Electronic transport in small strongly localized structures

by A. B. Fowler
J. J. Wainer
R. A. Webb

We review some recent results on the low-temperature transport properties ($T < 4$ K) of very small silicon metal-oxide field-effect transistors in the insulating regime of conduction. Our devices are lithographically patterned to have widths as small as 0.05 μm and lengths as short as 0.06 μm. These small transistors exhibit new and unexpected sample-specific fluctuation behavior in the gate voltage, temperature, and magnetic field dependence of the conductance. We discuss both resonant tunneling and Mott variable-range hopping, the two main transport mechanisms in these devices at low temperature.

Introduction

Over the past seven years we [1–7] and others [8, 9] have been studying the conductance of small (mesoscopic) systems at temperatures below a few degrees Kelvin, where tunneling of electrons from one site to another is the dominant transport mechanism. Tunneling is only observed when the electron wavefunction is confined by the potential disorder to a small region ξ, the localization length, in the sample. When ξ is much smaller than the sample length L, the system is strongly localized. In this case the electron wavefunctions are of the form $\psi \sim \exp -(x/\xi)$, where in the two-dimensional case [10–14] (width w, $L \gg \xi$), ξ is about 20–70 nm [15]. Samples of width $w \sim \xi$ are defined to be strictly one-dimensional. At low temperatures, $T < 0.1$ K, the number of localized electron energy states that contribute to the total electrical conductivity of small samples can be quite small, much less than 50 in many cases. This means that the transport properties are not determined by some average statistical behavior that characterizes the localized states, but are dominated by the individual character of a few localized states. These states in turn are determined by local microscopic variables such as impurities, defects, etc. Therefore, the conductance of one sample may be very different from that of a nominally identical sample.

Our experiments employ metal-oxide-semiconductor (MOS) silicon structures in which the electrons are confined within a 5-nm region of the semiconductor surface by perpendicular electric fields and are laterally confined by lithography. By changing the voltage on the metal gate, the energy of the electronic states in the surface channel relative to the chemical potential or Fermi energy can be altered over a wide range. At large gate voltages the transport becomes metallic in character and the electrons diffuse from the source to drain contacts. At low gate voltages the electrons are localized in the band tail of the two-dimensional conduction band by random charges in the oxide. The ability to vary the chemical potential or Fermi

energy (μ) relative to the energy of the localized states permits detailed examination of the various conduction mechanisms. In devices where both the length and width are much greater than the localization length, two-dimensional strong localization effects have been studied for more than twenty years [10–13]. It is well established that Mott variable-range hopping (VRH) correctly describes the conduction mechanism in these samples. For a general overview of localized transport, see Shklovskii and Efros [14].

There are only three types of conduction processes that one might expect to occur in the strongly localized regime at temperatures below ~1 K. These are shown schematically in **Figure 1.** The first is thermally activated tunneling from one localized site to another, or Mott hopping, and the conductance from such processes decreases exponentially with decreasing temperature. The second is direct tunneling of the electrons from the source contact to the drain contact. In the samples we studied, the separation between the source and drain contacts is greater than 500 nm, and this process does not appear to be important. The third and most important low-temperature transport process is that of resonant tunneling, a process whereby an electron directly tunnels elastically from the source contact to one localized state in the sample and subsequently tunnels elastically out of that state to the drain contact. The last two tunneling processes can occur at zero temperature.

In the case of resonant tunneling, the energy of the resonant localized state must be equal to the chemical potential in the Fermi surface of the contacts for maximum tunneling probability. This probability for electron transmission decreases very rapidly when the chemical potential changes by an amount greater than the energy linewidth of the resonant state. Although it is perhaps not so obvious, the same is true for hopping. However, if the number of paths through the sample is sufficiently limited that the number of available states within $k_B T$ of μ is small, only a few hops will dominate the resistance, and this resistance will be a strong random function of the relative values of μ and the energies of the states. In both resonant tunneling and hopping, each and every sample has its own characteristic and different dependence of conductance on μ or on gate voltage. There is no ensemble averaging in this case.

In the following sections we discuss our experiments on resonant tunneling and Mott hopping. Others have observed similar results in similar samples, but in this paper we discuss primarily our own work, which has previously been reported. Results rather than experimental techniques or derivations are emphasized.

Samples where hopping dominates

The appropriate mesoscopic samples for studies of hopping conductivity are relatively long but narrow Si accumulation layers. If the most probable hopping length, R_0, is long compared to the geometrical width of the sample, the conduction is quasi-one-dimensional. Hopping lengths for two dimensions are given by $R_0 \sim [\xi/TD(E)]^{1/3}$, so low temperatures, low density of states $D(E)$, and long localization lengths are required for one-dimensional behavior. In silicon two-dimensional systems, the density of states in the ground-state subband is $D(E) \sim 1.6 \times 10^{14}$ $eV^{-1}cm^{-2}$ and is constant for a wide range of Fermi energies [13]. At very small values of the chemical potential, it is now well established that the density of states decreases exponentially with decreasing gate voltage. If $\xi \sim 30$ nm and $T \sim 0.1$ K, R_0 is about 50 nm.

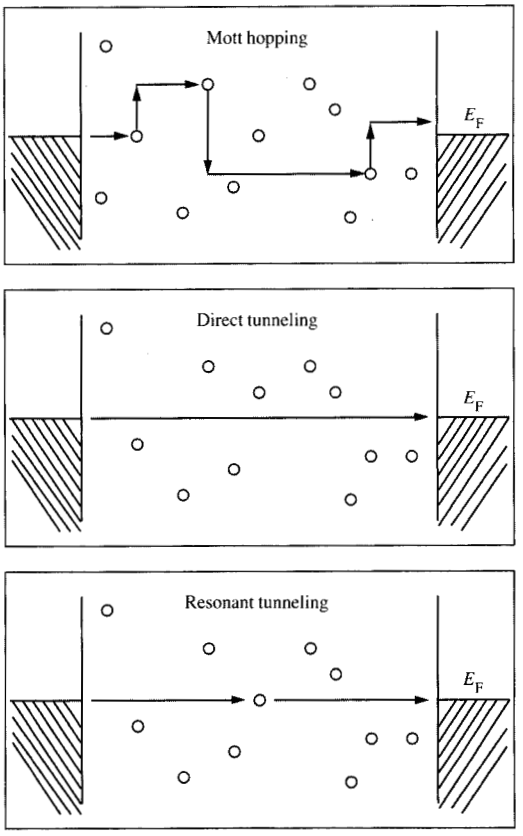

Figure 1

The three major conduction mechanisms at low temperature are shown schematically. In hopping the electrons proceed stepwise down the chain. To tunnel from site to site they must pick up thermal energy, so that the process is activated. For direct tunneling the electrons make a transition directly from filled states in the cathode to empty states in the anode. In resonant tunneling the electrons may be considered to tunnel from the cathode into the resonant state, resonate in the state, and then tunnel out to the anode. For a MOSFET device, changing the gate voltage moves the band edge and the localized states relative to the chemical potential of the contacts.

