Conductance fluctuations near the localized-to-extended transition in narrow Si metal-oxide-semiconductor field-effect transistors

M. A. Kastner, R. F. Kwasnick,* and J. C. Licini
Department of Physics and Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

D. J. Bishop
AT&T Bell Laboratories, Murray Hill, New Jersey 07974-2070
(Received 2 February 1987)

A unified theoretical picture has emerged to explain the fluctuations in the conductance of disordered, one-dimensional resistors. At zero temperature T, and if the localization length ξ is shorter than the sample length L, current is carried by resonant tunneling which gives rise to exponentially large conductance fluctuations as the Fermi energy is varied. On the other hand, if $\xi > L$, the fluctuations are always of size e^2/h. As T is raised, the fluctuations at first remain exponentially large for $L > \xi$, but are the result of one-dimensional phonon-assisted hopping. At still higher T, the inelastic diffusion length L_{in} plays the role of sample length, and when L_{in} becomes as short as the localization length the states become delocalized. For temperatures above this localized-to-extended transition, the fluctuations of conductance are always of order e^2/h in a sample of length L_{in}. We present the results of experiments on narrow (~ 70 nm) inversion layers in Si metal-oxide-semiconductor field-effect transistors. At low T there are exponentially large fluctuations at low conductance and e^2/h fluctuations at high conductance. It is shown that the dependence of current on Fermi energy, temperature, and source-drain voltage for low conductance can be explained by one-dimensional phonon-assisted hopping. The samples are too long to observe resonant tunneling. It is pointed out, however, that several predicted features are not observed in the limited temperature range studied. At high conductance the fluctuations are of size e^2/h in a sample of length L_{in}. An estimate is made of the value of the conductance at which the localized-to-extended transition occurs: That is, the value at which the conductance and its fluctuations are both $\sim e^2/h$ in a sample of length L_{in}.

I. INTRODUCTION

In 1977 Thouless[1] began a revolution in the way we think about conduction in metals. He used a scaling argument to show that all wires, no matter what their cross-sectional area, would become insulators at $T=0$ if they were made long enough to have resistance greater than h/e^2, about 30 000 Ω. Thouless's predictions were confirmed by experiments on narrow wires (for reviews see Refs. 2 and 3). His scaling ideas led to spectacular theoretical progress[4,5] which has recently been reviewed.[6] This, in turn, stimulated experiments on two-dimensional (2D) conductors,[7] showing that in 2D, as in one, all states are localized at zero temperature.

Central to the scaling theory is the idea that there is only one parameter in the problem: the conductance G itself. It was, therefore, devices in which G could be varied in the midst of an experiment that provided the clearest test of the theory. The experiments[8] on metal-oxide-semiconductor field-effect transistors (MOSFET's) showed the logarithmic temperature dependence predicted for 2D and its dependence on resistance.

It was clear to several groups that much could be learned by using MOSFET's to study conduction in 1D as well as 2D. Conventional MOSFET's have subbands arising from the quantization of electronic motion normal to the Si surface. Each of these 2D bands has a constant density of states. However, if the motion in a MOSFET were quantized for two directions but still free-electron-like in the third, there would be square-root singularities in the density of electronic states. In addition to such ideal behavior, the effects of disorder could be explored and the predictions of the scaling theories tested. However, as is the case for recent work on small metal wires,[9] the most exciting results to come out of experiments on narrow MOSFET's were completely unexpected.

A variety of devices have been fabricated to probe 1D behavior in MOSFET's. The first was reported by Fowler, Hartstein, and Webb.[10] It consists of a lightly doped n-type Si substrate, n^+ source and drain contacts, and an Al gate over the thermal oxide. What would be a standard accumulation-mode device is modified by implantation of p^+ regions on each side of the channel. The depletion regions between these p^+ contacts and the substrate are then used to constrict the conducting channel. Fowler, Hartstein, and Webb[10] and Pepper and Uren[11] reported that at low T such devices displayed exponentially large fluctuations as a function of gate voltage, V_G, that is, the Fermi energy of the electron gas.

Almost simultaneously, Wheeler *et al.*[12] reported measurements on inversion-layer devices which are produced using fairly conventional photolithography but with gate electrodes as narrow as 300 nm. Their technique preserved the high mobility of these devices ($\sim 20\,000$ $cm^2/V\,s$). These workers found small fluctuations in the conductance with V_G in their high-conductance inversion layers. Skocpol *et al.*[13] used electron beam lithography to produce an extremely narrow gate electrode. Then, using the narrow nichrome gate wire as a mask, the oxide and Si adjacent to the device were removed by reactive-ion etching, forming a mesa structure. These devices revealed small fluctuations in conductance with V_G at relatively high conductance G, similar to the results of Wheeler *et al.* They also displayed fascinating time-dependent phenomena which will be discussed later. Using techniques described in Sec. III based on the wire-fabrication techniques of Giordano and Prober[2] and White and Flanders,[14] Kwasnick *et al.* fabricated[15] MOSFET's with gate wires as narrow as 50 nm. In these devices the electrostatic field, alone, is used to confine the electrons in the narrow inversion layer. These devices display exponentially large fluctuations of conductance with V_G near threshold and smaller fluctuations at higher V_G. Warren, Antoniadis, and Smith[16] have produced an array of 250 parallel one-dimensional inversion layers using a dual-gate structure in which the lower gate was a grating of 200-nm period. These devices also display small fluctuations in the conductance as a function of V_G for large G.

Each of these devices has strengths and weaknesses. The Pepper/IBM device requires no sophisticated lithography, apparently results in very narrow confinement, and shows none of the history-dependent effects characteristic of some other devices, but the geometry is poorly defined. The device of Wheeler *et al.* also uses conventional lithography and therefore preserves the high mobility, but the dimensions cannot be made very small. The devices of Skocpol *et al.* have very well-defined geometry and are the narrowest MOSFET's made to date, but the creation of a mesa which accurately determines the geometry also introduces surface states, and the electron beam lithography causes a deterioration of the mobility. In the devices to be discussed in this paper the high mobility of the Si(100) surface seems to be fairly well preserved. In principle, since the confinement results from the electrostatic field, the geometry can be modeled quite accurately. However, there are history-dependent effects which are not, as yet, well understood. The devices of Warren *et al.* may result in narrower inversion layers than those of Kwasnick *et al.* but the interpretation of the results for an array of conductors is less straightforward than for a single inversion layer.

The amazing fact is, however, that despite the great differences in the devices used, the results are qualitatively similar. The devices all show small fluctuations at high conductance G and, for those cases which have been carefully studied, exponentially large fluctuations at small G. In this paper we review what is known about the origin of the conductance fluctuations in Si MOSFET's. In Sec. II we outline the theoretical ideas which explain in a unified way the fluctuations in the limits of low and high G. Section III discusses the techniques we have used for fabrication of 1D MOSFET's. In Sec. IV we discuss the modeling of the confinement of the narrow inversion layer, which has been done with varying degrees of sophistication. In Sec. V we report the results of a detailed study of the exponentially large fluctuations in the low G regime. We include there, also, a discussion of history-dependent and switching phenomena similar to the telegraph noise observed by Ralls *et al.*[17] Finally, in Sec. VI we compare the results of experiments on the exponentially large fluctuations with theoretical predictions and point out the issues which are still unresolved. In particular, we use the conductance fluctuations to tentatively identify the transition from localized to extended states.

II. THEORETICAL BACKGROUND

For a truly one-dimensional conductor, in which the electrons are in a single quantum state for motion in both transverse directions, Mott and Twose[18] argued that any disorder causes all states to become localized. The localization length, the mean distance an electron moves before being backscattered, in this case is of the order of the mean free path for elastic collisions l. Thouless[1] explained that in wires of finite cross sectional area A, disorder would still give rise to localization but that the localization length ξ would, in general, be longer than l. Because of the increased phase space for scattering transverse to the length of the conductor, $\xi \sim l N_t$ where N_t is the number of transverse eigenstates occupied at the Fermi energy E_F (the number of transverse scattering channels). For large A, $N_t \sim k_F^2 A$ where k_F is the Fermi wave vector.

Consider, first, the $T=0$ conductance of a sample of a quasi-one-dimensional disordered metal between ideal metal contacts. Following the pioneering work of Lifshitz and Kirpichenkov,[19] Azbel[20] examined the case in which the length L of the quasi-one-dimensional sample is much longer than ξ, as depicted in Fig. 1. In the limit of zero voltage, the $T=0$ conductance is limited by tunneling from one contact to the other. Azbel identified three cases. (i) The Fermi energy E_F is not de-

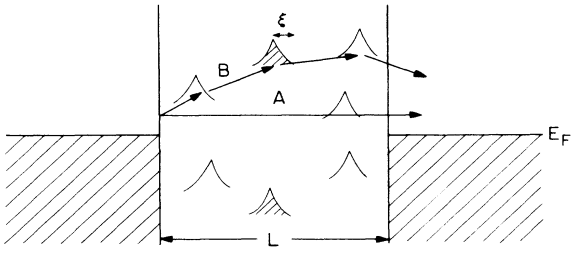

FIG. 1. Sketch of a disordered one-dimensional conductor between ideal contacts. The decay length ξ of the localized states is much less than the sample length L. States near the middle of the sample which have $T^* \sim 1$ are shaded. Path A depicts resonant tunneling, path B phonon-assisted tunneling.

generate with any of the localized states in the sample, and the tunneling transmission $T \sim e^{-2L/\xi}$. (ii) E_F is degenerate with a state at the center of the sample, and, by symmetry, $T \sim 1$. (iii) E_F is degenerate with a state closer to one contact than the other, and $T \sim e^{-4x/\xi}$ where x is the distance of the state from the center of the sample. The latter two cases are those of resonant tunneling. The transmission as a function of E_F, therefore, has a spectrum consisting of peaks with exponentially varying amplitude above a background level $\sim e^{-2L/\xi}$. For the resonances with $T \sim 1$, the Landauer formula[21] gives peak conductance $G \simeq e^2/h$. Azbel pointed out, also, that the widths of the peaks vary as $e^{-|L-2x|/\xi}$ because the width of a resonance is determined, through the uncertainty principle, by the time to tunnel to the closest contact. A sketch of the conductance versus E_F is given in Fig. 2(a).

Consider, now, what happens when the sample length

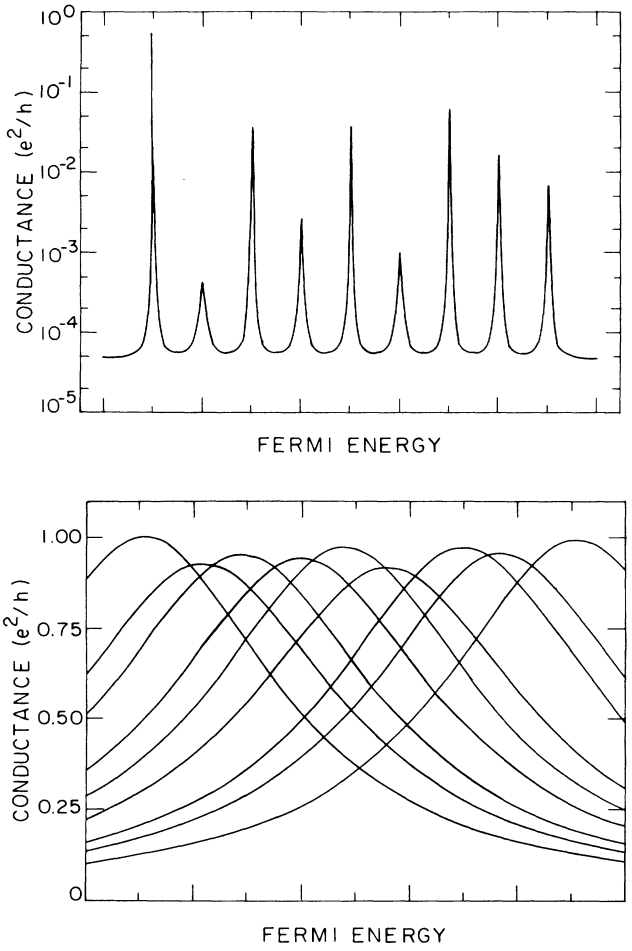

FIG. 2. (a) Sketch of logarithm of conductance G as a function of Fermi energy from resonant-tunneling model of Azbel. The resonances that are exponentially large are also exponentially narrow. (b) Sketch of overlapping-resonance picture of Thouless. Note that the measured conductance is the sum of the contributions from the individual resonances sketched in (b), whereas in (a) the total G, itself, is plotted.

is decreased or ξ is increased. In the limit $L < \xi$ every state gives $T \sim 1$ because every state is within ξ of the center of the sample. Furthermore, the states are no longer exponentially narrow relative to their spacing. Rather the resonances are broader than their energy spacing and overlap. Such a situation is sketched in Fig. 2(b). Since the conductance at any single value of E_F has contributions from many overlapping resonances, one has

$$G = \frac{e^2}{h} \delta E \frac{dN}{dE} , \tag{1}$$

where dN/dE is the density of levels and δE is the level width. In other words, $\delta E dN/dE$ is the number of resonances embraced by a single resonance. Equation (1) is, of course, just the result obtained by Thouless.[1] In his presentation, δE is replaced by h/t, and the Einstein relation is used to show that t is the time to diffuse to the contacts, i.e., $\delta E = hD/L^2$.

