Resonant Tunneling through Single Electronic States and Its Suppression in a Magnetic Field

T. E. Kopley, P. L. McEuen, and R. G. Wheeler

Section of Applied Physics, P.O. Box 2157, Yale University, New Haven, Connecticut 06520

(Received 20 May 1988)

We observe resonant tunneling through single electronic states in a novel semiconductor device. At low temperatures, Lorentzian peaks of order but less than e^2/h in magnitude are observed in the conductance versus Fermi energy of the device. As the temperature is increased, the peaks change in amplitude and width consistent with thermal broadening of the energy distribution of the electrons. An external magnetic field monotonically decreases the amplitude and reduces the widths of the peaks in accordance with a model based on isolated resonant wells.

PACS numbers: 73.40.Qv, 73.40.Gk

Resonant tunneling through a single energy level in a barrier has been a subject of considerable interest and controversy since the advent of quantum mechanics. Recently workers have attributed experimental results to resonant tunneling through single states in such diverse systems as quantum dots,[1] metal-oxide-semiconductor field-effect transistors (MOSFET's),[2,3] metal-insulator-metal tunnel junctions,[4] and narrow quantum Hall-effect devices.[5] Stimulated by these results, theoretical investigators have extended solutions of the one-dimensional resonant-tunneling problem to the more realistic case of a single state connecting higher-dimensional reservoirs.[6,7] These studies predict a Lorentzian peak in the conductance versus energy of maximum amplitude e^2/h centered around the energy of the resonant state. In addition, simple arguments presented here predict that if the resonant state is associated with a single isolated potential well, a magnetic field will decrease the amplitude and width of the corresponding conductance peak. Previous experiments have observed neither this magnetic field dependence nor the expected Lorentzian shapes. We have observed these characteristic signatures of isolated-state resonant tunneling in a novel MOSFET device.

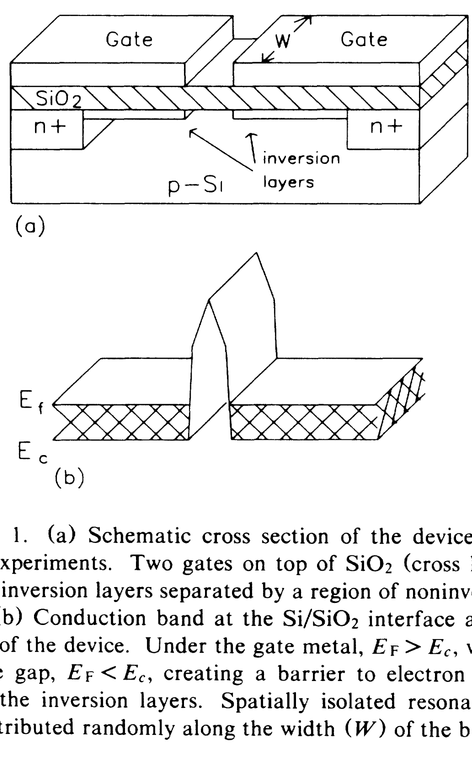

FIG. 1. (a) Schematic cross section of the device used in these experiments. Two gates on top of SiO_2 (cross hatched) induce inversion layers separated by a region of noninverted silicon. (b) Conduction band at the Si/SiO_2 interface along the length of the device. Under the gate metal, $E_F > E_c$, while under the gap, $E_F < E_c$, creating a barrier to electron flow between the inversion layers. Spatially isolated resonant states are distributed randomly along the width (W) of the barrier.

The tunneling system used in this study consists of two silicon inversion layers separated by a narrow noninverted region formed by patterning a small gap (~ 0.2 μm) in the gate metal of a MOSFET [see Fig. 1(a)]. The MOS structure consists of a 150-Å-thick aluminum gate on a 240-Å-thick oxide thermally grown on a (100) silicon wafer doped with $\simeq 10^{15}$-cm^{-3} boron. The aluminum gate metal is patterned with use of a hybrid deep-ultraviolet–electron-beam lithography process.[8] The conduction-band bending associated with the split gate geometry is shown schematically in Fig. 1(b). Conduction between inversion layers must occur through the noninverted silicon barrier, a region where the Fermi level lies below the conduction-band edge. The thickness and height of this barrier decrease with increasing gate voltage because of increasing fringing fields in the gap region.

