NANOFABRICATION

Building complex semiconductor structures with features as small as 100 atoms across brings quantum effects to life but challenges fabrication technology.

Henry I. Smith and Harold G. Craighead

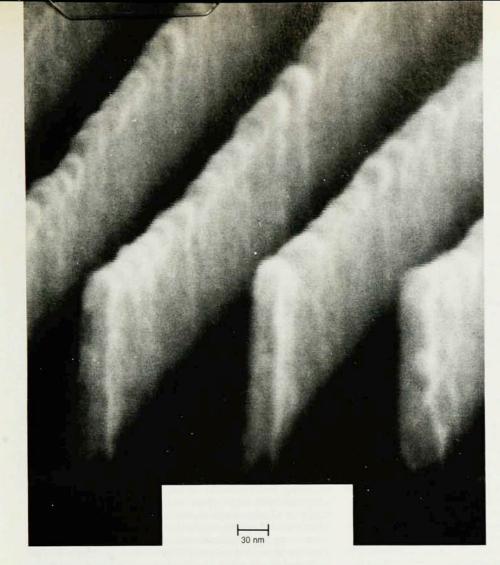
A decade ago the fabrication and study of electron devices whose smallest features were just under 1 micron in size represented the forefront of the field. Today that forefront has moved down an order of magnitude to 100 nanometers, engendering new terminology based on the prefix *nano*, from the Greek word for dwarf: "nanoscale devices," "nanolithography," "nanofabrication."

We can distinguish two classes of nanoscale devices: scaled-down conventional devices and devices based on quantum effects. Field-effect transistors with channel lengths under 100 nm are examples of the former. The goal in building such short-channel devices is usually shorter transit time and hence higher operating frequency. Also, as the transit time becomes comparable to certain relaxation times, new modes of operation become feasible. Scaled conventional devices are generally operated in regimes of temperature and electric field for which the electron coherence length is smaller than the minimum device dimensions. This allows one to treat charge carriers (electrons or holes) as classical particles.

In quantum-effect devices, on the other hand, the wave nature of the electron is fundamental to device performance or to the phenomena under study. Thus, one or more device dimensions are usually comparable to the de Broglie wavelength, which is typically much smaller than the phase-breaking length (that is, the coherence length). Because quantum phenomena become more pronounced and energy level separations increase as dimensions are reduced, such devices "improve" as their structures are shrunk, for example, from 100 to 10 nm. In contrast, in scaled conventional devices the benefits achieved as dimensions are shrunk (for example, shorter transit times) tend to be offset eventually by losses due to parasitic effects. Roughly speaking, scaled conventional devices begin to encounter diminishing returns as dimensions shrink below about 100 nm. From the point of view of fabrication technology, therefore, quantum-effect devices represent a greater challenge, calling upon us to develop precise and reliable methods for making complex semiconductor structures with minimum features approaching 10 nm, less than 100 atoms wide. (See figure 1.)

The major challenges in fabricating both scaled conventional and quantum-effect devices are growing the substrate materials and patterning the structures. The growth of high-purity semiconductor materials and the engineering of the band structures have been discussed

Henry Smith is a professor of electrical engineering at the Massachusetts Institute of Technology, in Cambridge. **Harold Craighead** is director of the National Nanofabrication Facility, at Cornell University, in Ithaca, New York.



Nanostructure consisting of lines less than 200 atoms wide. In x-ray nanolithography, backscattering from the substrate is absent, and dense patterns of sub-100nm features can be readily achieved. Here we see 30nm-wide lines of PMMA resist material that was masked during x-ray exposure (see figure 2). The mask's 200-nm-period grating pattern, which covered an area of about 1 cm², was produced by gold liftoff following exposure of the photoresist to the interference pattern from two 350-nm-wavelength laser beams intersecting at 122°. Figure 1

before in Physics today (see Venkatesh Narayanamurti's article on crystalline semiconductor heterostructures, October 1984, page 24, and the May 1985 special issue on optoelectronics). In this article we focus on nanostructure patterning, commonly known as nanofabrication, and illustrate the current state of the art with a variety of examples.

Smaller electronic devices

Nanofabrication is most frequently done with the planar process illustrated in figure 2. This process, which is applicable down to feature sizes approaching 10 nm, consists of two steps: lithography and pattern transfer. In lithography, a radiation-sensitive film known as the resist is exposed to some pattern of radiation, which can be electrons, ions or photons. The radiation alters the resist, enabling one to develop the pattern as a relief image in the resist. This image is then transferred into the substrate by etching, growth, doping or liftoff, as indicated in the figure. The critical issues include achieving resolution and accuracy in the pattern definition step, maintaining fidelity in the pattern transfer step and controlling damage in the small structures.

In scanning-electron-beam lithography a small electron source is imaged on a substrate through a series of electromagnetic lenses. The technique is highly developed and is widely used commercially in generating photolitho-

graphic masks. In nanostructure research it is the most widely used and versatile lithography tool. Because of the availability of high-quality electron sources and optics, electron beams can be focused to diameters less than 2 nm and accurately scanned to define patterns of arbitrary geometry. Electron beams focused down to about 0.5 nm have produced niobium lines as fine as 8 nm wide on thin carbon foils and holes as small as 1 nm in ionic salts. However, such demonstrations have not yet led to useful techniques for fabricating nanoscale electronic devices, and so we will not discuss them further here.

Figures 3 and 4 depict two kinds of nanoscale electronic devices in which scanning-electron-beam lithography was used to define the gate electrodes, the finest features in the devices. Figure 3 shows a metal-oxide-semiconductor field-effect transistor, or MOSFET, similar in structure and function to those found in commercial integrated circuits except that here the length of the polysilicon gate is 100 nm rather than the 1000 nm typical of commercial circuits. Switching times of 13.1 psec have been observed in such devices, as has velocity overshoot. (For a discussion of velocity overshoot and modulation doping see Karl Hess's article on page 34.) Short-channel field-effect transistors built on AllnAs/GaInAs modulation-doped material have achieved switching times as short as 4 picoseconds with gates 120 nm long.

The electron micrograph in figure 4a is a top view of a

quantum-effect device that uses resonant tunneling.4 The two parallel lines through the middle of the micrograph are metal electrodes, or gates, which produce potential barriers to electrons in the two-dimensional electron gas that exists between the source and the drain. As the gate voltage is raised, electrons from the source side can tunnel to the drain side through the states in the well between the electrodes. This is evident in the electrical characteristics shown in figure 4b. (See also Federico Capasso and Supriyo Datta's article on page 74.) In most resonanttunneling devices, transport is perpendicular to the plane of the layers, rather than in the plane as here. An important aspect of this and other quantum-effect devices is that both the strength of the effects and the temperature at which they can be observed increase as the electrode size decreases. Thus there is a strong motivation to develop means to push the linewidths well below the 60 nm of the device shown in figure 4.

In an experiment on electron-beam-defined Al/Al₂O₃/ Al tunnel junctions using wires as narrow as 50 nm, Theodore Fulton and Gerald Dolan observed large oscillations in current-voltage curves and attributed them to

single-electron charging events.5

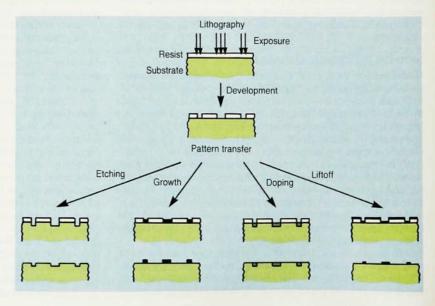
Experimenters recently used electron-beam lithography and reactive ion etching to create metallic "nanoconstrictions" less than 10 nm in diameter.6 They did this by evaporating metal onto two sides of a Si₃N₄ membrane into which a small hole had been etched. A number of interesting effects, including two-level fluctuations and the evolution of 1/f noise, were observed. These nanoconstrictions are some of the smallest structures ever studied by transport measurements.

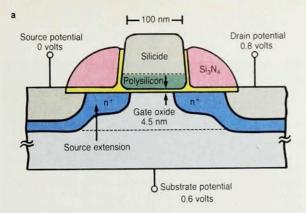
Overcoming scattering

One problem with electron-beam lithography is that a large fraction of the incident electrons, after passing through the resist, are scattered back through it. Because the range of these backscattered electrons is much larger than typical feature sizes and separations, they create a background exposure that limits the tolerable exposure range and causes resist exposure parameters to depend on the atomic number of substrate material and the proximity of other pattern features. To minimize such effects in sub-100-nm lithography, some researchers have gone to very high voltages, above 100 keV. At such voltages the range of most backscattered electrons is many micrometers, so that the background exposure is spread out and relatively uniform with respect to fine features. Other researchers have used thin (10-100 nm) membranes as substrates to greatly reduce the deleterious backscattering. In fact, the finest features obtained by electron-beam lithography, about 2 nm, have been on such thin membranes. However, for nanoscale semiconductor devices, membrane substrates are of very limited utility. Another strategy is to employ very low voltages, less than 1 keV, so that the range of scattered electrons is under 100 nm.

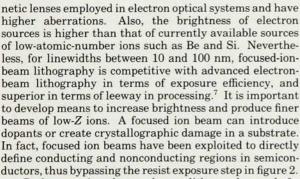
Focused-ion-beam lithography circumvents most of the problems of backscattering and overlapping exposure from nearby features. When an ion passes through a resist film it loses energy in a quasicontinuous fashion, exposing the resist by exciting low-energy secondary electrons that have short ranges of about 5 nm. The majority of large-angle scattering events occur at the end of an ion's trajectory, where the energy is low and little exposure occurs. Focused-ion-beam lithography with Be2+ ions has been used to create 250-nm-thick gold lines 50 nm wide and 50 nm apart,7 and lines as narrow as 12 nm have been obtained in thin PMMA resist using a beam of gallium ions.8 (Polymethyl methacrylate is a polymer that undergoes chain scission with ionizing radiation.) Focused-ion-beam lithography requires electrostatic lenses, which are not as well developed as the electromag-

The planar process is the most effective means of producing a relief structure, a pattern of selective doping or a pattern of one material on top of another. In the lithography step, a radiation-sensitive film known as the resist is exposed to some pattern of radiation, which can be electrons, ions or photons. The radiation alters the resist, which is usually a polymer, enabling the pattern to be developed as a relief image in the resist. This image can then be transferred to the substrate by etching, growth, doping or liftoff. Figure 2





Mosfet schematic diagram (a) and micrograph (b). The cross-sectional schematic is of a silicon transistor with a 100-nm channel length, part of a ring oscillator circuit fabricated using electron-beam lithography. To make the electron micrograph, the contact metal and oxide were removed. The two structures that go diagonally across the image from upper left to lower right are the source and drain contacts; the gate is in the gap between them. The switching times of such short-channel silicon transistors are competitive with those of field-effect transistors made with III–V semiconductors. (Micrograph courtesy of T. H. Philip Chang and Dieter P. Kern, IBM.) Figure 3



Both scanning-electron-beam lithography and focused-ion-beam lithography expose patterns in a sequential, dot-by-dot fashion, allowing great flexibility in creating patterns of arbitary geometry. Controlling a high-energy, charged beam to achieve absolute accuracy and pattern coherence over large areas requires sophisticated particle optics and computer control. For minimum linewidths of 50 nm, the maximum writing field is only about 100 microns. Patterns covering larger areas are done as mosaics of small fields. This requires precise mechanical motion of the sample stage, tracked by laser interferometry. Discontinuities in lines that span the boundaries between two such fields must be small compared with the minimum linewidth. With sufficient care the most advanced instruments can reduce discontinuities to levels approaching 10 nm.

Exposure time

Because exposure is a sequential process with both ion and electron beams, the exposure time depends on the beam current, the sensitivity of the resist and the pattern area. To write 50-nm lines would require an electron beam of about 10 nm diameter. The beam current would range between about 10 and 200 picoamps, depending on the type of source (either thermionic or field emission) and the



electron optics. The resist most widely used for nanometer lithography, PMMA, requires about 10 microcoulombs per square centimeter, which corresponds to just 63 electrons per $10\times 10\text{-nm}$ area. Increasing the resist sensitivity by more than about a factor of 10 would lead to problems of statistical noise. Thus we can assume an electron sensitivity of 1–10 $\mu\text{C/cm}^2$ for 50-nm linewidths. Knowing the area of a pattern and the beam current, one can easily calculate the minimum writing time.

Some research problems encounter the difficulty of long exposure time; this will be a problem for future practical applications as well. One way of addressing this difficulty is to increase the electron-beam current by several orders of magnitude while keeping the energy low-under 1 keV-to avoid the detrimental effects of electron backscattering. T. H. Philip Chang and his coworkers at IBM are using this approach.10 They employ field-emission tips and electrostatic lenses having bore diameters of about 1 micron. Their calculations indicate that in a 2-keV beam 5 nm in diameter they should be able to get a current of about a microamp, which is about 5 orders of magnitude higher than what conventional 2-keV field-emission systems provide. Their plans call for overcoming the well-known problem of unstable field emission by piezoelectrically adjusting the tip position to maintain constant current, an idea borrowed from the technology of scanning tunneling microscopy.

Another approach to the problem of long exposure time is to expose patterns in parallel by passing radiation through a mask, which can take considerably less time than scanning a beam. Either a shadow image of the mask (proximity printing) or a focused image of the mask (projection printing) can expose the resist. Many schemes for doing this with electrons, ions and x rays have been demonstrated or proposed. For nanolithography—that is, lithography producing sub-100-nm features—the use of x-ray shadow printing has proven the most successful, as figure 1 illustrates. For such "x-ray nanolithography" the wavelength should be either about 1.3 nm or 4.5 nm, where

the 1/e absorption depth of x rays in a polymeric resist is about 1 micron. In addition to the benefits of parallel exposure (such as the ability to expose mask fields of a few square centimeters), back reflection and scattering of x rays are negligible. As a result, image contrast is high and processing does not need to be as tightly controlled.

With synchrotron sources of soft x rays, exposure times may be a few seconds, whereas with conventional electron-bombardment sources exposure times can be one to a few hours. Sources based on laser-produced plasmas are under development, and these promise exposure times within a factor of 10 of what synchrotrons offer.

The patterning of an x-ray mask can be done by various combinations of electron-beam, ion-beam and photon-based lithography. The durability, alignment and distortion of masks remain issues for nanolithography. Schemes for aligning masks to better than 10 nm appear feasible.¹¹

Quantum effects

Figure 5a shows a cross section of a lateral-surface superlattice device, and figure 5b shows the metal grid gate produced by x-ray nanolithography and liftoff. 12 A small potential is applied between the source and the drain. As the potential on the gate is raised, the electron Fermi wavelength decreases. Back diffraction from the periodic potential occurs whenever an integral number of half wavelengths is equal to the period of the grid. This is manifest in the plot of conductance as a function of gate voltage in figure 5c. Alternatively, one can say that the artificial grid superlattice breaks the normal band structure into minibands and minigaps, provided the electron coherence length is larger than a few periods of the grating. Each region of negative slope (that is, negative differential transconductance) corresponds to the Fermi level passing through a minigap. (See Capasso and Datta's

The gate potential also controls the depth of modulation. In fact, at sufficiently negative gate voltage the two-dimensional electron gas is depleted under the gate metal, leaving an array of isolated, parabolic potential wells, or "quantum dots," each containing only a few electrons. The structure shown in figure 5 yields 2.5 billion quantum dots per square centimeter. Using x-ray nanolithography it is possible to produce such arrays over areas of several square centimeters. The electronic states of these quantum dots have been probed by far-infrared magnetoabsorption¹³ and magnetocapacitance, ¹⁴ techniques that make use of the dependence of absorption and capacitance on the magnetic field. Each of these measurements required areas of several square millimeters.

The 200-nm-period grid on the x-ray mask used to produce the device shown in figure 5 was patterned using ultraviolet holographic lithography, which ensures pattern coherence over large areas and requires only a few seconds of exposure. Jorg Peter Kotthaus and his coworkers at the University of Hamburg, in West Germany, have used holographic lithography to make grating structures with periods of a few hundred nanometers directly on semiconductor substrates and have studied the optical and

transport properties of the electrons confined in the resulting quantum wires. ¹⁵ Holographic lithography using the deep-uv line of ArF at 193 nm yields a grating of 100-nm period (50-nm lines and spaces) ¹⁶ and hence should ultimately give quantum dots of about 25-nm diameter, with a density of 10 billion per square centimeter!

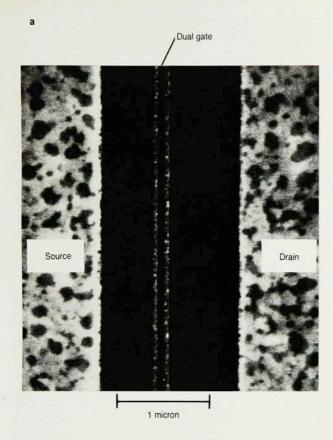
Mark Reed and his coworkers at Texas Instruments in Dallas have fabricated arrays of quantum dots by etching columns less than 100 nm in diameter down through a two-dimensional electron gas in a multilayer compound semiconductor.¹⁷ However, because the etching process involves ion bombardment, crystallographic damage on the sidewalls and surfaces presents serious problems to this approach. Arrays of quantum dots interacting in a predictable way with their near neighbors may someday find applications in massively parallel computation. Schemes based on cellular automata architecture have been considered.¹⁸

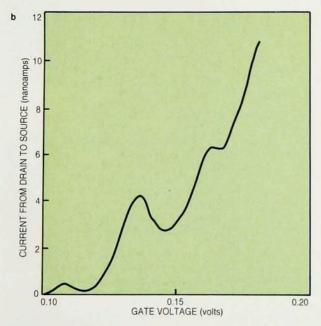
The quantum-effect devices discussed above all use lithography and liftoff to define metal electrodes, and a field effect to produce the potential that confines electrons. Often the confinement is narrower than the lithography, a highly beneficial feature for quantum-effect devices. Many researchers have produced one-dimensional and zero-dimensional quantum confinement by etching relief structures into layered III–V semiconductors or by selectively damaging the surrounding area with ion bombardment. An electrode-controlled field-effect scheme is easier to fabricate and naturally lends itself to fundamental studies because one can vary the confining potentials with external voltage sources.

Recent experiments on quantized ballistic conductance involve noteworthy examples of field-effect confinement.19 (See the news story in Physics today, November 1988, page 21.) The separation of the electrodes in these experiments was about 270 nm. However, by applying a negative bias to them one could pinch down and even cut off the conducting channel. The conductance decreased in steps of $2e^2/h$ as the channel was narrowed. The channel was not only narrow but also short, so that electrons could travel through it ballistically-that is, without any scattering. The allowed modes in the channel are then the "waveguide" modes, which are quantized. The conductance is also quantized, increasing by $2e^2/h$ whenever the Fermi energy is raised to the point where the next-higherorder waveguide mode is allowed. The magnitude of the steps, $2e^2/h$, can be obtained from the uncertainty relation $\Delta E \Delta t \sim \hbar$ and the Pauli exclusion principle, which dictates that only two electrons (one for each spin state) can occupy any given waveguide mode. The detection of current (that is, $e/\Delta t$) implies that an electron has scattered inelastically inside the output reservoir (so that $\Delta E = Ve$), thereby vacating a waveguide mode and enabling another electron from the input reservoir to occupy it

Gregory Timp and his coworkers at AT&T Bell Laboratories have explored ballistic electron waveguide junctions and bends,²⁰ and applications analogous to microwave devices have been proposed.

If electron waveguides are made much longer than the electron mean free path, transport is no longer





ballistic, but quantum effects are still observable. For example, at MIT we have fabricated arrays of about 100 parallel electron waveguides, or "quantum wires," in modulation-doped AlGaAs/GaAs by ion etching a 200-nm-period grating into the AlGaAs. Surface states on the etched sidewalls caused depletion of the two-dimensional electron gas, resulting in channels about 30 nm wide. As we swept the Fermi wavelength by means of a gate on the back surface of the gallium arsenide sub-

Resonant-tunneling transistor in a planar configuration on AlGaAs/GaAs. a: Scanning electron micrograph showing a top view of the source, the drain and two Schottky metal gate electrodes, each 60 nm wide, separated by 60 nm. By modulation doping, a high-mobility two-dimensional electron gas is created that connects source and drain. The electrodes, or gates, produce potential barriers to the free flow of electrons from source to drain. As the gate voltage is raised, electrons from the source side can tunnel to the drain side by way of states in the potential well between the two gates. The Ti/Pt gate electrodes were produced by liftoff following electron-beam lithography. The lithography was done at the National Nanofabrication Facility, at Cornell University. b: Plot of drain-to-source current as a function of gate voltage. The three peaks correspond to resonant tunneling. Figure 4

strate, the conductance oscillated above and below the value for a scaled two-dimensional conductor. This occurred because the density of states in a waveguide has a discontinuity whenever the Fermi wavelength is reduced to the point where the next-higher-order lateral mode is allowed. The sudden increase in the number of available states into which an electron can scatter causes a drop in mobility. The influence of dimensionality (that is, whether the conductor is one dimensional or two dimensional) on conductance enables one to modulate mobility with an external electrode and may lead to future device applications.²²

The future

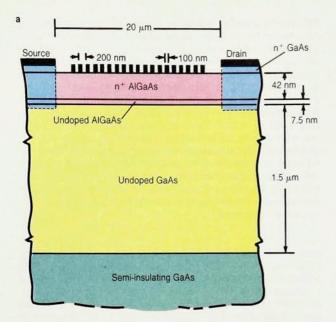
The challenges of fabricating nanoscale devices, both conventional and quantum effect, are formidable from the points of view of semiconductor material growth, electrical contacts, lithography and pattern transfer. No single lithography technique is a panacea. An optimal strategy is to employ electrons, ions or x rays wherever they are most effective in solving the problem at hand. Pattern transfer techniques are often a more serious limitation than nanolithography. For example, to achieve profile control, one must almost always use etching techniques that involve bombardment with ions having energies of a few hundred electron volts. Such bombardment causes damage to both semiconductor surfaces and deep into the bulk. We will need to develop new methods of damage-free etching compatible with linewidths under 100 nm.

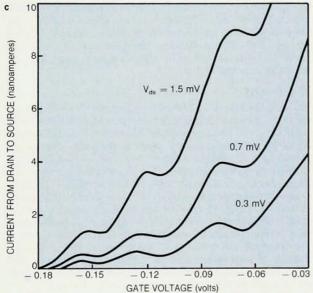
Studies of quantum effects have generally been done at 4 K or below. For practical applications in electronic systems it is highly desirable to work at 77 K, which necessitates using semiconductors with lower effective electron masses (InAs or InGaAs, for example) and minimum features closer to 10 nm than 50 nm. Also, it will likely be necessary to control linewidth variations to less than about 10% of the minimum linewidths so that quantum levels are predictable and repeatable. Such requirements represent enormous challenges for nanofabrication, which we are only beginning to address.

The work at MIT is supported by the Joint Services Electronics Program, the Air Force Office of Scientific Research and the National Science Foundation. The National Nanofabrication Facility at Cornell is supported by the National Science Foundation.

References

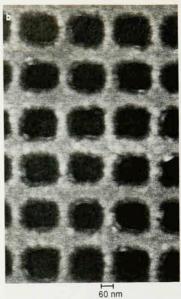
- A. N. Broers, IBM J. Res. Dev. 32, 502 (1988). M. Isaacson, A. Murray, J. Vac. Sci. Technol. 19, 1117 (1981).
- G. A. Sai-Halasz, M. R. Wordeman, D. P. Kern, S. Rishton, E. Ganin, IEEE Electron Dev. Lett. 9, 464, 633 (1988).
 G. Sha-





hidi, D. A. Antoniadis, H. I. Smith, IEEE Electron Dev. Lett. $\bf 9, 94~(1988).$

- U. K. Mishra, A. S. Brown, M. J. Delaney, P. T. Greiling, C. F. Krumm, IEEE Trans. Microwave Theory Tech. 37, 1279 (1989).
- K. Ismail, D. A. Antoniadis, H. I. Smith, Appl. Phys. Lett. 55, 589 (1989).
- 5. T. A. Fulton, G. J. Dolan, Phys. Rev. Lett. 59, 109 (1987).
- 6. K. Ralls, R. A. Buhrman, Phys. Rev. Lett. 60, 2434 (1988).
- W. Chu, A. Yen, K. Ismail, M. I. Shepard, H. J. Lezec, C. R. Musil, J. Melngailis, Y. C. Ku, J. M. Carter, H. I. Smith, J. Vac. Sci. Technol. B 7, 1583 (1989).
- R. L. Kubena, F. P. Stratton, J. W. Ward, G. M. Atkinson, R. J. Joyce, J. Vac. Sci. Technol. B 7, 1798 (1989).
- M. A. Gesley, F. J. Hohn, R. G. Viswanathan, A. D. Wilson, J. Vac. Sci. Technol. B 6, 2014 (1988).
- T. H. P. Chang, D. P. Kern, M. A. McCord, J. Vac. Sci. Technol. B 7, 1855 (1989).
- J. Itoh, T. Kanayama, N. Atoda, K. Hoh, J. Vac. Sci. Technol. B 6, 409 (1988). A. Moel, M. L. Schattenburg, J. M. Carter, H. I. Smith, J. Vac. Sci. Technol. B 7, 1692 (1989).



60 1111

Lateral-surface superlattice device.
a: Cross-sectional schematic showing the 200-nm-period Schottky-barrier grid gate and the grown semiconductor layers. b: Electron micrograph of the 200-nm-period grid gate, fabricated by x-ray nanolithography and liftoff. c: Current from drain to source as a function of gate voltage, for three values of drain-to-source voltage V_{ds}. The peaks indicate electron back diffraction, which occurs whenever an integral number of half wavelengths matches the period of

the grid. Figure 5

- K. Ismail, W. Chu, A. Yen, D. A. Antoniadis, H. I. Smith, Appl. Phys. Lett. 54, 460 (1989).
- C. T. Liu, K. Nakamura, D. C. Tsui, K. Ismail, D. A. Antoniadis, H. I. Smith, Appl. Phys. