At zero temperature the periodic waveform would be a sawtooth with a slope $C_g/(C + C_g)$ on one side and $-C_g/C$ on the other. At higher temperatures the corners of the sawtooth become rounded. If the middle of the negative-going slope is taken as the operating point, then a small variation in the gate voltage results in a variation $V_{out} = (C_g/C)V_g$ across the double junction. In other words, there is a voltage gain of C_g/C . Experimentally, a voltage gain of 2.8 has been achieved with metallic junctions and a gain of 5 in semiconducting systems. With the foreseeable advances in lithography, single-electron transistors with a gain of 100 should become available in the next few years.

A single-electron transistor has a large input impedance and is ideally suited to measure charge. Transistors designed to measure charge are referred to as electrometers. An electrometer need not have a voltage gain greater than one to function. It is the very large power gain that is used in an electrometer. A small fraction of the elementary charge, e , at the input produces a measurable change in the output voltage. The sensitivity of the best electrometers made so far is on the order of $10^{-4} e/\sqrt{Hz}$. This sensitivity is about one million times greater than is attainable with commercially available electrometers. Theoretical estimates of the ultimate sensitivity of the electrometer are $10^{-6} e/\sqrt{Hz}$. This is comparable to another quantum-noise-limited detector, the dc SQUID, which is capable of measuring magnetic flux with a sensitivity of $10^{-6} \Phi_0/\sqrt{Hz}$.

Stray capacitances are an important consideration when using an electrometer. The charge that is to be measured is put on the input lead, that is, the wire that leads up to the gate capacitor. This charge sees the gate capacitance in parallel with the stray capacitance of the input lead. For

the proper operation of the electrometer, the gate capacitance should dominate. The gate capacitance is on the order of 1 fF, which limits the stray capacitance of the input lead to something less than this quantity. In practice this restricts the use of an electrometer to circuits where it can be placed a few microns away from the charge that is to be measured. In other words the electrometer must be integrated into the same circuit that is to be measured. Even with this restriction, the utility of the single-electron electrometer has been demonstrated in several experiments. Electrometers are used to measure the tunneling of single electrons on and off small metal islands in more complicated single-electron circuits. These results are important for evaluating the role of defects and of higher order quantum processes in circuits.

The idea behind using single-charge tunneling devices to make a current standard is that we can modulate the Coulomb blockade of a circuit in such a way that exactly one electron passes through the circuit. By doing this over and over with frequency f , one can generate a current proportional to the frequency, $I = ef$. All of the current standards consist of at least two small capacitance islands in series. A gate capacitor is coupled to each of the islands. By adjusting the voltage on the gates one can selectively suppress the Coulomb blockade of the various islands. If the current is to flow from left to right then the gates are adjusted so that it becomes energetically favorable for one more electron to tunnel on to the rightmost island. It is possible to adjust the gates so that tunneling to any of the other islands is forbidden by the Coulomb blockade. Tunneling is a quantum process that occurs at a certain rate (on the order of $(R,C)^{-1}$). The gates are held in the situation where it is favorable for an electron to tunnel for a long enough time that it is virtually certain that an electron has tunneled. The gates are then adjusted so that it is energetically favorable for the electron to tunnel to the next island to the left. This process is continued until the electron is brought entirely through the array of low capacitance islands. Then the process is started over again.

Current standards have been made from metal tunnel junctions and semiconductors. **Figure 7** shows the current through a four junction circuit (called a turnstile) as a function of the bias voltage. Note that there is a finite bias voltage range where the current is equal to ef . The turnstiles that have been made until now have been modulated at about 10 MHz. This results in a very small current of a few picoamps. The turnstile is operated so slowly to be certain that the electrons have tunneled. The current that is measured in an experiment is never exactly the electron charge times the frequency. There is always a small discrepancy, $I = ef + \delta I$. The precision that has been achieved so far is $\delta I / (ef) = 10^{-3}$. This extra current is due to thermal activation over the Coulomb blockade and to quantum co-tunneling through the Coulomb blockade. More complicated circuits are being investigated that will reduce the effects of these unwanted processes. Many of these devices in parallel would then produce a current of a more manageable magnitude. The precision that has

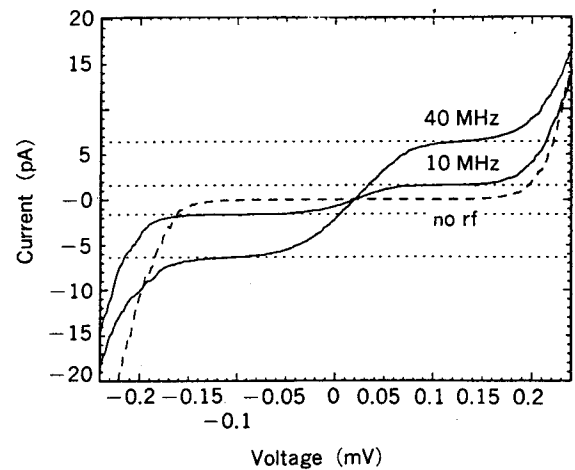


Fig. 7 The current through a turnstile as a function of the bias voltage for three different drive frequencies. There are plateaus in the current where $I = ef$.

been achieved is worse than what can be obtained with a conventional current standard. However, these are the first measurements and there is optimism that someday the amp will be defined by a device that regulates the motion of electrons one at a time. That would give the current standard a comparable fundamental foundation as the quantum-Hall-effect resistance standard and the Josephson voltage standard.