Thouless realized that the transition from localized to extended states occurs when the resonances just begin to overlap, i.e., when $G \simeq e^2/h$. As pointed out by Azbel, one expects exponentially large conductance fluctuations with E_F on the localized side of the transition [Fig. 2(a)]. It is clear from Fig. 2(b) that the fluctuations will be much smaller on the extended side.

Lee and Stone[22] and Altshuler[23] used diagrammatic techniques to show that the conductance in metals fluctuates from one sample to another, one Fermi level to another, or one magnetic field to another, and that the magnitude of the fluctuations is always $\sim e^2/h$ independent of geometry or sample size at $T = 0$. This result is surprising in the context of the overlapping-resonance picture of Fig. 2(b). If the resonances were distributed in energy according to Poisson statistics, one would expect

$$\Delta G = (Ge^2/h)^{1/2} , \tag{2}$$

which is, in general, larger than e^2/h. This result, which is a consequence of Campbell's theorem,[24] can be comfortably appreciated by analogy with shot noise.[24]

Altshuler and Shklovskii[25] explained how correlations in the energies of the resonances reduce the fluctuations from the size of Eq. (2). They relied on the Wigner-Dyson theory[26] of random matrices which was developed to solve the similar problem of the distribution of energies of resonances observed in nuclear scattering. The theory of random matrices is quite sophisticated. However, the results confirm a very simple suggestion of Wigner. This "Wigner surmise"[26] was that, given a level at E, the probability of finding the next level at $E+S$ is proportional to S for small S. This probability can be written $P(S) \propto S = \exp(-\ln S^{-1})$. The interaction of particles in a two-dimensional Coulomb gas is proportional to $\ln r^{-1}$, so the statistics of the level distribution is the same as that of the spatial distribution of the particles in a $d = 2$ Coulomb gas with $kT = 1$. The logarithmic dependence on the energy separation of the levels (or of the particle separation in the Coulomb gas) leads to the result that the fluctuations in the number of levels in an energy range E_c (or particle

density) follows

$$\langle [\delta N(E_c)]^2 \rangle \propto \ln \langle N(E_c) \rangle \ , \tag{3}$$

which is clearly much smaller than the result ($\propto N^{1/2}$) for Poisson-distributed levels given in Eq. (2).

Altshuler and Shklovskii[24] pointed out that the energy range used for calculating $N(E_c)$ is only defined to within $\pm\delta E$ for levels of width δE so that

$$\langle [\delta N(E_c)]^2 \rangle \propto \ln(E_c/\delta E) \ , \tag{4}$$

where E_c is the energy correlation range of the fluctuations. It is clear, however, that for the case of overlapping resonances, $E_c \sim \delta E$ giving fluctuations that are independent of $N(E)$ and δE.

Stated simply, level repulsion tends to keep the levels equally spaced in energy. This dramatically reduces the fluctuations from the size expected for Poisson-distributed levels. The level-repulsion becomes more effective as δE, the level width, increases, because the matrix elements between levels are larger. As a result, the fluctuations remain constant rather than growing with G. The fluctuations in the density of states turns out to be only one of two components. The second comes from fluctuations in D or, equivalently, δE. Imry[27] and Lee[28] have provided alternative intuitive arguments based on the generalized Landauer formula[21,29,30] and using properties of random matrices. We have here emphasized the approach of Altshuler and Shklovskii because it ties together the simple pictures of Azbel, for the localized regime, and Thouless, for the extended regime.

We thus have a unified theoretical picture of conductance fluctuations at $T=0$. In the limit that $L > \xi$ the states are localized and the fluctuations should be exponentially large. On the other hand, when $L < \xi$ the states are extended because of their quantum-mechanical coupling to the eigenstates of the ideal metal contacts, and the fluctuations should have magnitude e^2/h. In the delocalized regime the effect of increasing temperature is the same as that of reducing the sample length. An electron loses temporal phase coherence after suffering inelastic collisions, so the length over which the conductance is to be computed quantum mechanically (using the Landauer formula) is reduced from L to $L_{\rm in}(T)$, the inelastic length. This results in a gradual reduction of the conductance fluctuations when $L_{\rm in}$ is shorter than both ξ and L. More dramatic, however, is the prediction that decreasing temperature should cause a transition from extended to localized states when $L_{\rm in} = \xi$. This happens when the sample of length $L_{\rm in}$ has fluctuations $\Delta G \sim G \sim e^2/h$.

Increasing temperature is also expected to alter the Azbel resonant-tunneling mechanism in a fundamental way. Azbel developed the model to explain the large fluctuations seen by Fowler, Hartstein, and Webb[10] and by Pepper and Uren.[11] But there were both experimental and theoretical difficulties with this interpretation. The model predicts that the peaks in the conductance should be quenched with increasing temperature because inelastic scattering interferes with the coherent resonant-tunneling process. This was never observed. Instead, the conductance increases with T even at the maxima. Furthermore, the disordered samples were all, at that time, much too long to expect resonant tunneling from one contact to another.

Stone and Lee[31] have explained why resonant tunneling is not the conduction mechanism observed in long quasi-one-dimensional MOSFET's at currently accessible temperatures. Recall that the highest conductance resonances are extremely narrow because the tunneling time for these is exponentially long. As a result of this long tunneling time, electron-phonon collisions occur before the electron can utilize the high resonant conductance. As a result, thermally activated hopping is the dominant conductance mechanism. This is depicted in Fig. 1. It is still true, however, that resonant tunneling will dominate at sufficiently low T or in sufficiently short conductors as has recently been demonstrated by Fowler *et al.*[32]

Following the experiments of Fowler *et al.*[10] and Kwasnick *et al.*[15,33] which provided a fairly complete characterization of the fluctuations in the strong localization regime, Lee[34] proposed a new model based on thermally activated hopping. Using numerical calculations for a one-dimensional chain of uniformly spaced sites with random energies, Lee showed that the major features observed in the experiments could arise from variable-range hopping. His calculations showed that the average conductance increased with T as $\exp[-(T_0/T)^{1/2}]$ as expected from the usual Mott[35] argument. However, the resistance displayed exponentially large fluctuations with the chemical potential (Fermi level) of the electron gas. The fluctuations arise because, for any given value of E_F, the electron must encounter each one of a particular sequence of hops to traverse the one-dimensional sample. The resistance of a hop is an exponential function of the site separation and energy difference, both of which are random variables. The resulting distribution of hopping resistances is enormously broad (it is log-normal). At low T it becomes very likely, therefore, that a single hop, somewhere in the sample, limits the entire current.

This picture leads to several specific predictions. (i) At fixed E_F the temperature dependence of the resistance should be piecewise activated. The activation energy is fixed as long as a given hop limits the current. But when the temperature is lowered enough, the resistance of that hop becomes sufficiently high that the electron will hop to a more distant site with smaller activation energy. This mechanism is, of course, the origin of variable-range hopping itself, and the activation energy, on average, decreases in the way predicted by the $\exp[-(T_0/T)^{1/2}]$ dependence. But because of the one dimensionality, the activation energy should decrease in steps as T is lowered since a single hop will always limit the current.

(ii) The resistance R as a function of E_F is predicted to consist of regions in which R increases or decreases exponentially with E_F. This behavior is also easily understood in the model in which the current is limited by a single hop. From the work of Miller and Abrahams[36]

and Ambegaokar, Halperin, and Langer[37] we know that the highest resistance hops occur when the activation energy is large because the initial or final state has an energy far from E_F or because the energy difference between the two sites (if on opposite sides of E_F) is unusually large. The first case gives a resistance which is exponential in E_F/kT; in the second case the resistance is independent of E_F. Whether the latter situation (which gave plateaus in R versus E_F in the calculations of Lee) can be important had been earlier called into question by Thouless.[38] Since the predicted plateaus have not been seen we will not discuss this case in what follows.

(iii) One expects the dependence of the current on the voltage to be exponential and rectifying. Because one hop limits the current, the entire voltage drop occurs at that hop. In the limit of small voltage ($eV < kT$) the occupancy of the limiting site with energy E_i determines the current and is given by a Boltzmann factor $\exp[-(E_i - E_F)/kT]$. This situation is changed, however, when the voltage drop exceeds kT/e. For higher voltages the activation energy is expected to decrease linearly with the applied voltage for the following reasons: in the nonequilibrium situation, $eV > kT$, the current is limited by the rate at which electrons reach the highest-energy state rather than by its occupancy. (As noted above, we ignore the case of initial and final states on opposite sides of E_F.) Since the entire voltage (electrochemical potential) difference falls across the small region near the limiting state, the energy of that state relative to E_F changes by a fraction which we call f of eV, and f is of order one. This linear decrease of the activation energy results in an exponential dependence of current on voltage at fixed T. The fraction f will be different, in general, for the two directions of current flow. Therefore, the nonohmic behavior will give rise to rectification because of the microscopic properties of the hopping system, even though the sample is macroscopically symmetric.

Shklovskii[39] has discussed the effects of high fields in the three-dimensional hopping. His arguments for the percolation backbone in $d = 3$ are even more appropriate to a one-dimensional system. He pointed out that, beginning at very low applied voltage, the entire voltage falls across a single hopping resistance. However, as the voltage is increased, that resistance decreases exponentially until it is equal to that of the second-largest resistor in the chain. For higher voltages still, the voltage increment is shared equally by the two resistors until their resistances are reduced to that of the third largest and so on. One, therefore, predicts that the current will be piecewise exponential in voltage with a slope (on a logarithmic plot) that decreases in steps as more hops share the voltage drop.

As we will show, the experiments in the strong-localization regime are, in most aspects, in qualitative agreement with the one-dimensional-hopping model of Lee. However, several features, like the piecewise change in the exponent of the current-voltage curve, have not yet been observed.

We summarize the theoretical predictions for a wire long enough that its resistance is greater than h/e^2 as follows: at zero temperature the current will be limited by resonant tunneling from one contact to the other. At some finite temperature the rate for phonon-assisted tunneling will exceed that for resonant tunneling and variable-range hopping will become the dominant conductance mechanism. For these two low-temperature regimes, the conductance fluctuations will be exponentially large. At a still higher temperature, the inelastic length L_{in} will become as short as the localization length. At this temperature the states become delocalized, and at all higher temperatures the fluctuations in each subsample, of length L_{in}, will be of order e^2/h.

III. FABRICATION PROCEDURE

Figure 3 shows an overview of the method used to fabricate the gate wire. A photoresist (PR) step is used as a mask for reactive-ion etching (RIE), which creates a 50-nm step in the 100-nm thick thermally grown SiO_2. After removal of the photoresist, metal is evaporated at a glancing angle into the step creating the $\sim$50-nm wide gate wire.