In these experiments, the conductance (G) as a function of gate voltage (V_g) is measured at a fixed temperature and magnetic field (oriented perpendicular to the plane of the device) with use of a lock-in amplifier with a measurement voltage much less than kT/e. Since V_g is related to the Fermi energy (E_F) of the inversion layers ($dE_F/dV_g \simeq 5$ meV/V for the devices reported here), G is measured as a function of E_F of the inversion layers. Figure 2 is a plot of G vs V_g for device B5 at $T=0.5$ K. The device begins to conduct at $V_g \sim 7$ V while the inversion layers, which can be probed independently, begin to conduct at $V_g \sim 0.25$ V. This delayed turn-on behavior due to the barrier is also observed at room temperature. Structure in G vs V_g begins to appear at $T \sim 5$ K and becomes more pronounced as the temperature is decreased.

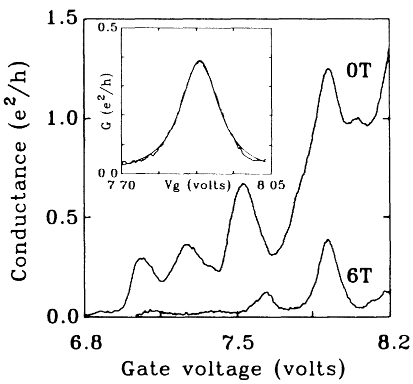

FIG. 2. Conductance vs gate voltage at $T=0.5$ K of device B5 for $B=0$ and 6 T. The 6-T data are shifted down in V_g by 0.04 V to line up the largest peak. [The positions of peaks in V_g varied by small ($\sim$0.1 V) amounts between gate voltage scans due to "switching" events (Ref. 9) that change the local potentials of the resonant states.] Inset: Largest peak at 6 T and a Lorentzian fit with $\Gamma_0=1.30$ meV (0.118 V in V_g units).

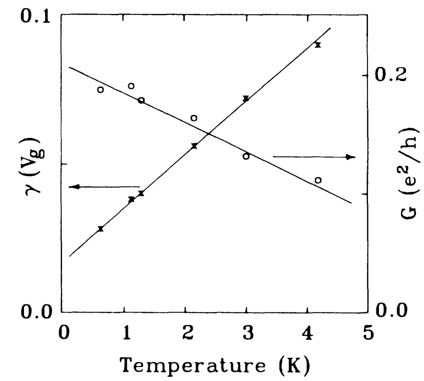

FIG. 3. Temperature dependence of the height and the FWHM γ of an isolated peak from device B6. The inferred zero-temperature width is $\Gamma_0=0.25$ meV. The line through the peak-height data is a guide to the eye.

We attribute this structure to resonant tunneling through spatially isolated states in the noninverted silicon barrier. For the measured range of device widths (1 μm $< W < 80$ μm), the structure is qualitatively independent of W and there is no direct correlation between W and the number of peaks per unit V_g. This suggests that the barrier is nonuniform and the observed conduction is through resonant levels in regions where the barrier is lowest.