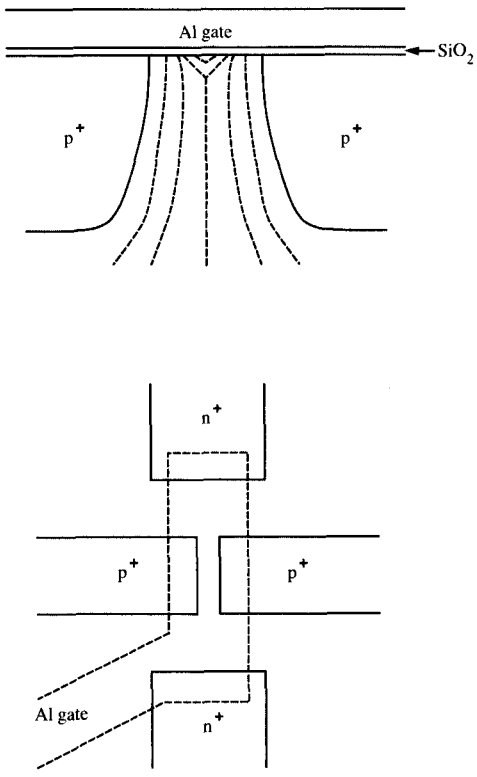

Figure 2

A schematic view of pinched samples. The upper figure is a section across the sample. The p+ diffusions create a well parallel to the surface. The surface potential resulting from the gate field continues the electrons to the surface. The lower figure is a plan of the structure showing the location of the n+ contacts, the aluminum gates, and the p+ diffusions.

The pinched sample configuration that we have used [1–6] is shown schematically in **Figure 2.** In this type of structure, the electric field from the p^+ diffusions confines the electron accumulation layer into a channel much narrower than the width between the diffusions. Although the geometrical spacing between p^+ diffusions is 1 μm, we estimate the conduction channel width to be of the order of 20–50 nm at low gate voltages [1].

Figure 3 shows the main features of the conduction at $T =$ 1.56, 0.416, and 0.107 K. The conductance, G, increases exponentially as either the temperature, T, or the gate voltage, V_G, is increased (V_G is roughly proportional to the chemical potential μ). Conductance data over a much narrower range of gate voltages at 0.05 K are displayed in **Figure 4.** What appeared to be noise at low T and low V_G in Figure 3 is actually reproducible but random fluctuations in the conductance as a function of V_G. Here the fluctuations, ΔG, are dramatic, and $\Delta G/G$ are much larger than those found in the metallic samples, where they are generally of the order of 10^{-3}–10^{-4} of the conductance. The peaks in the conductance curves persist to temperatures above 0.5 K at approximately the same V_G, albeit broadened and overlapped with other peaks. Many of the peaks are reproducible over a period of weeks or even months as long as the device is kept below ~10 K and not shocked electrically so as to change the oxide charge. Other devices, fabricated from the same Si wafer and nominally identical, exhibit the same behavior except for differences in the structural detail of the fluctuations.

The reason that each device is different is obvious if one considers the number of states in the entire sample within k_BT of the chemical potential, μ. If we use the 2D density of states, 10^{14} $cm^{-1}eV^{-1}$, and a temperature of 100 mK so that k_BT is 8×10^{-6} eV, the total number of states within k_BT of μ in these devices of area $\sim 5 \times 10^{-9}$ cm^{-2} is 8. These would be randomly arranged along the conductance channel in energy and position. If those within $\pm 5k_BT$ take part in conduction, there are still only of the order of at most 100

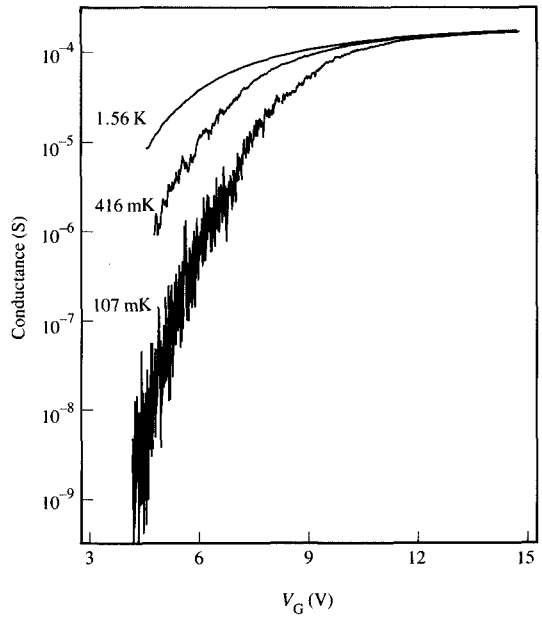

Figure 3

The gate voltage dependence of the conductance is shown at three temperatures. It can be seen that the conductance increases rapidly as a function of gate voltage and also as a function of temperature, especially at low gate voltages. It is also apparent that the conductance fluctuations increase rapidly as temperature decreases.

sites or one on the average of every 100 nm. Thus, it is not surprising that this would look like a 1D chain and that conduction should depend strongly on the sample and on μ. At the highest temperatures shown, the tails of many peaks overlap and contribute to the conductance at any fixed V_G, and the conductance begins to take on the character of a statistical average.

In the case of very large samples in two and three dimensions, it has been well established that when Mott variable-range hopping (VRH) is the dominant transport mechanism, the temperature dependence of the conductance, G, takes the form $\ln G \sim -(T/T_{0d})^{\langle 1/(d+1)\rangle}$, where d is the dimension of the conduction path and T_{0d} is a characteristic temperature that depends upon ξ, $D(E)$, and dimensionality [13]. Surprisingly, the same temperature dependence is observed in our small device, provided that the temperature dependence of the conductance is measured at a local minimum in the conductance structure. We find experimentally [1–3] that at temperatures $0.2\ \text{K} < T < 3\ \text{K}$ and gate voltages below ~6 V, the conductance can best be described by $\ln G \sim -(T/T_{01})^{1/2}$, or 1D VRH. At gate voltages above 6 V, our conductance data are better fit by a smaller exponent, perhaps $\ln G \sim -(T/T_{02})^{1/3}$, or 2D VRH. The reason for this transition from 1D to 2D behavior is that increasing V_G increases $D(E)$, hence decreasing the most probable hopping distance. In addition, increasing the gate voltage widens the conduction channel. Recent computer simulations by Xie and Das Sarma [16] indicate that for our device there will be a continuous transition from one- to two-dimensional VRH as the channel width increases relative to the most probable hopping length. Indeed, careful analysis of the data near $V_G = 5.9$ V indicate that at low T the best-fit exponent is 1/2 and at higher T it is 1/3, consistent with a crossover from 1D to 2D behavior as the most probable hopping distance decreases with increasing T. We point out, however, that the temperature range over which we fit the two-dimensional behavior is only one and a half decades; thus, the 1D VRH to 2D VRH transition may simply represent the transition from strongly localized to weakly localized one-dimensional transport. In fact, at gate voltages above ~10 V, we clearly observe weakly localized behavior because the change in the conductance, ΔG, takes the form $\Delta G = -CT^{1/2}$ for all temperatures above 0.1 K. This type of power-law temperature dependence for the conductance is frequently found in 1D weakly localized systems.

At temperatures above ~3 K and gate voltages below 9 V, we find that the conductance mechanism can be best described by thermal activation (simply activated behavior) or $G \sim \exp(E_0/k_B T)$, where E_0 is an activation energy that depends slightly on gate voltage and represents an energy that is characteristic of the potential disorder in the sample. As we have already pointed out, below 3 K the low-gate-voltage strongly localized regime shows the temperature

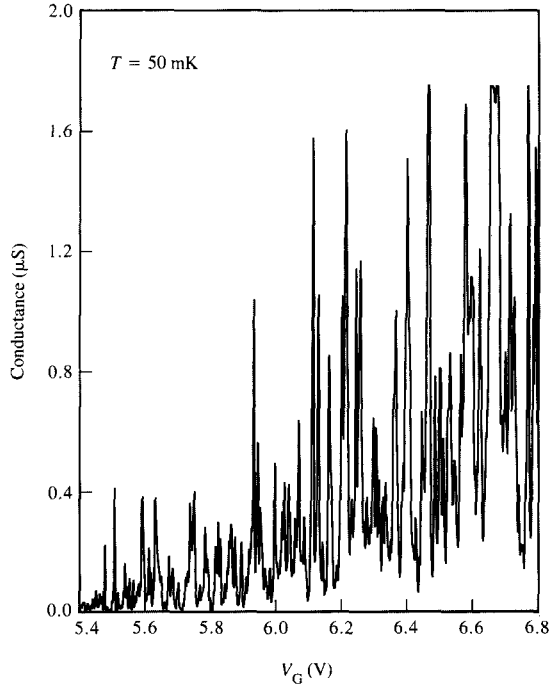

Figure 4

Conductance as a function of gate voltage at 50 mK. Here the peaks vary by two orders of magnitude or more.

dependence that a simple Mott-type argument for $R_0 > w$ would suggest. However, Kurkijarvi [17] argued that in any 1D hopping sample the conductance should be simply activated at all T, because one of the hops in the chain will always be exponentially less likely than others and this single hop will limit the total conductance of the entire sample. There have been arguments [18, 19] as to whether such "blockades" will exist in short samples, but for all our data taken at a minimum in the conductance structure below $V_G \sim 6$ V at $T < 3$ K, we found what looked like statistically averaged results with a $\ln G \sim T^{-1/2}$.

To explore the transport mechanisms further [5, 6], more detailed measurements were made on the behavior of the conductance peaks at temperatures as low as 20 mK. The T dependence of a few peaks contained in a very limited gate voltage range is shown in **Figure 5.** At 36 mK a single peak at about $V_G = 4.5175$ V lies above the noise level. (Note that these experiments were made with voltages across the contacts of order 2 μV, so the noise signal was at about 10^{-15} A.) As the temperature is increased, the amplitude of this central peak grows, other peaks appear and grow faster, the widths of all the peaks broaden, and the background

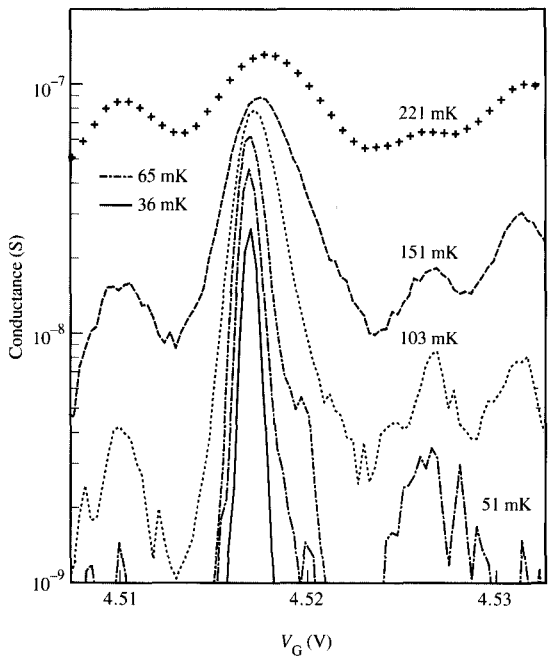

Figure 5

Conductance as a function of gate voltage for T = 36, 51, 65,103,151, and 221 mK. At 36 mK there is only one peak. The peaks lost in the noise grow more rapidly than this peak, and all peaks broaden. As a result, log G is fairly smooth at 221 mK.

conductance (the minimum of the peaks) increases faster than the peak conductance. **Figure 6** shows the temperature dependence of the peak conductance for a few peaks that were well isolated from others at low temperatures. All appear simply activated [i.e., $G \sim \exp(-E_0/k_B T)$] for $T < 0.2$ K, with the smallest peaks having the greatest activation energies. Above about 0.2 K the data for all four peaks fall approximately on a single curve of the form $\ln G \sim T^{-1/2}$, a dependence typical for a statistically averaged 1D variable-range hopping system. At temperatures between 0.05 K and 0.1 K, the peak at 4.5175 V seems to show a weaker temperature dependence than expected from simple activation with the slope, $d\ln G/dV_G$, decreasing as the temperature approaches the activation energy (~0.083 K for this peak). This decrease is actually due to a change in the occupation probability, $f(E)$, of electrons in the localized state, and is given by the well-known Fermi–Dirac distribution function, $f(E) = (e^{(E-\mu)/k_B T} + 1)^{-1}$. A similar effect should occur for the other peaks, but their activation energies are higher and the transition is not as evident in Figure 6.

To understand the origin of the conductance fluctuations in the 1D strongly localized limit, it is useful to model the transport process as being similar to conduction in a 1D random resistor network [20–22]. In this case, the conductance is limited by the largest resistor (the critical hop) in the chain, and the conductance associated with hopping from localized state i to localized state j is given by

$$G = G_{ij} \exp\left\{-2x_{ij}\xi^{-1} - \frac{1}{2k_B T}[|E_i - \mu| + |E_j - \mu| + |E_i - E_j|]\right\}, \quad (1)$$

where ξ is the localization length, x_{ij} is the spatial separation of two states, E is a site energy, and μ is the local chemical potential. If this hop is the lowest conductance hop in the chain, the conductance of the entire sample is also given by Equation (1). If the chemical potential is adjusted to any energy at a constant temperature such that $E_i < \mu < E_j$ (assuming $E_i < E_j$), the conductance G is a constant independent of μ (a flat-topped peak). If $\mu > E_j$ or $\mu < E_i$, the conductance on the side of the peak decreases exponentially with changing μ and is given by $|d\ln G/d\mu| \sim (k_B T)^{-1}$.

This problem of hopping conductivity in a small random network has been extensively studied using numerical simulation techniques by P. A. Lee et al. [23] and others [24–27]. Under the assumption that the localized states have

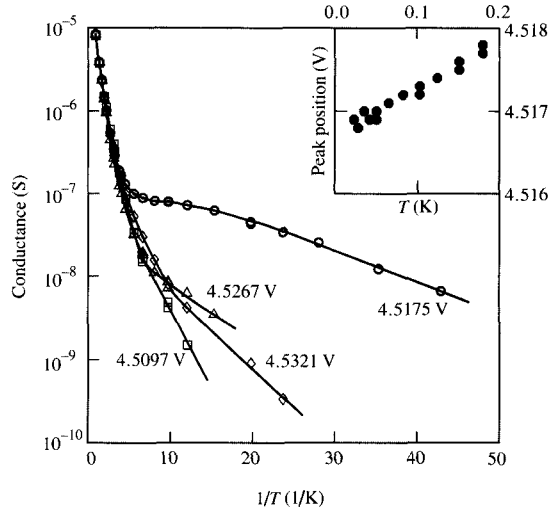

Figure 6

Temperature dependence of some of the peaks shown in Figure 5; the highest peak has the lowest activation energy. Above approximately 200 mK, all peaks follow a "$T^{-1/2}$" law. The inset shows the dependence of the position of one peak on temperature.