Since it is possible to make transistors based on charging effects, it is also possible to make logic circuits. Two types of single-electron logic have been proposed. One proposal is to use logic circuits similar to those used in semiconductor circuits, merely replacing the semiconductor FETs with single-electron transistors. The main advantage here would be that since the single-electron transistors are smaller than ordinary transistors, more could be put on a chip. In this scheme, bits of information would still be represented by voltages as they are in ordinary semiconductor logic. The other proposed type of single-electron logic has a different character. Rather than representing digital bits as dc voltage levels, a bit would be represented as the presence or absence of a single electron. In this way we take advantage of the natural digital nature of charge. The basic building block consists of four tunnel junctions in series as is shown in **Fig. 8**. It is essentially the same device that is used in the current standard application. In the logic application the gates are not modulated periodically, rather the voltages that appear at the inputs determine whether an electron passes through the device. The element shown is an RS flip-flop and is used as a one-bit storage element. The capacitances and bias voltages are chosen so that a Coulomb blockade exists at every island and no motion of electrons is possible. If a voltage then appears at S, it suppresses the Coulomb blockade on the upper island and one electron tunnels to the central island. This electron represents a bit of information. The electron will stay on the central island even if the voltage at S

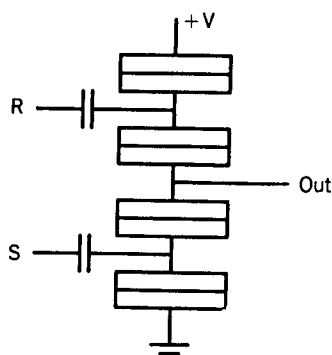


Fig. 8 Schematic drawing of an RS flip-flop.

returns to its original value. The flip-flop can be reset by applying a voltage to R. This suppresses the Coulomb blockade of the lower island and allows the electron to tunnel away. Other logic elements have been proposed, but until now no single-electron logic elements have been tested. There is still much work to be done to determine the limits of this logic. Speculation on the ultimate performance of these devices will be discussed in the next section.

In the logic circuits just described it was assumed that the tunnel barriers were held constant during operation. In the semiconductor devices the tunnel barrier heights can be adjusted. This is essentially the same thing as adjusting the source-drain conductance of a conventional FET. Tunnel barriers can be removed or new barriers introduced merely by changing the voltage on a gate. Thus, the hardware that performs a calculation could be adjusted based on the result of a previous calculation. Experiments where the tunnel barriers in semiconducting single-electronic circuits were modulated at megahertz frequencies have been performed. These experiments show that the combination of single electronics and conventional FETs allow for new and intriguing possibilities.

6. Problems, Limitations

The basic principles of single electronics are well understood, the first simple circuits have been realized in practice and have been found to work. It is, however, much too early to decide whether practical large-scale circuits can really be developed. Only a very limited effort by a small number of groups, mainly at universities, has been devoted to the subject. These experiments have identified some problems that need to be solved.

A major point of difficulty is the presence of "offset charges" on islands between tunnel barriers. When measurements are performed on single-electron transistors, one typically finds that the maximum Coulomb gap does not occur for zero gate voltage, but at some offset value. If there are more islands in a circuit, they all appear to have a different offset charges, that have to be compensated for the circuit to function. These random offset charges are

probably caused by charged defects in the neighborhood of the conductors, that induce fractional surface charges on islands in much the same way as the fabricated intentional gates do. Another cause of offset charge might be connected with the different work functions of different metal crystal faces. No systematic study has been performed yet. It is not clear whether fractional charges are completely random over a range that exceeds e , or are limited to smaller values. It has been suggested that offset charges could be compensated by mobile charges in a semiconductor substrate that are frozen in on cooling down. So far, in circuits like the six-junction turnstile, every island needs a separate gate that, except for the center island, serves only to compensate the offset charge. Cross-capacitances make this compensation a difficult procedure. In circuits with many junctions and islands it seems hardly practical to put adjustable gates everywhere. So, the offset charges will have to be eliminated or circuitry has to be designed where the functioning is independent of their presence.

A problem that seems related to that of offset charges is that of $1/f$ noise and the stability of the circuits. One often finds that during a measurement the current or voltage switches. One finds, on analysis, that the optimum gate voltage for maximum Coulomb gap has suddenly shifted. The behavior is consistent with an instantaneous shift in position of a charged defect. Because of the large amplitude (corresponding to large fractions of e) and the sudden appearance it is probable that the shifting charge is in or very near the tunnel barrier. A large $1/f$ component is often found in the noise that seems to have the same origin. A systematic study of these effects is needed. The experience obtained in silicon technology with defects in FET structures may be valuable here.

The junction fabrication techniques that are now employed: shadow evaporation for metallic junctions and multiple-gate confinement for semiconductors, are not attractive for larger-scale integration. New fabrication techniques will have to be developed. One of the interesting directions is towards lateral junctions, consisting of metal-tunnel electrodes in the same plane with a very small separation.

The response time of the circuits is, as stated before, limited by the R_iC time of the junctions and the other elements. The value of R_i cannot be lowered, as the principle of charge localization would be violated. So, R_i will be at least 100 k Ω in all circuits based on these principles. The response time will go down as capacitances are reduced with improved micro- or nanofabrication.

7. Future Prospects

Single electronics is not for this century. Only for a limited number of small-scale applications, most of them in physics laboratories, will there be interest in the next decade. There is, however, for those that are interested in the long-term future of microelectronics a good reason to

follow the development of single electronics in the coming period: In single electronics one encounters limits and practical problems that may be very relevant for all future devices with dimensions on the scale of nanometers. Single electron devices offer a sensitive test for nanofabrication techniques, for the control of defects on such scales and for the understanding of the coupling of quantum effects with devices that are within microscopic distances from each other. Single electronics may be of direct long-term interest as well. It offers a potentially feasible operation mode for logic and memories on the nanoscale.

The future of single electronics completely depends on the further development of nanofabrication techniques. It is to be expected that reproducible structures on the order of 10 to 20 nm will be produced by electron beam lithography within the next 5 years. This will lead to capacitance values around 10^{-17} F, and a charging energy of 10 meV which is equivalent to a thermal energy of 100 K. For logic and memories, or for a current standard, one needs a very strong suppression of unwanted tunneling processes. Therefore the operating temperature would still have to be about 20 K. The response time would be around 1 ps.

A clear effort has started in several laboratories towards non-lithographic fabrication of structures on the scale of a few nanometers. Scanning-tunneling-microscope techniques are employed, as well as very narrowly focused electron beams that are used for direct deposition. When real nanofabrication comes available, the capacitance values of structures can be of order 10^{-18} F. The charging energies will be 100 meV, operating temperatures go up to more than 100 K.

What will be the performance of such future single-electronics elements? A common figure of merit for logic elements is the product of the switching time and the power needed for switching. Short switching times and low power dissipation are both desirable. In single electronics the switching time-power product is the charging energy E_C , with a future limit of about 100 meV. There is a fundamental limit, set by quantum mechanics, that does not allow one to decrease dissipation and increase speed forever. This limit can be expressed as $P_{\min} \tau^2 = h$, where $P_{\min}$ is the minimum power dissipated and τ is the switching time. Thus, there is a trade-off between power and

speed. The only way we can increase the speed is to increase the power dissipation. The switching time will be 100 k Ω times the capacitance limit of 10^{-18} F, so around 0.1 ps. The power for switching is E_C divided by the switching time, some 0.2 μ W. Can such numbers be beaten by other types of devices? The product of the switching time and the switching energy is $(hC/e^2)(e^2/2C)$, close to Planck's constant, h . Single electronics are already working at the quantum limit. No other device can be better, given the switching time. Others may be faster, but will need to dissipate more energy.

All future nano-elements will exhibit quantum effects. They will still be macroscopic objects containing many elementary particles, but experiments have demonstrated that macroscopic objects obey quantum mechanics just as surely as subatomic particles do. If you drive your car at a wall there is a vanishingly small, but finite probability that the car will quantum mechanically tunnel through the wall. One can also calculate the probability that an entire electronic circuit will quantum mechanically tunnel to another charge configuration. For conventional circuits the chance of this happening can be neglected. However, as the circuits get smaller, this probability can form a measurable contribution to the error rate. Experiments demonstrating the tunneling of the charge configuration of entire circuits (as opposed to just the tunneling of a single electron) have been performed. This phenomena is not unique to single electronics, where it is now studied, but generally occurs in all nano-fabricated systems. What is learned with single electronics will, at least in part, apply to them all. It seems that at this time single electronics offers the most direct access to this new world.

References

General review articles with references to the original publications are found in the Proceedings of the NATO advanced study institute on *Single Charge Tunneling*, edited by H. Grabert and M.H. Devoret, Plenum Press 1992.

More specialized research papers are published in the special issue of the *Zeitschrift fur Physik B* of June 1991, Volume 999.

A review paper on digital applications of single electronics is "Correlated discrete transfer of single electrons in ultras-small tunnel junctions," by K.K. Likharev, IBM J. Res. and Development, Vol. 32, p. 144 (1988)