Near a resonant state, the energy-dependent transmission coefficient of a barrier is a Lorentzian centered around the resonance energy E_0 (Ref. 10),

$$T(E)=\eta\Gamma_L\Gamma_R/[(E-E_0)^2+(\Gamma_L/2+\Gamma_R/2)^2].$$

η is a factor less than unity associated with reflections at the barrier boundaries,[11] E is the energy of the incident electrons, and $\Gamma_{L(R)}$ is the decay width, proportional to the leak from the resonant state to the left (right) contact. The transmission coefficient is related to the zero-temperature conductance of the system by the single-channel two-terminal Landauer formula[12,13]: $G(E_F)=(e^2/h)T(E_F)$. The validity of the single-channel Landauer formula for the case of a single state connecting higher-dimensional reservoirs has been theoretically investigated in the resonant-tunneling studies by Kalmeyer and Laughlin[6] and Xue and Lee.[7] Those studies, as well as the results presented here, indicate that the number of resonant states in the barrier determines the number of channels for the system independently of the properties of the reservoirs.

From the above equations, the peak height at resonance is given by

$$G(E_F=E_0)=4\eta(e^2/h)\Gamma_L/\Gamma_R, \quad \text{for } \Gamma_R \gg \Gamma_L,$$
$$G(E_F=E_0)=\eta(e^2/h), \quad \text{for } \Gamma_R=\Gamma_L.$$

Thus at zero temperature a conductance peak less than or equal to e^2/h in magnitude is expected for each resonant state. The data shown in Fig. 2 appears to be a superposition of such peaks centered at different energies. The application of a magnetic field (whose effect will be described later) isolates one of the peaks. Below $T\sim1$ K, this conductance peak is temperature independent and hence represents the intrinsic shape. This intrinsic shape is Lorentzian, as illustrated in the inset of Fig. 2.

As the temperature is increased, isolated peaks decrease in amplitude and increase in width as illustrated in Fig. 3 for a peak from device B6. This temperature dependence can be explained by the thermal broadening of the energy distributions of the tunneling electrons in the inversion layers without appeal to any inelastic mechanisms that would affect the tunneling process itself. In this model, the finite-temperature conductance is given by the convolution of the zero-temperature conductance with the energy derivative of the Fermi distribution function.[14] To a first approximation, Γ, the full width at half maximum (FWHM) of such a thermally broadened peak, is the sum of the width of the zero-temperature peak, $\Gamma_0=\Gamma_L+\Gamma_R$, and the width of $df(E,T)/dE=3.5k_BT$. In Fig. 3, γ, the experimentally observed width in terms of V_g, is plotted versus temperature. The slope $d\gamma/dT$ combined with the theoretical prediction of $d\Gamma/dT=3.5k$ gives a relation between changes in V_g and changes in E_F-E_0:

$$d(E_F-E_0)/dV_g=(d\Gamma/dT)/(d\gamma/dT)=16.8 \text{ meV/V}.$$

With use of this conversion factor, the y intercept gives the approximate zero-temperature width $\Gamma_0=0.25$ meV. The conversion factor is larger than $dE_F/dV_g=5$ meV/V inferred from capacitance measurements and Shubnikov–de Haas oscillations, indicating that the energy of the resonant level E_0 is decreasing as a function of V_g as $dE_0/dV_g=-11.8$ meV/V. This behavior qualitatively

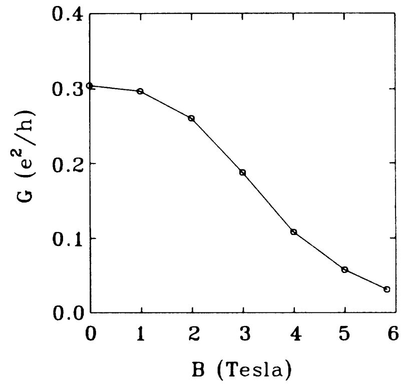

FIG. 4. Peak height vs magnetic field at $T=0.5$ K for the first peak from device B5.