a random distribution in energy and position, the conductance as a function of chemical potential was found to be quite similar to our results displayed in Figure 4. It turns out that there are only two important types of hopping conductance peaks. The first is the one described by Equation (1) and the second is the case when two of the random resistors in the network limit the conductance. In this second case, the conductance peaks have a sharp top and the peak position can move to either higher or lower chemical potentials with increasing temperature. In the case of a single resistor limiting the conductance, the peak position is independent of temperature. If there are two independent critical hops limiting the conductance such that $i \rightarrow j$ and $k \rightarrow l$, where $E_i < E_j < E_k < E_l$ (one of several possible localized-state energy distributions), the maximum conductance of the peak is given by $G_p \sim \exp[-(E_l - E_i)/k_B T]$ and is observed at an energy $\mu_p = (E_i + E_l)/2 + k_B T(x_{kl} - x_{ij})\xi^{-1} + (k_B T/2)\ln(g_{ij}/g_{kl})$. The peak-position shift can be to higher or lower gate voltages depending on the relative values at x_{kl} and x_{ij}. The inset to Figure 6 shows one example of the temperature dependence of the peak position observed in our experiments. For all our data on well-isolated peaks, we find that on average 33 percent of the peaks shift toward higher μ, 33 percent toward lower μ, and 33 percent do not change as the temperature is increased. On either side of the peak the conductance is again described by $|\, d\ln G/d\mu \,| \sim (k_B T)^{-1}$. The numerical simulations find a mix of sharp- and flat-topped peaks randomly distributed in μ, with each peak increasing exponentially in amplitude and becoming wider as the temperature increases. This agrees quite well with our experimental data (see Figure 3), with the exception that we never have observed flat-topped peaks. The lack of observation of flat-topped peaks suggests that all our conductance fluctuations are from hopping processes that involve more than one critical hop, or that the theory is not yet well enough developed to accurately describe the physics in our samples.

It is interesting to note that Equation (1) is *not* symmetric with respect to the direction of the applied voltage once the temperature dependence of the electron occupation probability is included [23]. This means that the current–voltage curves associated with a single hop are nonlinear and will cause rectification of an ac signal. This rectification in turn should generate large second-harmonic signals in ac measurements, as is indeed the case. Asymmetric I–V curves together with strong second-harmonic generation have been observed in these small structures [4, 8].

In principle the density of states appropriate for each peak can be derived from our measurements of the logarithmic slopes of the sides of the peaks. Under the assumption that $d\ln G/d\mu = (k_B T)^{-1}$, the density of states is given by [13] $D(E) = dn/d\mu = (C/e)(k_B T)^{-1}(d\ln G/dV_G)^{-1}$, where n is the 2D electron density and C is the gate capacitance of our device (115 pF/cm^2). In practice, we find that our measured

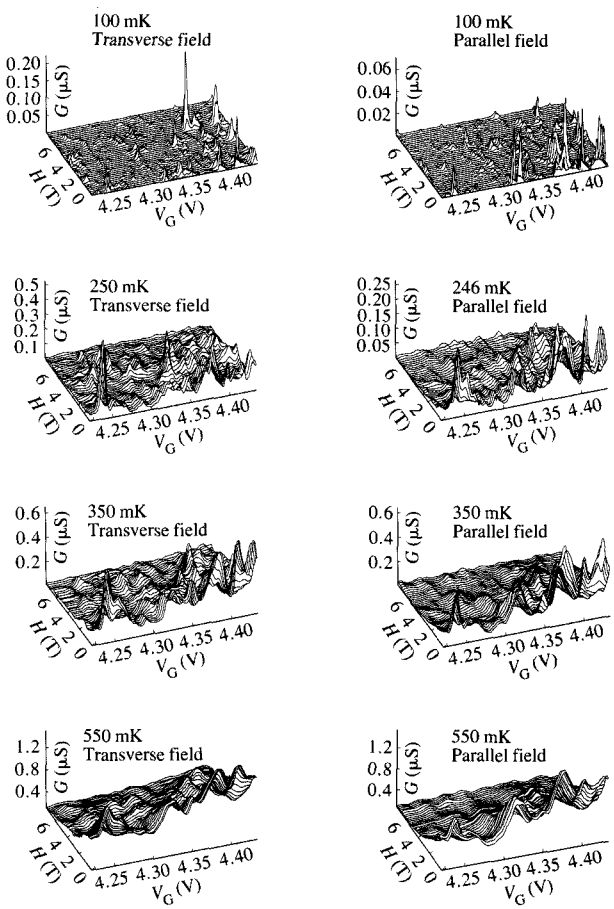

Figure 7

A collection of data showing the dependence of conductance on gate voltage for different temperatures, magnetic fields, and direction of B. Figures on the left are for B perpendicular to the sample surface; those on the right are for B parallel to the current direction.

slopes vary by a factor of three from peak to peak. The derived density of states is about 10^{-14} eV^{-1}cm^{-2}. We are unable to determine whether the factor of three variation in our computed $D(E)$ represents an actual fluctuation in $D(E)$ as V_G is varied or if it is associated with some other contribution to the physics of the individual peaks.

Recently we have made measurements [6] of the magnetic field dependence of these conductance fluctuations. Some of the striking results are shown in **Figure 7** for various magnetic fields, temperatures, and sample orientations relative to the direction of the applied magnetic field. It may be seen that for field changes of the order of a few tenths of a tesla there are strong variations in the peak amplitudes. Some peaks increase while others decrease in a seemingly random fashion. Within a few tesla, the structure is totally different from what is seen at zero field. The qualitative

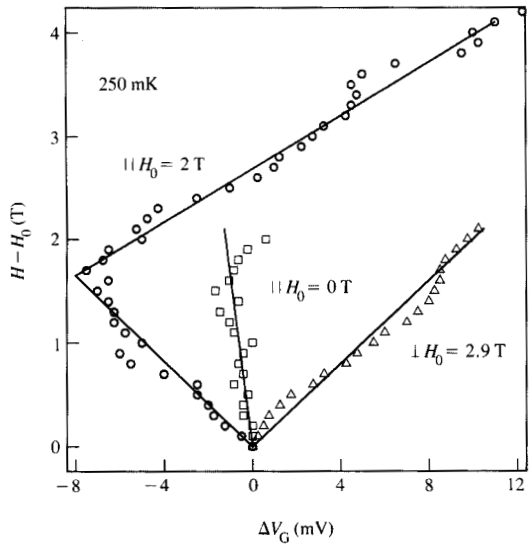

Figure 8

Magnetic field dependence of the location in gate voltage of the maximum in conductance for three separate and well-isolated peaks.

features of the magnetic field dependence of the conductance fluctuations are independent of the direction of the magnetic field, although in detail there are some differences. There are also, although it is not so apparent in Figure 7, shifts in peak positions with increasing magnetic field. Some peaks shift toward higher and some to lower values of V_G; others do not move.