The MOSFET's were produced on (100) p-type Si wafers of 3-Ω cm resistivity. First, a $\sim$0.5-μm-thick ("field") oxide was grown. Holes were etched in this oxide and the n^+ regions were produced by phosphorus diffusion, after which the top ($\sim$50 nm) P-doped oxide was etched away. The remaining oxide is left as an insulator over noncritical parts of the wafer. In a rectangular region overlapping each pair of n^+ contacts the oxide was completely removed and the 100-nm gate oxide was grown in dry O_2 at 1100 °C. The wafers were then

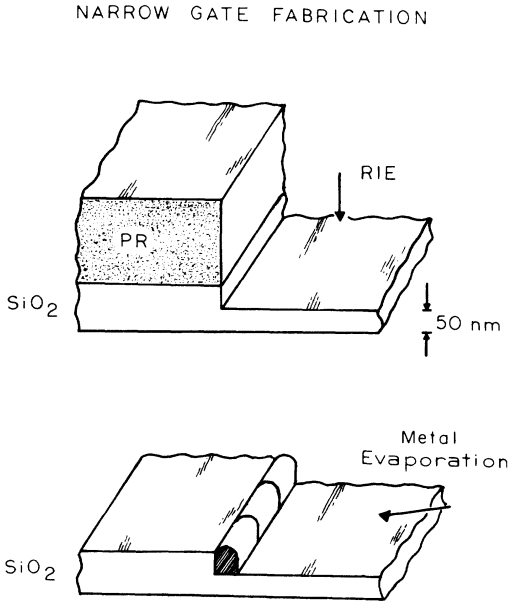

FIG. 3. Photoresist (labeled PR) is used as a mask for highly anisotropic reactive-ion etching (RIE) of thermal SiO_2. When half of the SiO_2 is removed, a 50-nm step is created in the oxide. Metal is evaporated into the step at a glancing angle to the surface to create the narrow gate wire.

patterned with photoresist to create the photoresist step extending from one n^+ contact to the other. Simultaneously, patterning was performed for test structures: gratings for secondary-emission monitor (SEM) examination of wires and wide MOSFET's for mobility measurements. The photoresist must have a smooth straight edge. However, interference effects during optical exposure of the photoresist results in a ragged foot at the base of the photoresist step. Plasma etching in an oxygen-helium mixture was used to remove the foot. The resulting photoresist step is used as a mask for reactive-ion etching at reduced power with CHF_3. The low power was chosen to minimize damage to the SiO_2. About half of the 100-nm-thick SiO_2 is removed leaving a step, an example of which is shown in Fig. 4.

After cleaning the oxide surface, electron-beam evaporation at a glancing angle was used to deposit the gate wire. In most of the experiments described here Al was used as the metal for the wire. Figure 5 shows a transmission electron micrograph (TEM) of a grating in SiO_2 onto which Al has been evaporated at an angle of $\sim 2.5°$ from the surface. One can see the wire (~ 70-nm wide) at the step between the thick (100 nm, darker colored) and thin (~ 50 nm, lighter colored) oxide. The right side of the micrograph shows the appearance of the surface after glancing-angle evaporation alone; the left side has been subjected to a brief etch to remove the small islands of Al which collect on surfaces away from the step. One clearly sees, on the right side, the shadow of the downward step of the grating. As can be seen in the magnified portion of the micrograph (Fig. 6) the grain size of the evaporated Al is comparable to the 70-nm wire diameter. This results in $\sim 20\%$ fluctuations in wire width and occasional breaks, one of which is seen in the figure.

In more recent devices we have used W instead of Al as the gate metal. Figure 7 shows a scanning-electron micrograph of a W wire and Fig. 8 shows one of an Al wire for comparison. The W wires are much more uniform and have fewer breaks resulting in a much higher yield of usable devices. However, the electronic properties of the resulting inversion layers do not appear to be limited by the wire morphology even for the Al wires.

Once gate wires were fabricated, holes were wet etched in the SiO_2 for contact to the n^+ regions and to the substrate, using photoresist to protect the gate wire and the rest of the wafer. This photoresist was removed and a new photoresist layer patterned to define Al contacts to the gate wire, n^+ regions, and the substrate. After Al evaporation, the photoresist was lifted off using acetone. This procedure prevents etching the gate wire. Finally, the wafers were annealed in forming gas at 410 °C.

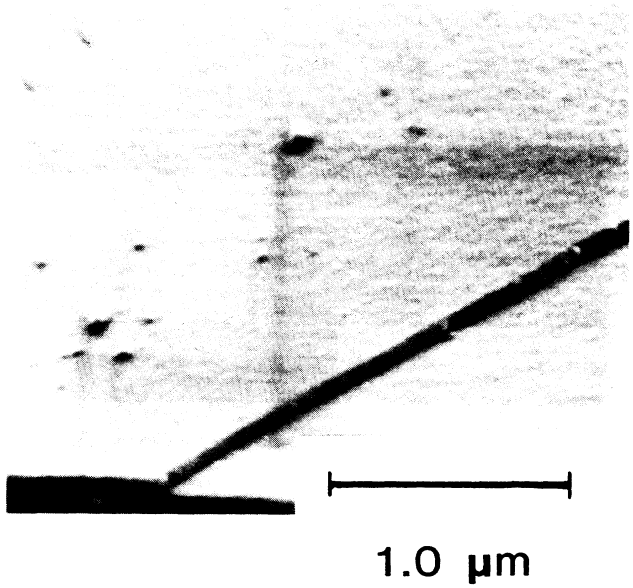

FIG. 4. Scanning electron micrograph of step reactive-ion etched in SiO_2.

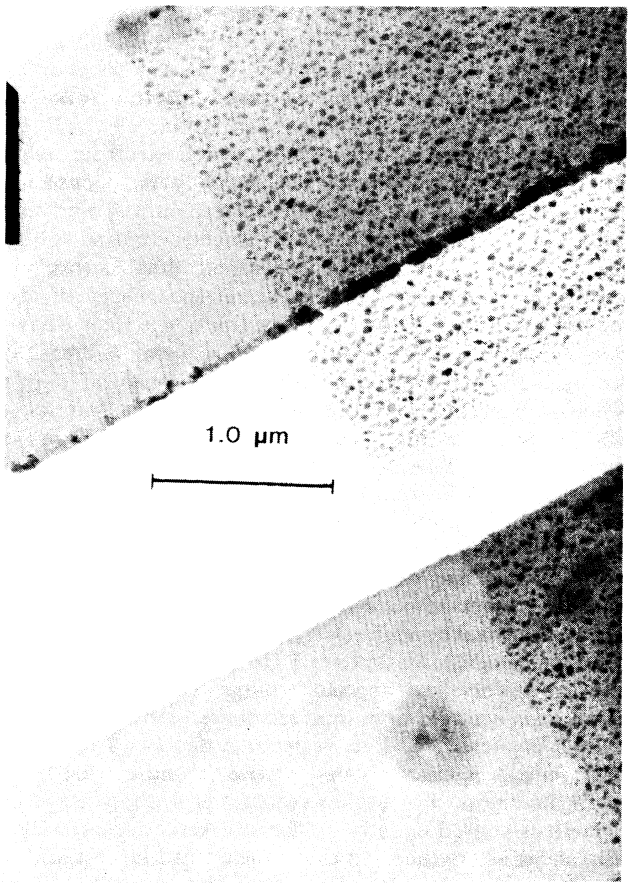

FIG. 5. Transmission electron micrograph of a section of a test grating for narrow-wire fabrication. The uniform darker region is the thick (~ 100 nm) oxide and the lighter region is the thin (~ 50 nm) oxide. Al, evaporated at a glancing angle to the surface from the lower right, collects in the upper grating step forming a wire. On the right is shown a region of the surface as it appears immediately after evaporation. One sees the shadow of the lower grating step and the Al islands that collect on the nearly tangential surfaces. The left side shows how a brief liquid etch removes the islands leaving only the narrow gate wire.

Figure 9 shows a schematic of the n^+ contacts and the gate wire with its large-area contacts. Using this geometry we create a device in which contact to the narrow inversion layer under the gate wire is made by wide inversion layers at each end. This should minimize the size of space-charge regions at the contacts.[40]

It is always difficult to determine the width of a submicron device. We have compared the transconductance at 77 and 4 K of narrow and wide devices prepared on the same wafer and find that the inversion layer is, like the gate wire, ~100-nm wide. This is consistent with models to be discussed next.

IV. MODELING OF THE POTENTIAL AND SUBBAND ENERGIES IN IDEAL 1D MOSFET'S

A question which must be addressed is whether the fluctuation phenomena observed in narrow MOSFET's are a consequence of quantum size effects or of disorder. To see if quantum size effects are possible, one must estimate the subband-energy spacings by modeling the potential in which the electrons move. An ideal narrow MOSFET can be modeled with various levels of sophisti-

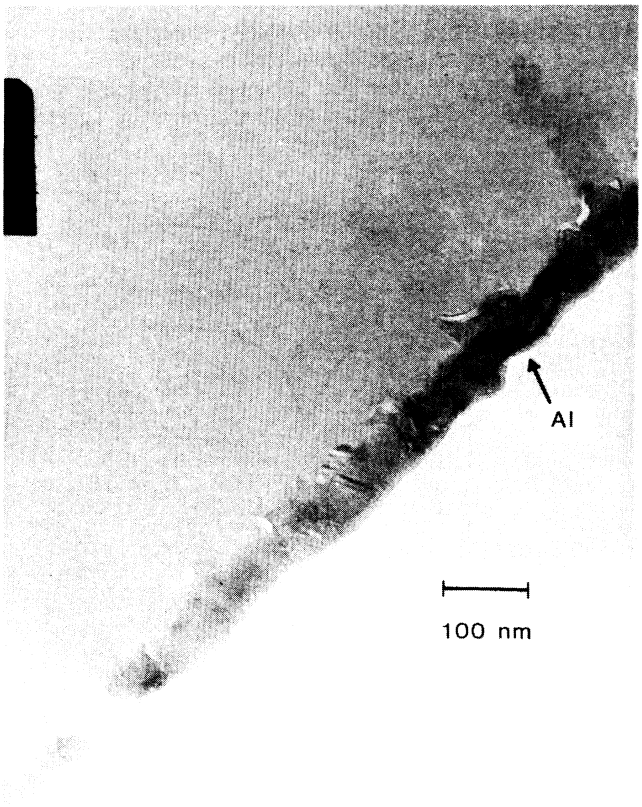

FIG. 6. Expanded region of the left side of the transmission electron microscopy (TEM) photograph of Fig. 5. Note that the individual grains of the Al wire have sizes comparable to the wires themselves. Note also that, because of the large grain size, breaks in the wire are not uncommon.

FIG. 7. Scanning electron micrograph of tungsten wire.

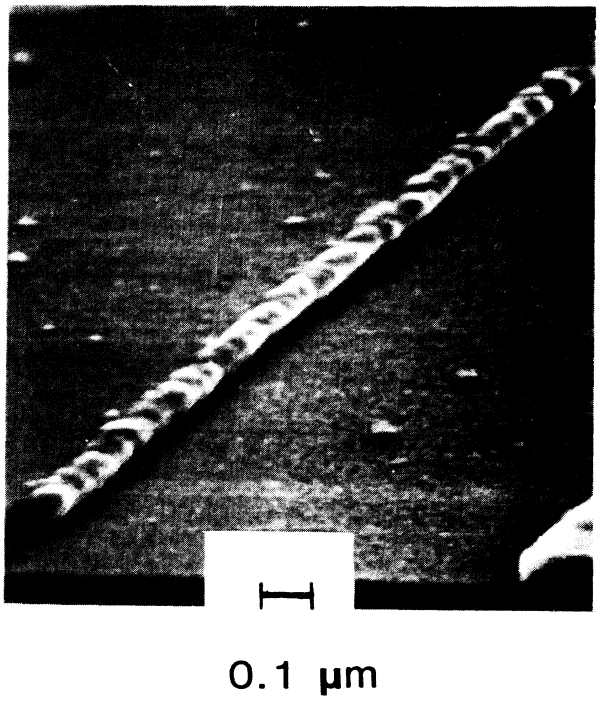

FIG. 8. Scanning electron micrograph of aluminum wire.