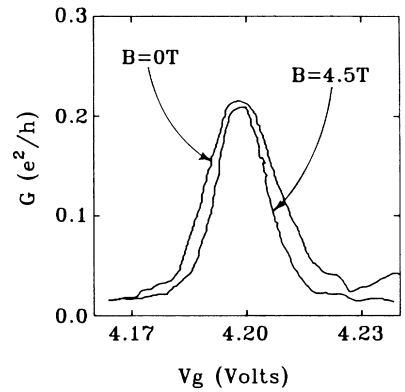

FIG. 5. Conductance vs gate voltage at $T=0.5$ K for an isolated peak from device B6 at $B=0$ and 4.5 T. The peak amplitude is unaffected by the field but the width is significantly reduced. The $B=4.5$-T data are shifted down in V_g by 0.017 V.

agrees with computer simulations of the electrostatics of a similar device geometry.[15] These simulations predict a decreasing barrier height with increasing E_F, an effect of increasing fringing fields in the gap area. This causes resonant levels in the barrier at a fixed energy below the conduction-band edge to decrease with increasing E_F.

The effect of a magnetic field upon the conductance peaks is the most striking evidence for a resonant-tunneling process through an isolated state. We postulate that the field decreases Γ_L and Γ_R by reducing the overlap of the resonant-state wave function with the electron wave function in the inversion layers. At zero field, $\Gamma_{L(R)} \sim \hbar\omega_0 \exp(-2\Lambda_{L(R)}/\xi)$, where ω_0 is the attempt frequency, $\Lambda_{L(R)}$ is the distance from the resonant state to the left (right) inversion layer, and ξ is the decay length. In a magnetic field, the resonant wave-function amplitude will decay at large distances as $\exp(-r^2/4l_B^2)$, and $\Gamma_{L(R)}$ will begin to be reduced when $4l_B^2 \sim \Lambda_{L(R)}\xi$, where $l_B = (\hbar/eB)^{1/2}$.

If an asymmetry $\Gamma_R > \Gamma_L$ exists because $\Lambda_R < \Lambda_L$, then Γ_L will decrease faster than Γ_R with increasing field because the effect of the magnetic field on the wave function increases with radial distance from the resonant site. Since G is proportional to Γ_L/Γ_R, the on-resonance conductance will decrease. This suppression of the conductance peaks by a magnetic field is shown in Figs. 2 and 4 for device B5. More striking is the field dependence of the isolated peak from device B6 shown in Fig. 5. The amplitude of the peak is *unaffected* by fields up to 6 T, implying that the ratio Γ_L/Γ_R is field independent and hence the resonant level is spatially centered between the inversion layers, i.e., $\Lambda_L = \Lambda_R$. The width of the peak is significantly reduced, however, since Γ_L and Γ_R are reduced. This definitively illustrates the resonant nature of the tunneling process: The transmission rates in and out of the resonant state decrease without decreasing the on-resonance conductance. The data also show that there are no observable spin effects on the peak shape or amplitude although the Zeeman energy at 4.5 T of ~ 0.5 meV is larger than $\Gamma = 0.29$ meV.

For the isolated peak of device B6, the magnetic field data coupled with the zero-field width can be used to estimate the parameters Λ ($=\Lambda_L = \Lambda_R$) and ξ. Assuming a reasonable value for $\hbar\omega_0$ (25 meV), $\Lambda/\xi \sim 2.6$ from the zero-field resonance width and $4l_B^2 \sim \Lambda\xi$ at $B \simeq 6$ T from the magnetic field data. Combining these gives $\Lambda \sim 340$ Å and $\xi \sim 130$ Å. The discrepancy between the inferred barrier length, 2Λ, and the lithographic dimensions of the gate is likely due to the spread of the inversion layers beyond the edges of the gate because of fringing fields.