One of the main reasons that the magnetic field has such a large effect on the behavior of the conductance is that it changes the energy of the localized state via the Zeeman effect, as has been studied extensively in 3D hopping systems by Kamimura [28]. The change in energy of an electron in a localized state is linearly proportional to the magnetic field, $E = E_0 \pm g\mu_B H/2$, where E_0 is the energy of the state before the application of the field, μ_B is the Bohr magneton, and g is the g factor of the electron. The energy of the state depends upon the direction of the electron spin relative to the direction of H. There are actually two separate spin bands (called Hubbard bands) and we expect that both of these will be partially occupied and contribute to our observed effects. Because of the Pauli exclusion principle, there are only four different types of single critical hops: 1) from a singly occupied localized site (either spin up or spin down) to an unoccupied site; 2) from a doubly occupied site to an unoccupied site; 3) from a singly occupied site to another singly occupied site; and 4) from a doubly occupied site to a singly occupied site. However, for large magnetic fields such that $g\mu_B H/2 > 5k_B T$, singly occupied antiparallel spin states will be suppressed [25, 28], and the hopping rate for processes 2 and 3 will be reduced. For transitions from a singly occupied site to an empty one, we would expect the energies to decrease in a magnetic field. For a hop from a doubly occupied site to a singly occupied one, they should increase. If the two hops giving rise to the peak are of the first type, the peak should move to lower V_G; if both are of the second type, the peak should move higher; if there is one of each, the peak should not move to first order. These three processes are shown for different peaks in **Figure 8.** The peak represented by the open circles first appears near 2 tesla and shifts linearly with increasing H toward lower V_G. Beyond 3.7 tesla the peak position begins to move linearly toward higher V_G until it disappears in our noise at $H \sim 6$ tesla. This change in direction of the peak-position shift is probably associated with a change of one of the critical resistors controlling the hop. Since the energies of the Zeeman shifts are known, the shift of position with gate voltage should be consistent with the density of states, as argued above for the temperature dependence of the peak positions in zero H. In fact, we find the slopes of the peaks that move with field, $dV_G(\text{peak})/dH$, to be consistent with a density of states one third of the value calculated from the capacitance and consistent with our earlier determination of $D(E)$. Zeeman shifts can also explain why the structure is totally changed within a few tesla. The Zeeman energies (~58 μeV/tesla) very quickly become comparable to the spacing between states (<4 μeV). The energies of doubly and singly occupied states get mixed in a different way by the field, and the random network is rearranged, changing the critical path and thus the structure we measure. Computer simulations by Kalia et al. [25] demonstrate the same results.

It should be noted that even though the results seem similar for the field perpendicular to the surface and parallel to the surface, there are differences. These can only arise from orbital (magnetic field effects on the phase of the electron wavefunction) rather than Zeeman effects. However, even with the mass of data collected in these experiments and with tedious calculations of various correlation functions, we are unable to make quantitative statements about the role that orbital effects play in our results. In other recent work in the strongly localized regime, Nguyen and coworkers [29] have observed fluctuations in the conductance that they argue can only be due to orbital effects. We believe that much more work needs to be done before the physics of hopping in small systems in the presence of a magnetic field can be completely understood.

Resonant tunneling

While the early experiments were being made and before they were understood, Azbel et al. [30] argued that the structure in the conductance as a function of gate voltage was the result of resonant tunneling through the localized

states near the center of the sample. This possibility had been proposed earlier by Lifschitz and Kirpichenkov [31]. The intriguing idea introduced by Azbel was that spectroscopy could be done using the gate to vary the energy of the localized states relative to the contacts. (This differs from other experimental approaches, which usually involve tunneling through a 2D oxide barrier [32] or through a quantum well structure [33].) However, detailed measurements of the long samples [3–5] showing simply activated behavior of the peaked structure made it apparent that a thermally activated tunneling mechanism was required. Stone and Lee [34] demonstrated that in fact the resonant tunneling current could not increase with temperature.

The fundamental idea of resonant tunneling is often discussed in elementary quantum-mechanics textbooks. If one has a double barrier with resonant states in the well between the barriers, then the transmission through the barriers is given by $T^L T^R$ when the energy of the incident electrons is not at resonance, and by T^L/T^R (if $T^L < T^R$) at resonance. Here T^L and T^R are the transmission coefficients of a wave through the isolated left and right barriers, respectively. If $T^L = T^R$, the total transmission, T, is unity.

For resonant tunneling through a disordered system, the leak rate of an electron from the localized state to left and right contacts is of the form $\Gamma^L = \Gamma_0 \exp -(2x/\xi)$ and $\Gamma^R = \Gamma_0 \exp[-2(L-x)/\xi]$, where L is the distance between contacts, x is the distance from the left contact to the localized state, and ξ is the localization length. It can be shown [35, 36] that for spinless electrons in any dimension,

$$G = \frac{e^2}{h} \frac{\Gamma^L \Gamma^R}{\delta E^2 + \left(\frac{\Gamma^L + \Gamma^R}{2}\right)^2}, \tag{2}$$

where $\delta E = E_0 - \mu$. G has a maximum of e^2/h for $\Gamma^L = \Gamma^R$, or $x = L/2$ and $\delta E = 0$. Equation (2) can be written as a kind of Lorentzian

$$G = \frac{e^2}{h} \frac{\Gamma_1^2}{\delta E^2 + \Gamma_2^2}.$$

At finite temperatures, if the Lorentzian is centered at E_0, the conductance must satisfy a Boltzmann equation, so that

$$G(E_0 - \mu) = \frac{e^2}{h} \int \frac{\Gamma_1^2}{(E_0 - \mu)^2 + \Gamma_2^2} \frac{df(E,\mu)}{dE}\, dE. \tag{3}$$

For $k_B T \ll \Gamma_2$, the derivative of the Fermi function $f(E, \mu)$ is a δ function, so that as $E_0 - \mu$ is varied (the gate voltage moves the resonant state relative to the source and drain Fermi surfaces),

$$G(E_0 - \mu) = \frac{e^2}{h} \frac{\Gamma_1^2}{(E_0 - \mu)^2 + \Gamma_2^2}. \tag{4}$$

Thus, at low temperatures the measurement of gate voltage dependence (proportional to μ) of the conductance of such a peak should be unchanging and represent the natural line shape of the resonance.

For $k_B T > \Gamma_2$ the Lorentzian looks like a δ function, and

$$G(E_0 - \mu) \sim \frac{df(E_0 - \mu)}{dE} = \frac{f(E_0 - \mu)[1 - f(E_0 - \mu)]}{k_B T}. \tag{5}$$

This has a maximum which decreases as $(k_B T)^{-1}$. Off the maximum, the slope of the peak $d \ln G / d(E_0 - \mu) = \pm (k_B T)^{-1}$. The line decreases and broadens with increasing temperature. As can be seen from Equation (2), the maximum conductance in the low-temperature regime is $\sim 4\Gamma^R \Gamma^L / (\Gamma^L + \Gamma^R)^2$, or

$$4 \exp(-2L/\xi) \left[\exp -2(x/\xi) + \exp -2\left(\frac{L-x}{\xi}\right) \right]^{-2}. \tag{6}$$

This has a maximum value of unity for $x = L/2$. However, the linewidth is proportional to the denominator and increases rapidly as x becomes significantly different from $L/2$. That is, the conductance maxima are strongest and narrowest for states near the center of the sample. By the time they are a localization length away, the magnitude of the peak is reduced by a factor of

$$\operatorname{sech}^2\left(\frac{L}{\xi} - 2\right).$$

The linewidth Γ_2 is increased by the same factor. Clearly, the effect of not being at the center is increased for large L: the shorter the sample, the better the chance of seeing a large peak. This had been suggested earlier by Azbel et al. [30], who found that the temperature at which resonant tunneling conductance became smaller than hopping conductance was proportional to L^{-2}. Clearly, shorter samples than the previous ones were required for tunneling studies, and a set of samples of the order of 0.5 μm long and 1 μm wide were chosen. These samples had heavily doped (metallic) polysilicon gates and an oxide thickness of 10 nm. (They were kindly supplied by M. R. Wordemann of the IBM Thomas J. Watson Research Center.) Measurement of the conductance as a function of gate voltage again showed sharp structure (see **Figure 9**). However, it was found that the temperature dependence was that expected from the arguments above. This is shown in **Figure 10,** where the peak values appear as a function of T^{-1} for peaks in two samples. Both show no T dependence at low temperature and a strong increase at high temperature. The largest peak, with a conductance equal to about 0.1 e^2/h, shows conductance that decreases in an intermediate region. The break occurs at about $k_B T = 4.3$ μeV (50 mK), which is an energy comparable to the width Γ_2. The high-temperature data, $T > 0.2$ K, fit a variable-range hopping law for two dimensions $\ln(G/G_0) = -(T/T_0)^{-1/3}$. What appears to be

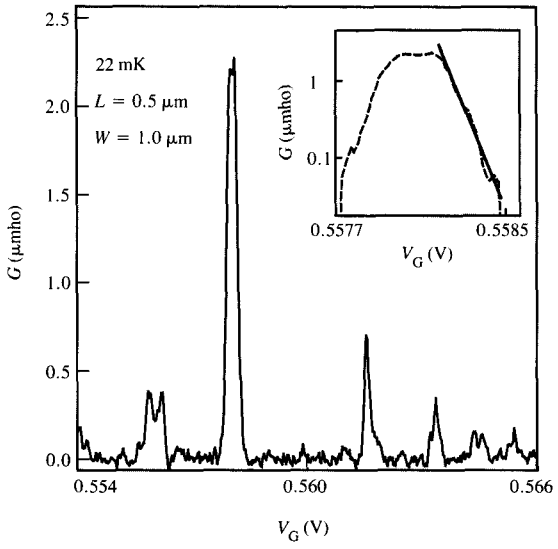

Figure 9

Structure in conductance as a function of V_G through short samples. The inset shows the largest peak as ln G vs. V_G. The solid line drawn along the high-V_G side was used to define d ln G/dV_G.

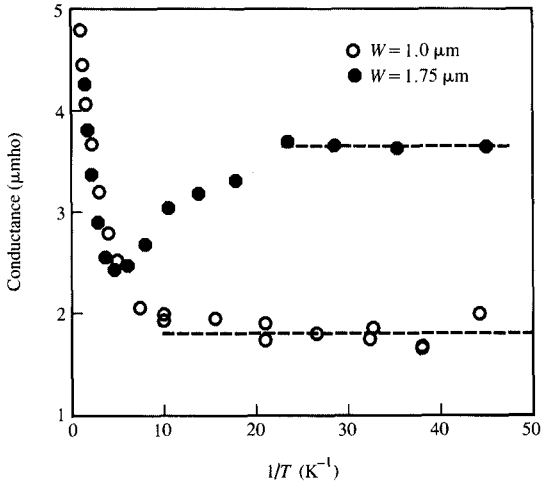

Figure 10

The temperature dependence of the maximum conductance of two tunneling peaks (G_{max} vs. T^{-1}). At high T both follow a "$T^{-1/3}$" law. At low T both are constant. The higher G peak decreases as T increases between 50 and 100 mK.

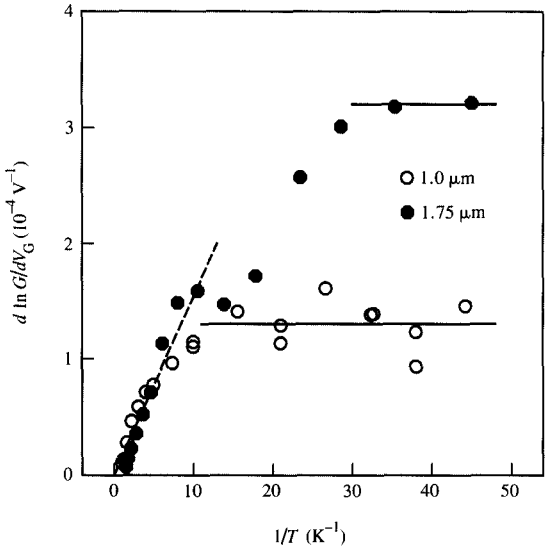

Figure 11

The temperature dependence of d ln G/dV_G for the same two peaks. Both are constant at low T and decrease roughly as T^{-1} at higher temperatures.

happening to the peak height for the high-conductance sample is that with increasing temperature the peak conductance remains constant, then decreases. However, the hopping increases exponentially and eventually dominates the conduction mechanism. The other peak is about a factor of two smaller. Its width Γ_2 is approximately a factor of two greater; therefore, it should not have decreased until $T \sim 100$ mK. By then the hopping begins to dominate the temperature dependence so that no minimum is seen.

More can be learned by looking at the slopes $d \ln G/dV_G$ for the two samples (see **Figure 11**), because these data are more readily separated from the background than the peak height. Here, although there is a difference in the low-temperature values, the expected $1/T$ temperature dependence can be seen at higher temperatures. The slopes of both peaks are relatively constant until some temperature T_1 is reached, then decreasing roughly as T^{-1}. For the higher-conductance peak, $T_1 \sim 55$ mK; for the other, it is $T_1 \sim 83$ mK. The ratio of these two temperatures is 0.66, and, according to the arguments given above, the peak heights should differ by the same ratio. The actual ratio of the peak heights is 0.5; thus, there is reasonable internal consistency of our data with the theory.

As in arguments made in the section on hopping, it is possible to deduce a density of states because $d^2 \ln G/dV_G dT$

is measured, dn/dV_G is known from capacitance, and $d^2 \ln G/d\mu dT$ is k_B^{-1}, so $D(E) = dn/d\mu$ can be found. It is roughly 1.8×10^{13} $eV^{-1}cm^{-2}$, or about 11 percent of the density of states of the unperturbed 2D silicon band. By using values of T_0 derived from the variable-range hopping regime, localization lengths of 42 nm for the low-conductance sample and 28 nm for the higher-conductance sample were found. One can use this density of states to estimate the width in energy, Γ_2, of the resonant state from the measured halfwidth of the peaks in gate voltage. For both peaks, the low-temperature halfwidth is roughly equal to $k_B T_1$, the energy at which the curves change to a temperature-dependent law. Within the errors of the measurement, we again find good agreement of the data with the theory.

It is interesting to consider the probability of finding a state within a bandwidth, Γ_2, of another state and also within ξ of the center of the sample. This probability is given by $2\Gamma_2 D(E) W\xi$, or about 0.05 for the parameters appropriate to our samples. Thus it would appear that the chance of finding two overlapping peaks is certainly less than 20 percent. All peaks, however, show reproducible fine structure, as shown in **Figure 12.**

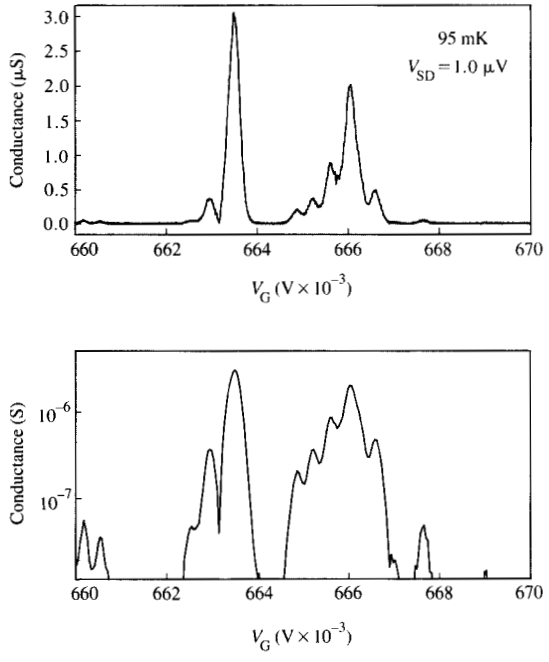

Figure 12

The gate-voltage dependence of several conductance peaks is shown plotted both as G and $\ln G$ vs. V_G. The multiple peak structure is quite apparent.

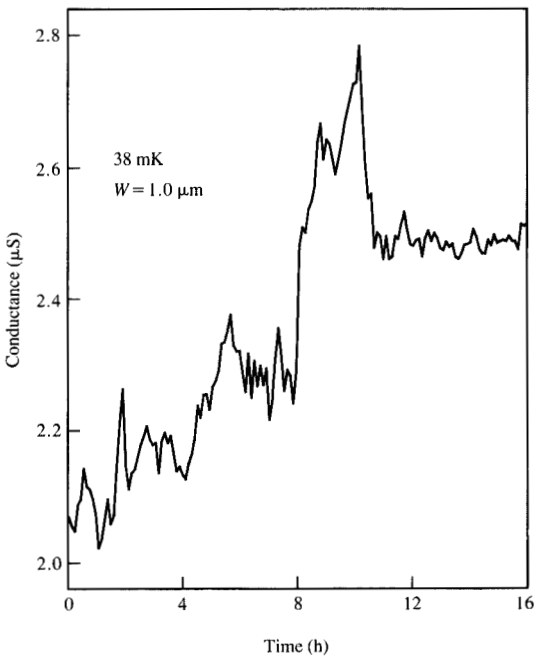

Figure 13

The conductance of a tunneling peak as a function of time.

A better understanding of the structure on each peak can be gained by studying the time dependence of the conductance. In our normal mode of data acquisition, we typically spent about two hours sweeping the gate voltage once over the range of a peak. Thus, all our measurements represent a time average. If a measurement is made at the peak maximum as a function of time, as shown in **Figure 13,** both high- and low-frequency fluctuations are evident which can be as large as 30 percent of the maximum value. These two observations may be explained by a phenomenon related to mechanisms proposed by Rogers and Buhrman [37]. At any moment there are other localized states in the sample which may or may not be occupied. The fraction of time they are occupied depends on the temperature and on their energies, *and* on the Coulomb interactions between them. In samples of the size we studied, there are roughly four other states within $\pm 5k_B T$ of the resonant state at 50 mK. Of the 32 possible configuration of electrons in these states, only a few are very likely, but each will result in the resonant state being shifted to slightly different energies by the Coulomb interactions. The Coulomb interactions [38], which are screened by the metal gate electrode, are of the

order of the bandwidth of the resonant state if these extra states are located about 0.2 μm from the resonant state. Thus, in the time-averaged experiments the structure may simply be a time average over the different configurations. The time dependence may result from jumps from one configuration to another. It clearly would be interesting to study the frequency spectra as a function of temperature. It should be related to the kinetics of the processes leading to the change of configurations. If the Coulomb interactions were increased by greatly increasing the oxide thickness, the nature of the peak structure should change markedly.

We conclude this section by noting that we have observed conductance fluctuations which seem to have the temperature dependence and other properties one would expect of resonant tunneling in such a system. Constants derived from the experiments are consistent with theoretical expectations. Finally, there is evidence of the effects of Coulomb interactions giving rise to structure in the tunneling.

Conclusions

We have discussed observations and results that a few years ago might well have been ignored by experimenters who did not recognize or understand the nature of the conductance fluctuations. Fluctuations were observed as early as 1965 in the transconductance of MOSFET devices, and were either ignored or attributed incorrectly to a host of causes. Even in fairly large samples the device-specific effects of the failure to ensemble average can be seen in the derivative of the conductance. In small hopping samples, the recognition that the conductance was sample-specific was unavoidable.

To advance from the present situation of having identified the relevant mechanisms to that of learning something new about the kinetic processes has been our most recent aim. In the case of magnetoresistance in hopping, this has proved difficult, because the Zeeman effects tend to mask other effects that seem more interesting to us. At present one has the feeling that the models that have been used to calculate the transport properties of large ensembles have been experimentally justified in small samples, both for tunneling and hopping. Furthermore, in resonant tunneling experiments much work remains to be done to understand the kinetics of the Coulomb interaction results, the I–V characteristic, and magnetic effects. Work at currents of 10^{-13}–10^{-15} A and at source–drain voltages of 1 μV and less at very low temperatures can be time-consuming, so it may be some time before these subjects are well understood. However, a new and powerful tunneling spectroscopy has been demonstrated for the first time. It has not been applied to direct tunneling nor to the dependence of direct tunneling on the position of the conduction band relative to the contact Fermi energies, because significantly shorter samples would be needed. With the recent advances in fabrication techniques such samples can now be created.

Acknowledgment

As papers have come out, we have acknowledged our indebtedness to colleagues; however, we wish to acknowledge again our continued interactions with Patrick Lee.

References

1. A. B. Fowler, A. Hartstein, and R. A. Webb, "Conductance in Restricted-Dimensionality Accumulation Layers," *Phys. Rev. Lett.* **48,** 196–199 (1982).
2. A. B. Fowler, A. Hartstein, and R. A. Webb, "Transition from 1-Dimensional to 2-Dimensional Hopping Conductivity in Silicon Accumulation Layers," *Physica* **117 & 118B,** 661–666 (1983).
3. A. Hartstein, R. A. Webb, A. B. Fowler, and J. J. Wainer, "One-Dimensional Conductance in Silicon MOSFET's," *Surf. Sci.* **142,** 1–13 (1984).
4. R. A. Webb, A. Hartstein, J. J. Wainer, and A. B. Fowler, "Origin of the Peaked Structure in the Conductance of One-Dimensional Silicon Accumulation Layers," *Phys. Rev. Lett.* **54,** 1577–1580 (1985).
5. R. A. Webb, A. B. Fowler, A. Hartstein, and J. J. Wainer, "Hopping Conduction in Quasi-One-Dimensional Systems," *Surf. Sci.* **170,** 14–27 (1986).
6. J. J. Wainer, A. B. Fowler, and R. A. Webb, "Magnetic Field Effects in Strongly Localized Quasi-1D MOSFETs," *Surf. Sci.* **196,** 134–138 (1988).
7. A. B. Fowler, G. L. Timp, J. J. Wainer, and R. A. Webb, "Observation of Resonant Tunneling in Silicon Inversion Layers," *Phys. Rev. Lett.* **57,** 138–141 (1986).
8. R. F. Kwasnick, M. A. Kastner, J. Melngailis, and P. A. Lee, "Nonmonotonic Variations of the Conductance with Electron Density in ~70-mm-Wide Inversion Layers," *Phys. Rev. Lett.* **52,** 224–228 (1984).
9. M. A. Kastner, R. F. Kwasnick, J. C. Licini, and D. J. Bishop, "Conductance Fluctuations Near the 2D Extended Transition in Narrow Si-Metal-Oxide Semiconductor Field Effect Transistors," *Phys. Rev. B* **36,** 8015–8031 (1987).
10. F. F. Fang and A. B. Fowler, "Transport Properties of Electrons in Inverted Silicon Surfaces," *Phys. Rev.* **169,** 619–631 (1968).
11. M. Pepper, "The Anderson Transition Inversion Layers: The Origin of the Random Field and the Effect of Substrate Bias," *Proc. Roy. Soc. Lond. A* **353,** 225–246 (1977).
12. C. J. Adkins, "Effective Medium Theory of Conductivity and Hall Effect in Two Dimensions," *J. Phys. C: Solid State Phys.* **11,** 3389–3393 (1978).
13. T. Ando, A. B. Fowler, and F. Stern, "Electronic Properties of Two-Dimensional Systems," *Rev. Mod. Phys.* **54,** 437 (1982).
14. B. I. Shklovskii and A. L. Efros, *Electronic Properties of Doped Semiconductors*, Springer-Verlag, Berlin, 1984.
15. S. Pollitt, "Further Investigations of the Spatial Extent of Localized-State Wave Functions," *Commun. Phys.* **1,** 207–212 (1976).
16. X. C. Xie and S. Das Sarma, "Transition from 1 to 2 Dimensional Fluctuation-Variable Range-Hopping Conduction in Microstructures," *Phys. Rev. B* **36,** 4566–4569 (1987).
17. J. Kurkijarvi, "Hopping Conductivity in One Dimension," *Phys. Rev. B* **8,** 922–924 (1973).
18. V. K. S. Shante, C. M. Varma, and A. N. Bloch, "Hopping Conduction in Quasi-One-Dimensional Disordered Compounds," *Phys. Rev. B* **8,** 4885–4889 (1973); V. K. S. Shante, *Phys. Rev. B* **16,** 2597–2612 (1977).
19. W. Brenig, G. H. Dohler, and H. Heyszenau, "Hopping Conductivity in Highly Anisotropic Systems," *Phil. Mag.* **27,** 1093–1103 (1973).
20. A. Miller and E. Abrahams, "Impurity Conduction at Low Concentration," *Phys. Rev.* **120,** 745–755 (1960).
21. V. Ambegaokar, B. I. Halperin, and J. S. Langer, "Hopping Conductivity in Disordered Systems," *Phys. Rev. B* **4,** 2612–2620 (1971).
22. M. Pollack, "A Percolation Treatment of dc Hopping Conduction," *J. Noncryst. Solids* **11,** 1–24 (1972).

23. P. A. Lee, "Variable Range Hopping in Finite One-Dimensional Wires," *Phys. Rev. Lett.* **53,** 2042–2045 (1984).
24. R. A. Serota, R. K. Kalia, and P. A. Lee, "New Aspects of Variable-Range Hopping in Finite One-Dimensional Wires," *Phys. Rev. B* **32,** 8441–8446 (1986).
25. R. K. Kalia, W. Xue, and P. A. Lee, "Hopping Magnetoconduction and the Random Structure in Quasi One-Dimensional Inversion Layers," *Phys. Rev. Lett.* **57,** 1615–1618 (1986).
26. J. A. McInnes and P. N. Butcher, "Numerical Calculations of Variable-Range Hopping in One-Dimensional MOSFETs," *J. Phys. C: Solid State Phys.* **18,** L921–L925 (1985).
27. V. L. Nguyen, B. V. Spivak, and B. I. Shklovskii, "Hopping-Conductivity Fluctuations in Small Samples," *JETP Lett.* **43,** 44–47 (1986).
28. H. Kamimura, "Theoretical Model on the Interplay of Disorder and Electron Correlations," *Suppl. Prog. Theor. Phys.* **72,** 206–231 (1982).
29. V. L. Nguyen, B. V. Spivak, and B. I. Shklovskii, "Aharonov–Bohm Oscillations with Normal and Superconducting Flux Quanta in Hopping Conductivity," *JETP Lett.* **41,** 42–45 (1985).
30. M. Y. Azbel and P. Soven, "Transmission Resonances and the Localization Length in One-Dimensional Disordered Systems," *Phys. Rev. B* **27,** 831–835 (1983); M. Y. Azbel, A. Hartstein, and D. P. DiVincenzo, "T Dependence of the Conductance in Quasi One-Dimensional Systems," *Phys. Rev. Lett.* **52,** 1641–1644 (1984); M. Y. Azbel and D. P. DiVincenzo, "Finite-Temperature Conductance in One Dimension," *Phys. Rev. B* **30,** 6877–6888 (1984); M. Y. Azbel, A. Hartstein, and D. P. DiVincenzo, "Conductance of One-Dimensional Random Systems," *Physica* **127B,** 252–256 (1984).
31. I. M. Lifschitz and V. Y. Kirpichenkov, "Tunnel Transparency of Disordered Systems," *Sov. Phys. JETP* **50,** 499–511 (1979).
32. S. J. Bending and M. R. Beasley, "Transport Processes via Localized States in Thin α-Si Tunnel Barriers," *Phys. Rev. Lett.* **55,** 324 (1985).
33. For a review, see *Synthetic Modulated Structures*, L. L. Chang and B. C. Giessen, Eds., Academic Press, Inc., New York, 1985.
34. D. A. Stone and P. A. Lee, "Effect of Inelastic Processes on Resonant Tunneling in One Dimension," *Phys. Rev. Lett.* **54,** 1196–1199 (1985).
35. V. Kalmeyer and R. B. Laughlin, "Differential Conductance in Three-Dimensional Resonant Tunneling," *Phys. Rev. B* **35,** 9805–9808 (1987).
36. W. Xue and P. A. Lee, preprint; available from P. A. Lee, Massachusetts Institute of Technology, Cambridge, MA 02139.
37. C. T. Rogers and R. A. Buhrman, "Nature of Single-Localized-Electron States Derived from Tunneling Measurements," *Phys. Rev. Lett.* **55,** 859–862 (1985).
38. O. Entin-Wohlman and Z. Ovadyahu, "Modifications of Hopping Transport Due to Electrostatically Enhanced Coulomb Repulsion," *Phys. Rev. Lett.* **56,** 643–646 (1986).
39. W. E. Howard and F. F. Fang, "Low Temperature Effects in Si FET's," *Solid-State Electron.* **8,** 82–83 (1965).

Received November 13, 1987; accepted for publication December 15, 1987

Alan B. Fowler *IBM Thomas J. Watson Research Center, P.O. Box 218, Yorktown Heights, New York 10598.* Dr. Fowler received B.S. and M.S. degrees from Rensselaer Polytechnic Institute and a Ph.D. from Harvard University. He has been with the IBM Research Division since 1958 and is the manager of the Low Temperature Group in the Semiconductor Department and an IBM Fellow. Dr. Fowler is a co-recipient of the John Price Wetherill Medal of the Franklin Institute, the David Sarnoff Medal of the Institute of Electrical and Electronics Engineers, and the Buckley Prize of the American Physical Society. He is an Alexander von Humboldt Preistrager, a Fellow of the American Physical Society and the Institute of Electrical and Electronics Engineers, and a member of American Association for the Advancement of Science and the Research Society of America.

Jonathan J. Wainer *IBM Thomas J. Watson Research Center, P.O. Box 218, Yorktown Heights, New York 10598.* Mr. Wainer received his B.S. in physics from the University of Massachusetts at Amherst in 1978 and his M.S. in physics from New York University in 1982. He worked at Bell Telephone Laboratories in Murray Hill, New Jersey, from 1978–1979 before joining IBM's Research Division at Yorktown Heights, where he is currently a senior associate engineer in the Low Temperature Quantum Phenomena Group in the Physical Sciences Department. Mr. Wainer's research interests are in the nonlinear transport properties of insulators.

Richard A. Webb *IBM Research Division, Yorktown Heights, P.O. Box 218, Yorktown Heights, New York 10598.* Dr. Webb received his B.A. from the University of California at Berkeley in 1968, his M.S. from the University of California at San Diego in 1970, and his Ph.D. from the University of California at San Diego in 1973, all in physics. He continued in a postdoctoral position at the University of California at San Diego until 1975, when he joined the research staff at Argonne National Laboratory, Argonne, Illinois. Dr. Webb joined IBM in 1978 and is currently a Research Staff Member and manager of the Low Temperature Quantum Phenomena Group at Yorktown Heights, New York. His main experimental interests are in the area of novel physics in very small structures at low temperatures. Dr. Webb has received three Outstanding Technical Achievement Awards for his work on macroscopic quantum tunneling at low temperatures in small Josephson junctions, variable-range hopping conduction in very small Si MOSFETs at low temperatures, and Aharonov–Bohm effects in normal metals. He is a Fellow of the American Physical Society.