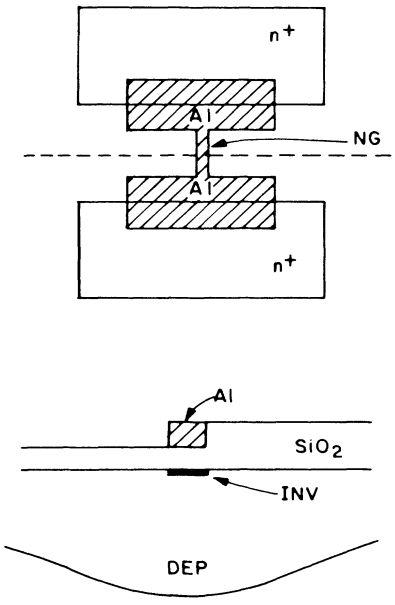

FIG. 9. Schematic of (above) top view of narrow-gate MOSFET. The width of the narrow gate (labeled NG) is exaggerated for clarity. The wide Al gates overlap the narrow (W or Al) wire. The contact to the narrow inversion layer is, therefore, made by wide inversion layers, minimizing contact effects. Also shown (below) is a schematic of a cross section through the device along the plane denoted by the dashed line (above). This indicates that the narrow inversion layer (INV) has about the same width as the narrow gate wire (in this case Al), but that the depletion region (DEP) is much wider.

cation. One might argue that the narrow MOSFET is simply a 2D MOSFET with a one-dimensional square-well confining potential. This may be a reasonable approximation for relatively wide inversion layers. Certainly, if the width W is much wider than both the gate oxide thickness t_{ox} and the depletion width L_D (several 100 nm), such an approximation would not be unreasonable. Wheeler *et al.*[12] have used this approach to estimate one-dimensional subband spacings for devices with $t_{ox} \simeq 40$ nm and $W \sim 800$ nm. On the other hand, for devices like ours, in which $t_{ox} \sim W \sim 50$ nm, such a square-well model is certainly inadequate.

Another rough estimate of the subband spacing may be obtained by assuming that the depletion region is semicylindrical and centered on the inversion layer. The potential of the depletion charge is then parabolic near the inversion layer. If the gate wire were narrow compared with t_{ox} and, of course, $t_{ox} \ll L_D$, the depletion depth, then the total potential felt by the inversion-layer electrons would be quadratic, to lowest approximation, and the subband energies for such a model are easily calculated.

The next level of sophistication uses classical computer simulations of the semiconductor. Programs to carry out such simulations are now widely available for use by the engineering community. We used a program[41] which solves Poisson's equation on a discrete grid approximating the device geometry, and uses Fermi-Dirac statistics to calculate the charge density. It does not, of course, account for changes in the electronic wave functions and charge densities caused by the confining potential.

The results of such a simulation are shown in Fig. 10 in which the potential is plotted as a function of distance transverse to the 70-nm-wide gate. Despite the shortcomings of the simulation, several reasonable results emerge. The width of the bottom of the potential well is less than ~ 100 nm from the threshold $V_T \sim 1.5$ V up to about 3 V. Of course, the width grows with V_G because, for a one-dimensional gate, the electric field is not perfectly screened by the inversion layer, as it is for a two-dimensional gate. Using the simulation we find that the charge density in the inversion layer drops off very sharply beyond a specific transverse distance for each V_G. This defines the width of the inversion layer W which is plotted against V_G in Fig. 11. Notice that the width grows very slowly for $V_G - V_T \gtrsim 1$ V.

Plotted also in Fig. 11 is the charge per unit length. The charge is quite linear with voltage for $V_G - V_T > 1$ V so we can determine the capacitance to be $C \simeq 0.8$ pF/cm. Describing the MOSFET as a parallel-plate capacitor of width 70 nm gives $C = 0.5$ pF/cm. On the other hand, assuming that it is half of a cylindrical capacitor gives $C \sim 1$ pF/cm. It is likely, therefore, that this value of the capacitance is reasonably accurate.

Note, in Fig. 10, that the potential energy has a very flat minimum. This results from the metallic screening

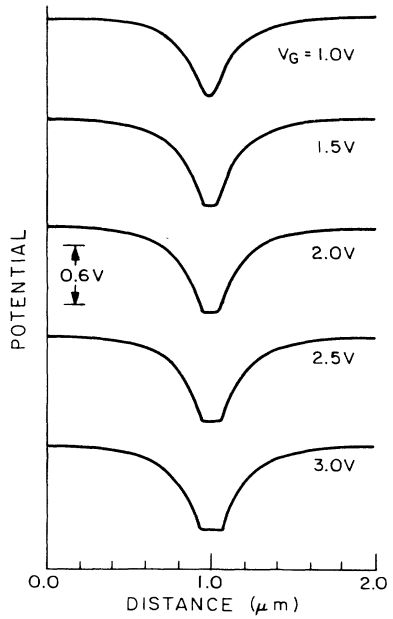

FIG. 10. Results of classical simulation of potential at the Si-SiO_2 interface in the direction transverse to the gate wire of the narrow MOSFET. The center of the wire is at 1.0 μm. The width of the bottom of the well is less than 100 nm up to $V_G \sim 3$ V.

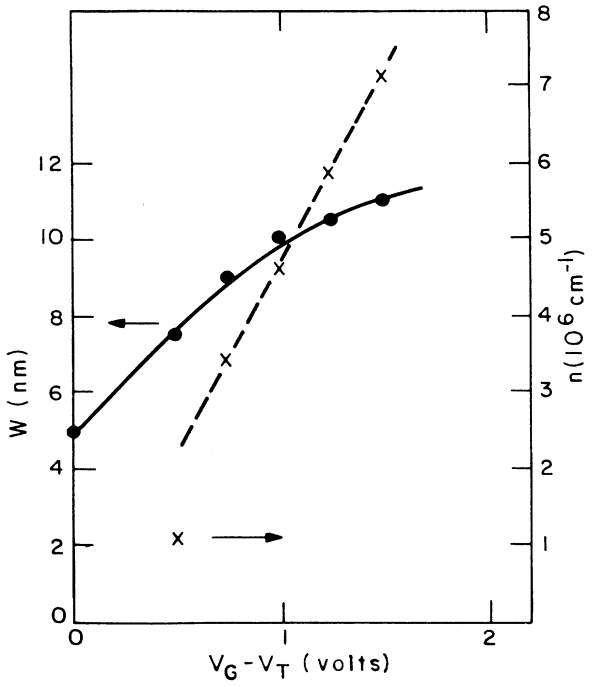

FIG. 11. Width W (circles) and total charge per unit length (crosses) of the inversion layer extracted from classical simulation. Curves are guides to the eye. Notice that n is quite linear with $V_G - V_T$ above ~ 0.8 V giving a constant capacitance ~ 0.8 pF/cm.

of the inversion layer. Therefore, a square well should be a reasonable approximation to the potential for the lowest subbands. In any event, all of the three semiclassical approaches discussed above give energy spacings for the lowest subbands of ~ 1 meV.

Laux and Stern[42] have carried out the first realistic calculations of subband spacings for one-dimensional MOSFET's. They predict subband spacings of $\sim 5-10$ meV for devices with inversion layers somewhat narrower than ours. Thus all simulations predict subband spacings larger than kT at 4 K. One would expect variations in the density of states as a result of this subband structure and concomitant conductance fluctuations if the one-dimensional MOSFET were ideal. However, the device is far from ideal. In analogy with the criterion for observation of Landau levels, we expect to detect subband structure only if $\hbar/\tau$ is less than the subband spacing, where τ is the elastic scattering time. From the mobility we find $\tau \sim 10^{-12}$ s so $\hbar/\tau$ is about the same size as the subband spacing, and the structure in the density of states might be barely observable. As we show below, however, in our devices there is strong evidence that disorder, rather than quantum size effects, is the origin of the conductance fluctuations.

V. RESULTS

A. Overview

We have measured the conductance of our narrow MOSFET's as a function of gate voltage V_G, drain-source voltage V_{DS}, and temperature. To obtain an overview of the fluctuation phenomena, look first at Fig. 12. This shows [Fig. 12(a)] the conductance as a function of gate voltage for a wide MOSFET fabricated on the same wafer as the narrow one. One sees the usual increase in mobility as T is lowered. The mobility at 4 K is ~ 8000 cm^2/V s, not as high as the values $\sim 20\,000$ cm^2/V s reported by Wheeler,[12] but reasonably high considering the high-energy radiation (from reactive-ion etching, electron-beam evaporation) to which this device, like our narrow MOSFET's, was subjected.

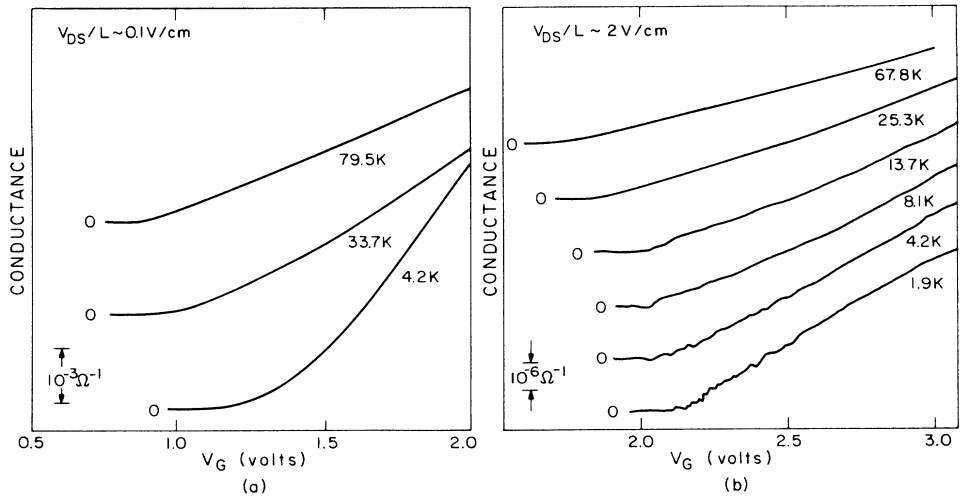

FIG. 12. Comparison of wide (a) and narrow (b) MOSFET's on the same Si wafer. The conductance is plotted as a function of gate voltage. Note that the zero of conductance is offset for each temperature and that the scale of conductance is $10^{-3}\ \Omega^{-1}$ for the wide MOSFET but $10^{-6}\ \Omega^{-1}$ for the narrow one. Source-drain fields indicated were small enough to ensure linearity in V_{DS}. In general, the threshold in 1D devices is higher than in 2D devices as seen here.

Turning to Fig. 12(b), one sees the same qualitative behavior for the narrow device as for the wide one, but with the development of noiselike, but reproducible, fluctuations of the conductance with gate voltage. It is important to understand that these fluctuations, which are largest near threshold, are variations with V_G *not with time.* We postpone until the end of this section the discussion of just how reproducible the structure is. Figure 13 shows an expanded view of the structure near threshold at 4.2 K and Fig. 14 shows how dramatic such a structure becomes at very low temperature. The peaks in Fig. 14 appear triangular in contrast with other reports.[43] Although the peak shape appears inconsistent with the predictions of Lee, we believe that such a conclusion must await experiments specifically designed to explore this question.

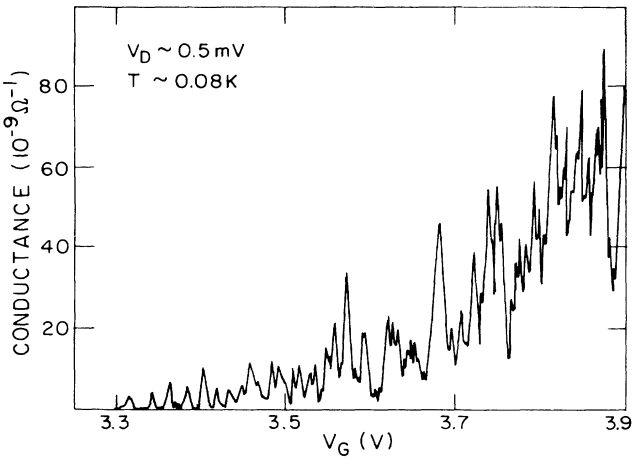

FIG. 14. Conductance vs gate voltage. The sample is held on a cold stage of a dilution refrigerator. The cold-stage temperature was 80 mK. However, the drain voltage is too large to be in the linear regime. The large voltage is required because the currents are so small.

B. Temperature dependence at fixed V_G and V_{DS}

To understand how the structure develops as T is lowered, we measured (Fig. 15) the temperature dependence of the current at fixed V_{DS} for several specific values of V_G. These are the maximum (PK) and minima (V1 and V2) indicated in Fig. 13. The current is thermally activated at all three gate voltages, but the activation energy varies from ~ 1 meV for the largest peak to ~ 3 meV for the deep minimum. From this we conclude, first, that the current is thermally activated at all values of V_G in the region of strong structure, even at the peaks, and, second, that the structure emerges at low T because the activation energy is a nonmonotonic function of V_G.

It is natural to ask whether the temperature dependence seen in Fig. 15 is a result of the shift of the structure to different V_G when T is changed. Figure 16 shows that this is not the case. The maxima and minima remain at fixed values of V_G up to the temperature at which the structure disappears.

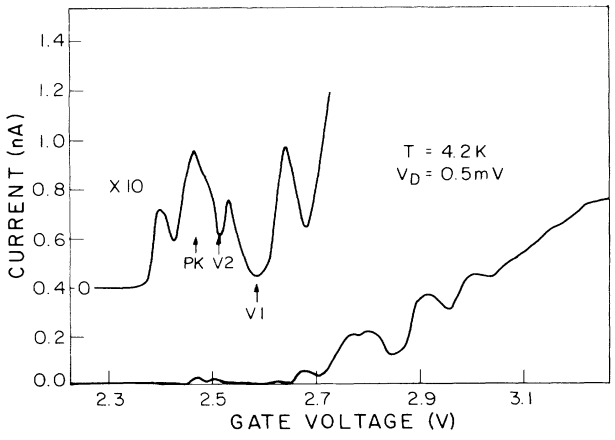

FIG. 13. Current vs gate voltage near threshold for ~ 70-nm-wide inversion layer. In the expanded (by 10) view a peak (PK) and two valleys (V1 and V2) are indicated by arrows for later reference.

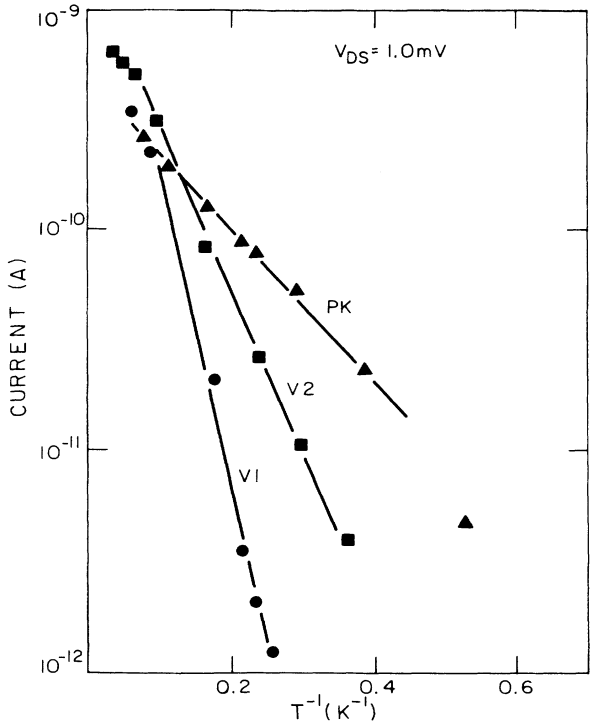

FIG. 15. Temperature dependence of current at fixed V_{DS} for the three values of V_G indicated by arrows in Fig. 13, the deepest minimum V1, a nearby peak PK, and a subsidiary minimum V2. The current is activated at all three points but the activation energy is ~ 0.7 meV for PK, 1.4 meV for V2, and 2.5 meV for V1. It then becomes smaller again at higher V_G.

Fowler, Hartstein, and Webb[10] argued that the temperature dependence of the conductance of their devices was not simply activated but rather followed the $\exp[-(T_0/T)^{1/2}]$ dependence expected for variable-range hopping in a one-dimensional, but constant, density of states. In Fig. 17 we plot the current at the minimum V1 against $T^{-1/2}$ to show that this dependence can fit our data, too. However, we do not believe that, for such a limited range of current, one can ascertain which temperature dependence is the correct one.

Notice that we see no evidence in Fig. 15 for the stepwise change in activation energy predicted by Lee.[34] Perhaps this, too, is just a result of the limited range of current. The reason that the current range is so limited is that, despite the high resistance of our narrow inversion layers, we must apply extremely low source-drain voltages V_{DS} ($<kT/e$) to be certain that the current is proportional to V_{DS}. Of course, parallel conductance channels would probably eliminate the piecewise-activated behavior. However, as we discuss in Sec. V C, the drain-voltage dependence militates against this explanation.

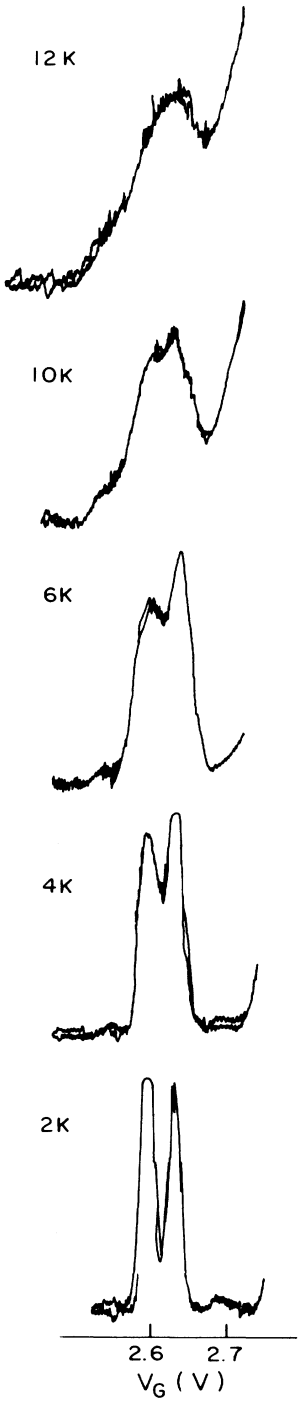

FIG. 16. Lowest-gate-voltage peaks at temperatures between 2 and 12 K. The structure does not shift to different positions in V_G, but the minima grow faster with increasing T than the maxima so that the structure disappears above ~25 K. Two scans of V_G are shown at each T showing that, for this small range of V_G, there is little hysteresis.

C. Dependence of current on V_{DS}

Our observation of a gate-voltage-dependent activation energy was not the first. Fowler, Hartstein, and Webb[10] had evidence that this was the way in which the exponentially large fluctuations developed. However, we were the first to report the dramatic dependence of the current on V_{DS}. Figure 18 shows the temperature dependence for the value of V_G corresponding to the deep minimum in Fig. 13, for various values of V_{DS}. The curve labeled $V_{DS}=1$ mV is the same as that labeled V1 in Fig. 15. As V_{DS} is increased the activation energy decreases dramatically. There is also evidence for a departure from the simple activated behavior at the lowest temperatures.

In Fig. 19 we have plotted the activation energy E_A as a function of V_{DS} for the two polarities. The activation

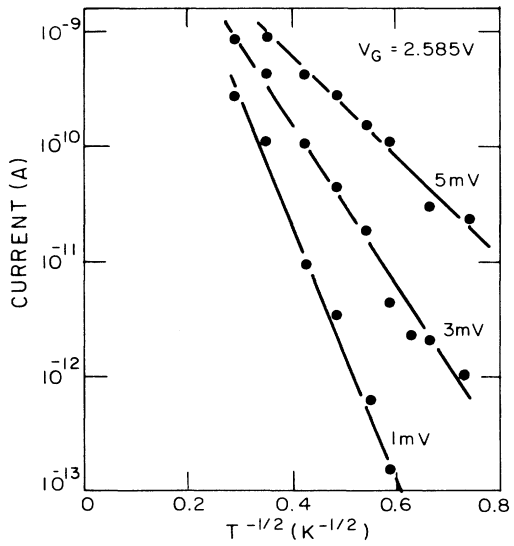

FIG. 17. Plot of log current vs $T^{-1/2}$ for $V_G=2.585$ V corresponding to the deep minimum (V1) in Fig. 13. The dependence on V_{DS} is discussed later.

energy decreases, roughly linearly with drain-source voltage. Thus

$$E_A \simeq E_0 - feV_{DS} \ , \tag{5}$$

where E_0 is the zero-voltage limit of the activation energy, and f is a fraction of order 0.3–0.4. Equation (5) requires that the current be exponential in V_{DS} for $efV_{DS} > kT$. Figure 20 shows that this is, indeed, the case.

At still higher V_{DS} the current begins to vary more slowly. The inset of Fig. 20 shows that $|I|$ varies linearly with $|V_{DS}|$ in this regime with the same differential resistance as observed at all $|V_{DS}|$ at high T where there is no nonlinearity at all.

The two sets of data in Fig. 20 are for opposite polarities of V_{DS}. This shows, quite dramatically, that there is rectification in this device which is macroscopically symmetric. The rectification varies nonmonotonically with gate voltage. This is one of the features explained by Lee's model.[34] Such nonlinearities and rectification have also been seen by Webb *et al.*[43] Because the rectification is so large, it may be that the current is limited by different hops for the two polarities. Further measurements are needed to explore this question.

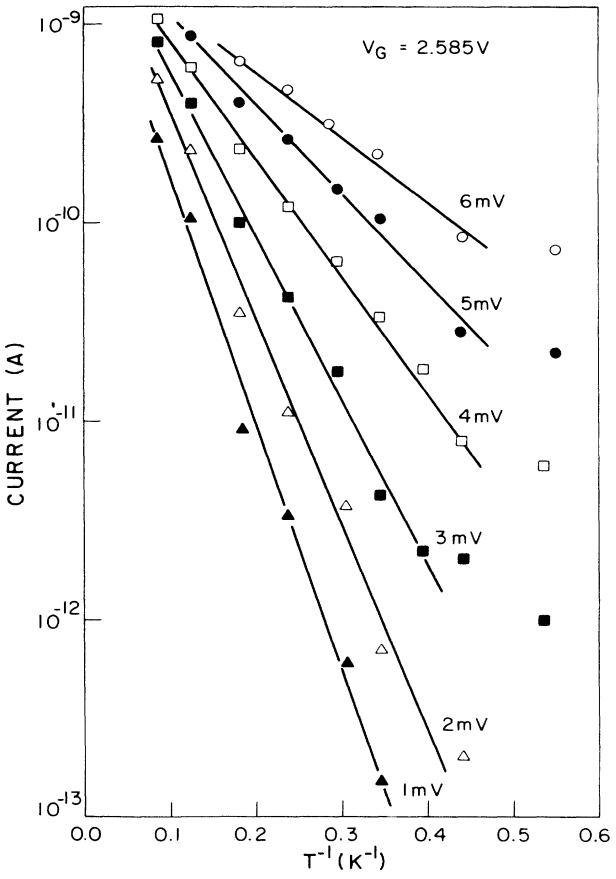

FIG. 18. T dependence of the current at the deep minimum (V1) in Fig. 13 for various values of V_{DS}.

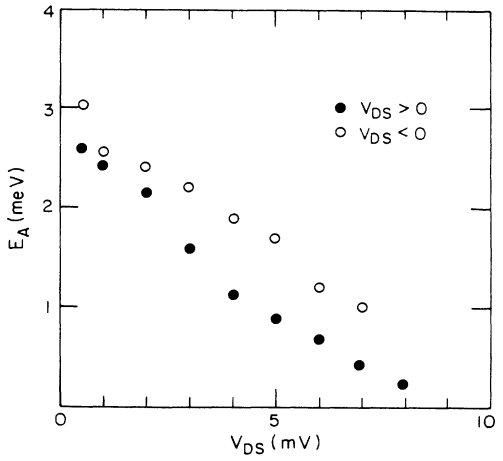

FIG. 19. Activation energy vs source-drain voltage from the data of Fig. 18 (solid circles), and for the opposite polarity of V_{DS} (open circles).

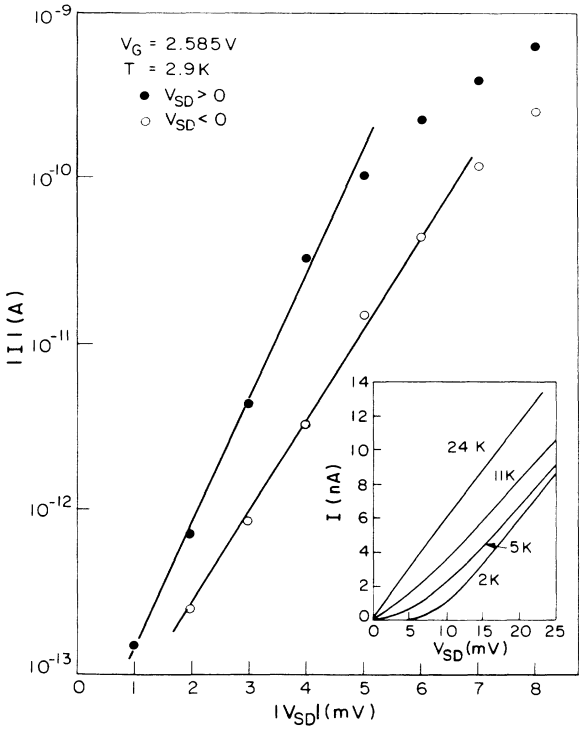

FIG. 20. Logarithm of current as a function of $|V_{DS}|$ for the two polarities at the deep minimum (V1) in Fig. 13. $|V_{DS}|$ is greater than kT/e for all the data presented here. The current is accurately exponential over more than two decades. Although the current varies more slowly at high $|V_{DS}|$, the stepwise decrease in the slope is not seen. The inset shows that this high-voltage regime corresponds to a constant differential resistance equal to the value measured at all $|V_{DS}|$ at high T. The large asymmetry in the current for the two polarities (almost a decade at ~ 4 mV) shows rectification in this macroscopically symmetric device.

The exponential dependence of I on V_{DS} does not, of course, extend to $V_{DS}=0$. When $feV_{DS}<kT$ the I-V characteristic must become linear as discussed in Sec. II. We cannot measure the extremely small current for $feV_{DS}<kT$ at 2 K for the deepest minimum in Fig. 13. However, by raising the temperature to 4.2 K one sees the crossover from linear to exponential behavior (Fig. 21). The crossover occurs at $eV_{DS}\simeq 0.5$ meV. Using $f\simeq 0.3$ from Fig. 19 we find that the crossover happens when feV_{DS} is of order kT at 4 K.

The hopping model predicts that the I-V characteristic should be piecewise exponential before becoming linear at the highest V_{DS}. We observe the predicted exponential behavior at low V_{DS} and the predicted linear behavior at high V_{DS}. We do not, however, observe the piecewise decrease in the exponential slope. This may be the result of the high temperature of our experiments. We hope to explore this issue in the future.

Figure 22 shows how some of the structure near threshold evolves with increasing V_{DS}. The apparent shift of the peaks results from the increasing average conductance at higher V_G. Comparing this with Fig. 16 makes it clear that the effect of V_{DS} is qualitatively similar to that of increasing temperature. This led us to worry a great deal about whether the nonlinear I-V characteristics were the result of heating of the electron gas. Such effects have been observed[44] in MOSFET's at fields comparable to those we are employing. However, several arguments have convinced us that the exponential I-V characteristic cannot arise from electron heating.

First, the temperature dependence for various V_{DS} is not that expected for electron heating. The heating of the electrons by the field would be negligible at high T,

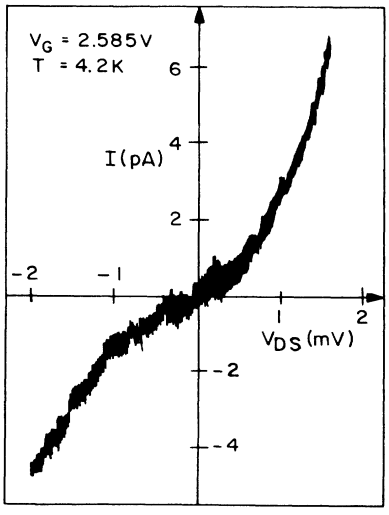

FIG. 21. Current vs source-drain voltage at the deep minimum (V1) in Fig. 13. By raising the temperature to 4 K the linear part of the I-V characteristic is seen at low $|V_{DS}|$, and the crossover to exponential behavior occurs at $|V_{DS}|\sim 0.5$ meV.

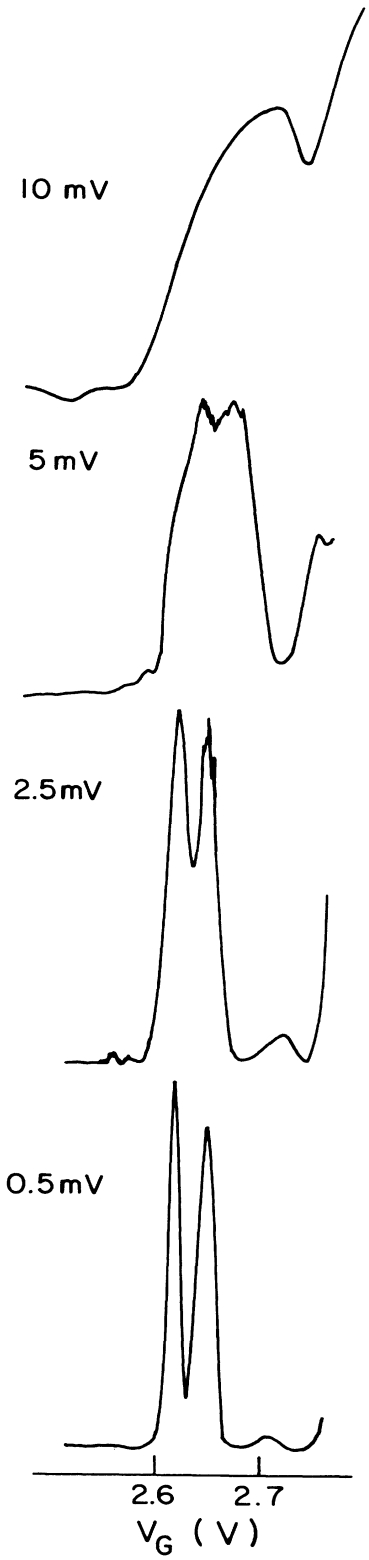

FIG. 22. Structure in the conductance vs gate voltage near threshold at various source-drain voltages. Comparison with Fig. 16 shows that the effect of increasing V_{DS} is qualitatively similar to that of increasing T. However, the quantitative dependence of current on V_{DS} cannot be explained by electron heating.

but would become dominant at low T. Therefore, the curves for different V_{DS} would be identical at high T and independent of T at low T. It is clear from Fig. 18 that, although there is evidence for flattening of the curves at the lowest temperatures, there is a strong voltage dependence at the high temperatures as well.

Second, the mobility of the electrons is too low to give substantial heating. The Joule heating of the gas of electrons is proportional to the conductivity,[44] and the heating per electron is therefore proportional to the mobility. Because of the strong localization, the average mobility of the electrons at the deep minimum in Fig. 13 is $\sim 10^3$ smaller than in two-dimensional MOSFET's. For the latter, the temperature rise at the highest fields employed in Figs. 17–22 (~ 10 V/cm) is 10 K. Therefore, the heating in the narrow devices should give rise to a temperature rise of only 10^{-2} K, assuming that the heat dissipation rate by electron-phonon scattering is comparable for wide and narrow devices.

Last, heating would result in thermal runaway which is not observed. The temperature dependence of the conductance is thermally activated, and for activated conductance G, an increase in electron temperature adequate to cause an increase of G by a factor of ~ 3 would cause runaway.[33] The conductance would switch to a high value and the current would be limited only by the external load resistor. Such thermal switching was never observed for our narrow MOSFET's. The evidence is compelling, therefore, for a nonthermal mechanism for the highly nonlinear dependence of current on V_{DS}.

D. History-dependent phenomena

As mentioned earlier, several history-dependent effects have been observed in our narrow MOSFET's. When the temperature is cycled to above ~ 200 K the structure near threshold is changed. The new pattern is qualitatively similar to the old, but the positions of maxima and minima like those in Fig. 13 are completely different.

One does not need to cycle the temperature to alter the structure. It also changes when the gate is raised to a potential well above threshold. For small variations in V_G, the pattern of fluctuations remains unaltered. This can be seen in Fig. 16, for example. However, for larger excursions of V_G, it changes dramatically. The careful reader will have observed that the expanded version of the structure in Fig. 13 is not exactly the same as the unexpanded curve. That is because the trace for the larger range of V_G was taken first.

These history-dependent effects are, of course, very disconcerting. In order to make measurements like those in Figs. 15–22 we were forced to equilibrate the sample at fixed V_G (or nearly fixed if V_G is to be varied) for several hours. On the other hand, the mutability of the pattern of fluctuations is evidence that the structure is a result of disorder. At relatively high temperatures, either ionic motion or electron trapping in the oxide can alter the random potential sampled by the electrons in the inversion layers. However, since variation of V_G at low T also changes the structure, electron trapping may be the most important cause of the variability of the random potential.

Ralls *et al.*[17] were the first to provide direct evidence that certain history-dependent effects result from trapping of the electrons in the oxide. They observed "telegraph" noise, switching between two discrete values of the conductance, for V_G well above threshold. From the dependence of the noise on Fermi energy (or V_G) and temperature they showed that it arises from the capture and release of a single electron by a specific trap.

In the regime explored by Ralls *et al.*[17] the fluctuations in conductance caused by trapping are small, of order a few percent. They are, in fact,[45] of the size predicted by the theory of Lee and Stone.[22] In the strong localization regime, the fluctuations resulting from such trapping events can be exponentially large. Figure 23 shows the dependence of current on V_{DS} at fixed V_G. The current repeatedly switches between two curves each of which displays the nonlinear dependence on V_{DS} as well as the rectification discussed above. For other minima we have observed that, at small V_{DS}, the change of conductance can be exponentially large. Figure 24 shows another example. In this case a peak in the plot of current versus V_G is clearly present when V_G is raised above threshold, but disappears when V_G is lowered again. Presumably, this is simply an example of a trap with an extremely long release time at the low temperature of the experiment.

VI. DISCUSSION AND CONCLUSIONS

As we have already emphasized, the variable-range-hopping model of Lee[34] is adequate to describe most of the behavior of the exponentially large fluctuations in the conductance near threshold in narrow MOSFET's: The activated temperature dependence is the result of localization; the nonmonotonic dependence of the conductance on Fermi energy which gives rise to the exponential fluctuations comes about because a small fraction of the hopping resistances limit the current in these one-dimensional devices; the exponential dependence on source-drain voltage has the same origin.

Some features of Lee's model have not been observed. One is the plateaus in conductance versus Fermi energy. These arise in the model from hopping between states on opposite sides of the Fermi energy, and, as pointed out by Thouless,[38] may not carry much current. Another is the piecewise activated behavior of the conductance at fixed Fermi energy leading, on average, to the $\exp[-(T_1/T)^{1/2}]$ dependence. This will be evident only at temperatures sufficiently low that variable-range hopping can compete with nearest-neighbor hopping.[35] The latter corresponds to hopping between spatially overlapping localized states. In other words, the activation energy can never exceed the energy spacing of levels located within ξ of each other. The last missing feature is the piecewise exponential dependence of current on source-drain voltage. This may be hidden because differences in slope of less than kT/e cannot be resolved. These effects will be sought in future experiments.

We have not yet discussed the universal conductance fluctuations seen at high V_G. These were first reported

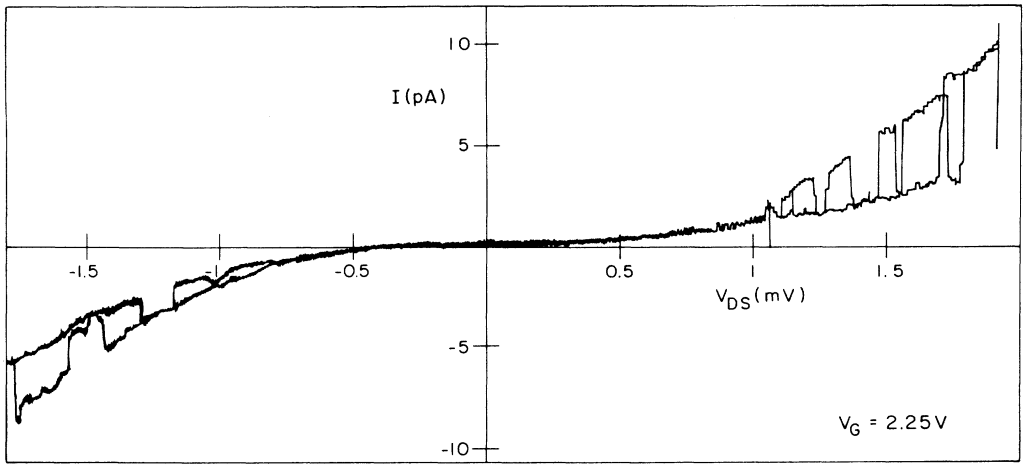

FIG. 23. Bistability in dependence of current on source-drain voltage at a minimum in the dependence of I on V_G. Two scans of I vs V_{DS} are shown to demonstrate the very large telegraph noise. The small steplike structure seen at positive V_{DS} is an experimental artifact.

by Webb *et al.* for small wires and rings.[9] Licini *et al.* were the first[45] to report such fluctuations in the magnetoresistance of MOSFET's and to show that fluctuations of similar size were seen when the Fermi energy is varied. Subsequently, Skocpol *et al.*[46] showed that the magnetoresistance fluctuations scaled in the expected way with the inelastic diffusion length and sample dimensions. Kaplan and Hartstein[47] showed that, as expected, the fluctuations depend only on the perpendicular component of the field.

The scaling theory of Thouless[1] predicts that as the sample length is increased (at $T=0$) the transition from extended to localized states will occur when the sample length equals the localization length, that is, when the conductance decreases to e^2/h. The theory of Lee and Stone[22] predicts fluctuations of magnitude e^2/h on the extended side of this transition. The theories of Azbel[20] and Lee[34] predict exponentially larger fluctuations on the localized side. One therefore expects that at the extended-to-localized transition the fluctuations are about the same size as the conductance itself.

As usual in the localization problem, we follow Thouless's suggestion[1] that finite temperature effects can be incorporated by taking $L_{\rm in}(T)$ as the sample length. At fixed T, both $L_{\rm in}$ and the localization length depend on the conductance per unit length, the 1D conductivity. It may be, however, that in a one-dimensional system the localization length increases more rapidly with conductivity than $L_{\rm in}$ does.[6] This point is discussed below. Therefore, we expect that at some value of V_G, in Fig. 13 for example, the system crosses over from the localized to extended regime as V_G and, therefore, the conductance is increased.

We now estimate the critical conductance G_c at which the transition occurs for our device. We have

$$G_c = \frac{e^2}{h}\frac{L_{\rm in}}{L}\ , \tag{6}$$

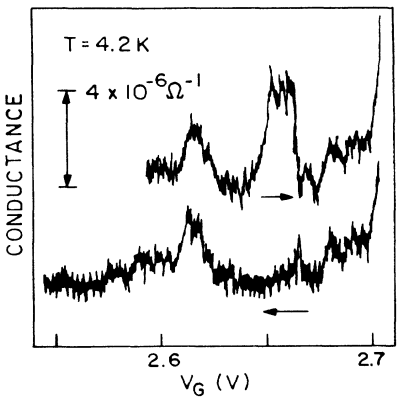

FIG. 24. Hysteresis in conductance vs V_G near threshold. One peak is completely absent for one direction of scan (note arrows) but not the other.

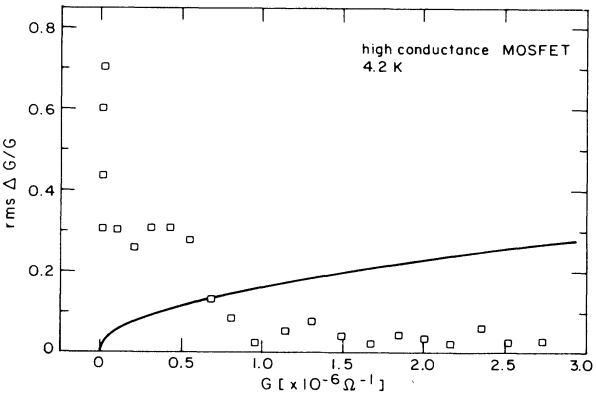

FIG. 25. Fractional fluctuations in conductance as a function of average conductance. The solid curve is $(Gh/e^2)^{1/2}$.

because, at the transition, the sample consists of $L/L_{\rm in}$ segments in series, each with conductance e^2/h. In addition, we use the result[30] that the fluctuations for the $L/L_{\rm in}$ resistors in series add incoherently. Converting resistance fluctuations to conductance fluctuations gives

$$\Delta G_c = \frac{e^2}{h}\left[\frac{L_{\rm in}}{L}\right]^{3/2}, \tag{7}$$

and eliminating $L_{\rm in}/L$ from Eqs. (6) and (7) yields

$$\frac{\Delta G}{G_c} = \left[G_c \frac{h}{e^2}\right]^{1/2}. \tag{8}$$

One sees from Fig. 13 that near $V_G = 3$ V, $G \simeq 10^{-6}\ \Omega^{-1}$ and $\Delta G/G \sim 0.2$, which satisfies Eq. (8) quite well. We, therefore, identify $10^{-6}\ \Omega^{-1}$ as the value of G at which the localized-to-extended transition occurs at 4 K. The data of Fig. 14 do not extend to high enough values of V_G to reach the transition at 80 mK.

To make the comparison more dramatic we plot in Fig. 25 the fluctuations in G against the average value of G. The average G is determined by a polynomial least-squares fit. Plotted also is $(Gh/e^2)^{1/2}$. One sees that the fluctuations increase rapidly once they exceed $(Gh/e^2)^{1/2}$.

Using $G_c \simeq 10^{-6}\ \Omega^{-1}$ in Eq. (6) gives $L_{\rm in}$ (and the localization length, as well) of the order of 100 nm. Wheeler *et al.*[12] found the inelastic length to be ~ 200 nm at 4 K for devices with mobilities ~ 2 times larger than ours. We conclude, therefore, that the value of 100 nm is quite reasonable. We plan experiments to explore the localized-to-extended transition for a wide range of temperatures in the near future.

This indication of the localized-to-extended transition at finite T raises some interesting questions. It is usually found[46] that $L_{\rm in}$ is proportional to the conductance per unit length G/L. Since the localization length ξ is also proportional to G/L, one might have expected to have localized states at all values of V_G or extended states at all V_G. That is, the transition would be seen as T is varied but not as G is varied. The observation of the transition suggests that either $L_{\rm in}$ increases more slowly than linear with G or ξ increases more rapidly than linear. Either of these may happen when k_F^{-1} approaches the elastic scattering length, for then the perturbation theory used to calculate these quantities is no longer valid.

ACKNOWLEDGMENTS

We acknowledge discussions with P. A. Lee and A. D. Stone. The fabrication of the narrow gate was carried out in the Submicron Structures Laboratory at MIT. We thank J. Melngailis, H. I. Smith, J. Carter, and other members of the laboratory for guidance and assistance. We also thank R. W. Mountain and colleagues at Lincoln Laboratory for performing initial fabrication steps. This work was supported, in part by the U.S. Joint Services Electronics Program under Contract No. DAA103-86-K-0002 and by National Science Foundation (NSF) Grant No. ECS-85-03443.

*Present address: General Electric Company, Corporate Research and Development, Schenectady, NY 12301.

[1]D. J. Thouless, Phys. Rev. Lett. **39**, 1167 (1977).
[2]N. Giordano, W. Gibson, and D. E. Prober, Phys. Rev. Lett. **43**, 725 (1979).
[3]N. Giordano, Phys. Rev. B **22**, 5635 (1980); G. Bergman, Phys. Rep. **107**, 1 (1984); P. Santhanam and D. E. Prober, Phys. Rev. B **29**, 3733 (1984).
[4]F. J. Wegner, Z. Phys. B **25**, 327 (1976).
[5]E. Abrahams, P. W. Anderson, D. C. Licciardello, and T. V. Ramakrishnan, Phys. Rev. Lett. **42**, 673 (1979).
[6]P. A. Lee and T. V. Ramakrishnan, Rev. Mod. Phys. **57**, 287 (1985).
[7]G. J. Dolan and D. D. Osheroff, Phys. Rev. Lett. **43**, 721 (1979).
[8]D. J. Bishop, R. C. Dynes, and D. C. Tsui, Phys. Rev. B **26**, 773 (1980); Phys. Rev. Lett. **44**, 1153 (1980).
[9]C. P. Umbach, S. Washburn, R. B. Laibowitz, and R. A. Webb, in *Proceedings of the 7th International Conference on Low Temperature Physics LT-17, Part II-Contributed Papers,* edited by U. Eckern, A. Schmid, W. Weber, and H. Wuhl (Elsevier, New York, 1984), pp. 717 and 718; Phys. Rev. B **30**, 4048 (1984).
[10]A. B. Fowler, A. Hartstein, and R. A. Webb, Phys. Rev. Lett. **48**, 196 (1982).
[11]M. Pepper and M. J. Uren, J. Phys. C **15**, L617 (1982).
[12]R. G. Wheeler, K. K. Choi, A. Goel, R. Wisnieff, and D. E. Prober, Phys. Rev. Lett. **49**, 1674 (1982).
[13]W. J. Skocpol, L. D. Jackel, E. L. Hu, R. E. Howard, and L. A. Fetter, Phys. Rev. Lett. **49**, 951 (1982); Physica **117B&118B**, 667 (1983).
[14]D. C. Flanders and A. E. White, J. Vac. Sci. Technol. **19**, 892 (1981).
[15]R. F. Kwasnick, M. A. Kastner, J. Melngailis, and P. A. Lee, Phys. Rev. Lett. **52**, 224 (1984).
[16]A. C. Warren, D. A. Antoniadis, and H. I. Smith, Phys. Rev. Lett. **56**, 1858 (1986).
[17]K. S. Ralls, W. J. Skocpol, L. D. Jackel, R. E. Howard, R. W. Epworth, and D. M. Tennant, Phys. Rev. Lett. **52**, 228 (1984).
[18]N. F. Mott and W. D. Twose, Adv. Phys. **10**, 107 (1961).
[19]I. M. Lifshitz and V. Ya. Kirpichenkov, Zh. Eksp. Teor. Fiz. **77**, 989 (1979) [Sov. Phys.—JETP **50**, 499 (1980)].
[20]M. Ya. Azbel, Solid State Commun. **45**, 527 (1983).
[21]R. Landauer, IBM J. Res. Dev. **1**, 223 (1957); Philos. Mag. **21**, 863 (1970).
[22]P. A. Lee and A. D. Stone, Phys. Rev. Lett. **55**, 1622 (1985).
[23]B. L. Altshuler, Pis'ma Zh. Eksp. Teor. Fiz. **41**, 53 (1985) [Sov. Phys.—JETP Lett. **41**, 648 (1985)].
[24]See, for example, E. Parzen, *Stochastic Processes* (Holden-Day, Oakland, California, 1962), pp. 149 and 155.
[25]B. L. Al'tshuler and B. I. Shklovskii, Zh. Eksp. Teor. Fiz. **91**,

220 (1986) [Sov. Phys.—JETP Lett. **64**, 127 (1986)].
[26]For references see Refs. 24 and 30, and M. Mehta, *Random Matrices and the Statistical Theory of Energy Levels* (Academic, New York, 1967).
[27]Y. Imry, Europhys. Lett. **1**, 249 (1986).
[28]P. A. Lee, Physica **140A**, 169 (1986).
[29]P. W. Anderson, D. J. Thouless, E. Abrahams, and D. S. Fisher, Phys. Rev. B **22**, 3519 (1980).
[30]Y. Imry, in *Directions in Condensed Matter Physics,* edited by G. Grinstein and E. Mazenko (World Publishing, Singapore, 1986); P. A. Lee, A. D. Stone, and H. Fukuyama, Phys. Rev. B **35**, 1039 (1987).
[31]A. D. Stone and P. A. Lee, Phys. Rev. Lett. **54**, 1196 (1985).
[32]A. B. Fowler, G. L. Timp, J. J. Wainer, and R. A. Webb, Phys. Rev. Lett. **57**, 138 (1986).
[33]For details, see R. F. Kwasnick, Ph.D. thesis, Massachusetts Institute of Technology, 1984.
[34]P. A. Lee, Phys. Rev. Lett. **53**, 2042 (1984).
[35]N. F. Mott and E. A. Davis, *Electronic Processes in Noncrystalline Materials,* 2nd ed. (Oxford University Press, Oxford, 1979), and references therein.
[36]A. Miller and E. Abrahams, Phys. Rev. **120**, 745 (1960).
[37]V. Ambegaokar, B. I. Halperin, and J. S. Langer, Phys. Rev. B **4**, 2612 (1971).
[38]D. J. Thouless, *Ill-Condensed Matter,* 1979 Les Houches Notes (North-Holland, Amsterdam, 1979); *The Metal Non-Metal Transition in Disordered Systems* (Scottish Universities Summer School in Physics, Edinburgh, 1978), p. 88.
[39]B. I. Shklovskii, Fiz. Tekh. Poluprovodn. **10**, 1440 (1976); [Sov. Phys. Semicond. **10**, 855 (1976)].
[40]P. H. Woerlee, G. A. M. Hurkx, W. J. M. Josquin, and J. F. C. M. Verhoeven, Appl. Phys. Lett. **47**, 700 (1985).
[41]J. A. Greenfield, S. E. Hansen, and R. W. Dutton, Stanford Electronics Laboratory Technical Report No. G-201-7, 1980 (unpublished).
[42]S. E. Laux and F. Stern, Appl. Phys. Lett. **49**, 91 (1986).
[43]R. A. Webb, A. Hartstein, J. J. Wainer, and A. B. Fowler, Phys. Rev. Lett. **54**, 1577 (1985).
[44]T. Ando, A. B. Fowler, and F. Stern, Rev. Mod. Phys. **54**, 437 (1982).
[45]J. C. Licini, D. J. Bishop, M. A. Kastner, and J. Melngailis, Phys. Rev. Lett. **55**, 2987 (1985).
[46]W. J. Skocpol, P. M. Mankiewich, R. E. Howard, L. D. Jackel, D. M. Tennant, and A. D. Stone, Phys. Rev. Lett. **56**, 2865 (1986).
[47]S. B. Kaplan and A. Hartstein, Phys. Rev. Lett. **56**, 2403 (1986).

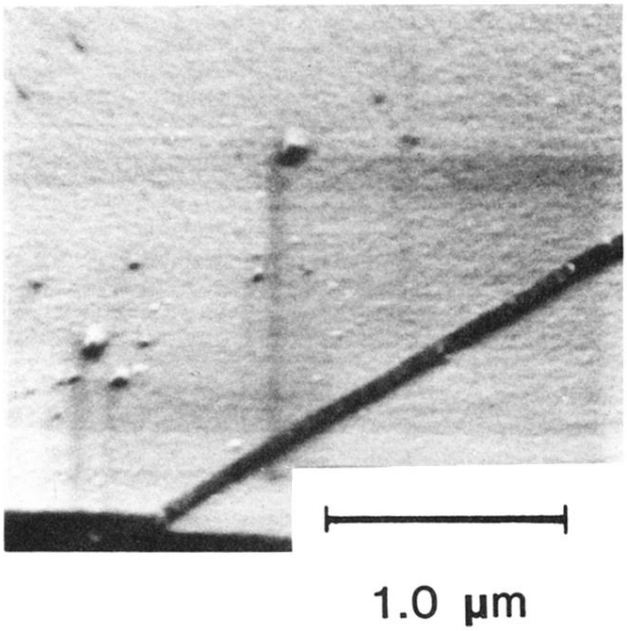

FIG. 4. Scanning electron micrograph of step reactive-ion etched in SiO_2.

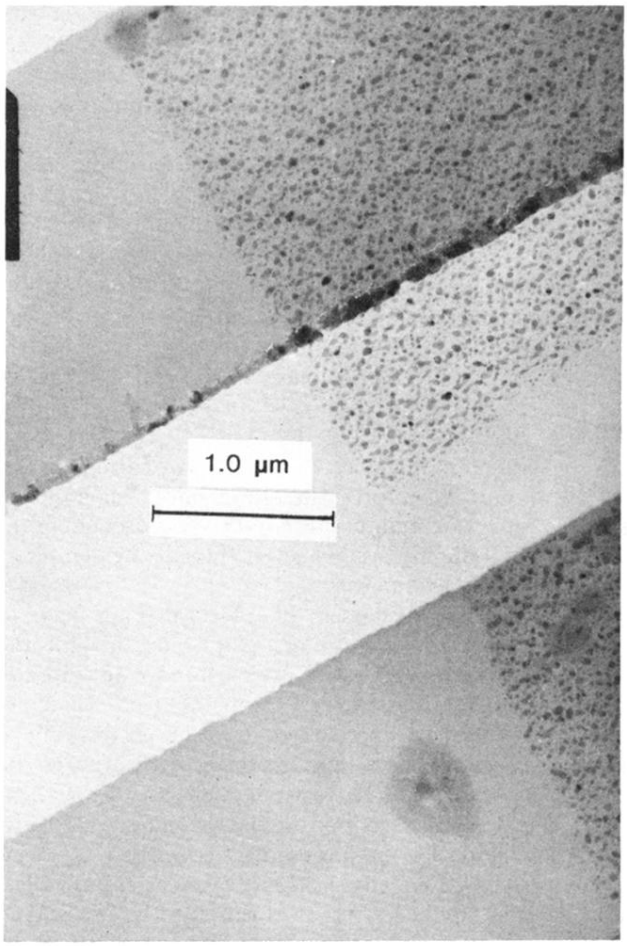

FIG. 5. Transmission electron micrograph of a section of a test grating for narrow-wire fabrication. The uniform darker region is the thick ($\sim$100 nm) oxide and the lighter region is the thin ($\sim$50 nm) oxide. Al, evaporated at a glancing angle to the surface from the lower right, collects in the upper grating step forming a wire. On the right is shown a region of the surface as it appears immediately after evaporation. One sees the shadow of the lower grating step and the Al islands that collect on the nearly tangential surfaces. The left side shows how a brief liquid etch removes the islands leaving only the narrow gate wire.

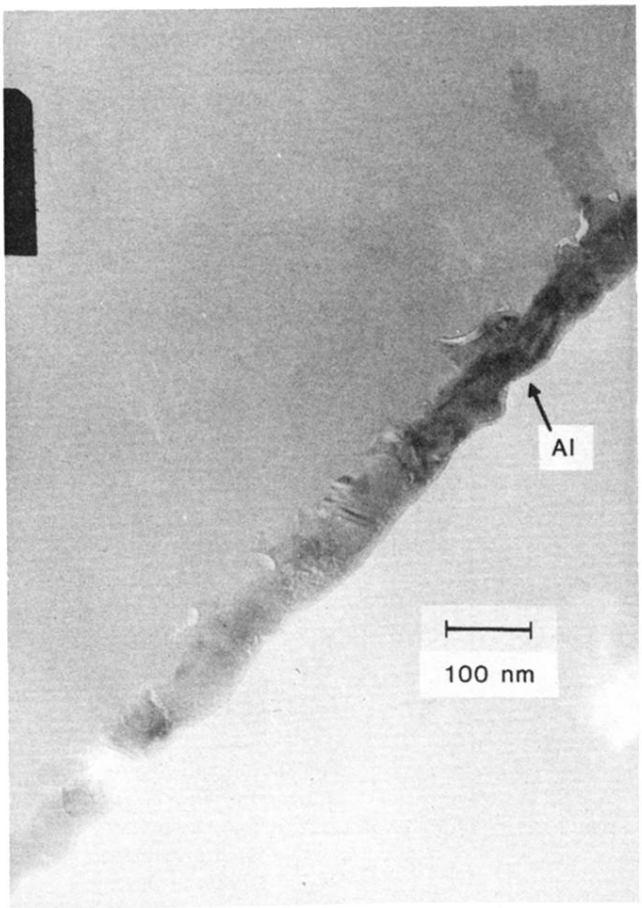

FIG. 6. Expanded region of the left side of the transmission electron microscopy (TEM) photograph of Fig. 5. Note that the individual grains of the Al wire have sizes comparable to the wires themselves. Note also that, because of the large grain size, breaks in the wire are not uncommon.

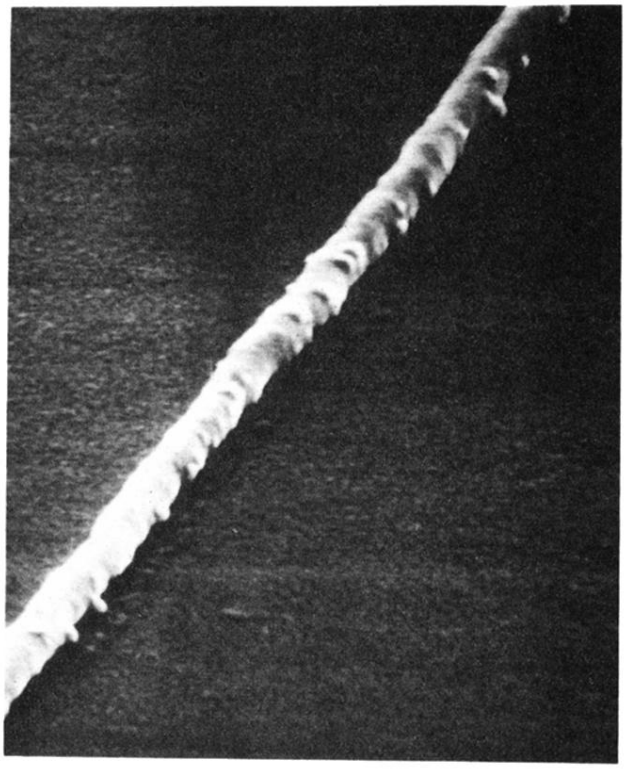

FIG. 7. Scanning electron micrograph of tungsten wire.

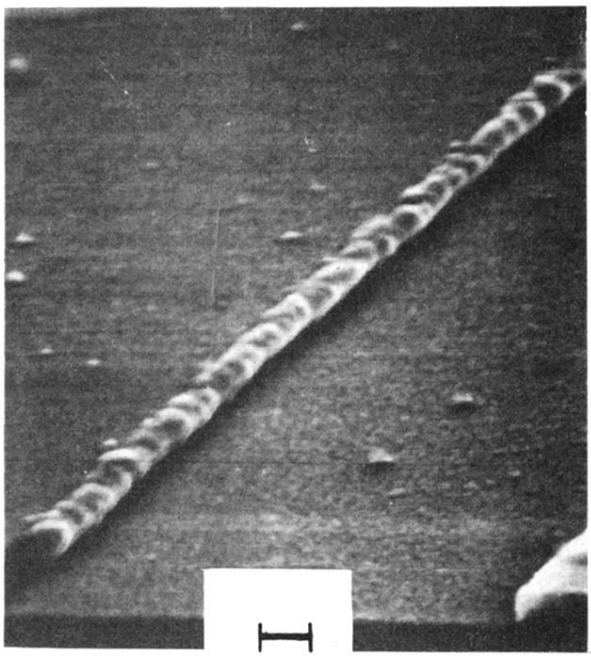

FIG. 8. Scanning electron micrograph of aluminum wire.