In all the devices studied, peak heights decrease with increasing magnetic field or are field independent. This is in sharp contrast to the resonant-tunneling studies of Wainer, Fowler, and Webb[16] in short-channel MOSFET's, where the resonant peaks are due to disorder-induced localized states in the conduction-band tail. These peaks can either increase or decrease in amplitude when a magnetic field is applied. Numerical simulations of resonant tunneling through a disordered barrier by Xue and Lee[7] also demonstrate this random field dependence. We believe the difference in the two cases lies in the nature of the resonant states, and that a single-well model, not a disordered (i.e., multiwell) model, is applicable for our devices. We postulate that these single wells are spatially isolated Coulombic sites in the noninverted silicon barrier region with bound-state energies lying in the band gap of the silicon. These may be mobile ionized impurities such as Na^+ located at the Si/SiO_2 interface. The number of resonant peaks we observe is consistent with the calculated number of mobile ions (~ 10 for $W = 1$ μm) within a decay length of the center of the barrier. However, the value of ξ (~ 130 Å) inferred is larger than the calculated decay length (~ 50 Å) of an electron bound to a single charge at the Si/SiO_2 interface.[17] This calculation does not include the effect of the electric fields associated with our split gate geometry on the Coulombic confining potential. These fields lower the barrier and increase the decay length in the direction of transport.

In conclusion, we have fabricated a novel system that clearly exhibits the characteristics of resonant tunneling through single states. Peaks in G vs E_F are observed to be Lorentzian for $kT \ll \Gamma_0$, and their temperature dependence can be described by the thermal broadening of the electron distributions in the inversion layers. A magnetic field reduces the amplitude and width of these peaks in accordance with simple predictions based upon a single resonant well.

We acknowledge helpful discussions with D. E. Prober and A. D. Stone and thank P. A. Lee for providing preprints and P. D. Dresselhaus for assisting with the low-temperature measurements. This work was supported by the National Science Foundation under Grants No. ECS-8509135 and No. DMR-8213080. The MOSFET mask set was fabricated at the National Nanofabrication Facility, which is supported by the National Science Foundation under Grant No. ECS-8619049, Cornell University, and industrial affiliates. One of us (P.L.M.) was partially supported by IBM.

[1]M. A. Reed, J. N. Randall, R. J. Aggarwal, R. J. Matyi, T. M. Moore, and A. E. Wetsel, Phys. Rev. Lett. **60**, 535 (1988).

[2]R. H. Koch and A. Hartstein, Phys. Rev. Lett. **54**, 1848 (1985).

[3]A. B. Fowler, G. L. Timp, J. J. Wainer, and R. A. Webb, Phys. Rev. Lett. **57**, 138 (1986).

[4]S. J. Bending and M. R. Beasley, Phys. Rev. Lett. **55**, 324 (1985), and references therein.

[5]J. K. Jain and S. A. Kivelson, Phys. Rev. Lett. **60**, 1542 (1988).

[6]V. Kalmeyer and R. B. Laughlin, Phys. Rev. B **35**, 9805 (1987).

[7]W. Xue and P. A. Lee, to be published.

[8]M. J. Rooks, S. Wind, P. McEuen, and D. E. Prober, J. Vac. Sci. Technol. B **5**, 318 (1987).

[9]K. S. Ralls, W. J. Skocpol, L. D. Jackel, R. E. Howard, L. A. Fetter, R. W. Epworth, and D. M. Tennant, Phys. Rev. Lett. **52**, 228 (1984).

[10]M. Ya. Azbel, Phys. Rev. B **28**, 4106 (1983).

[11]E. O. Kane, in *Tunneling Phenomena in Solids,* edited by E. Burstein and S. Lundqvist (Plenum, New York, 1969), p. 1.

[12]R. Landauer, Philos. Mag. **21**, 863 (1970).

[13]Y. Imry, in *Directions in Condensed Matter Physics,* edited by G. Grinstein and G. Mazenko (World Scientific, Singapore, 1986), p. 101.

[14]H. L. Engquist and P. W. Anderson, Phys. Rev. B **24**, 1151 (1981).

[15]A. C. Warren, Dimitri A. Antoniadis, and Henry I. Smith, IEEE Electron. Device Lett. **7**, 413 (1986).

[16]J. J. Wainer, A. B. Fowler, and R. A. Webb, Bull. Am. Phys. Soc. **33**, 702 (1988).

[17]B. G. Martin and R. F. Wallis, Phys. Rev. B **18**, 5644 (1978).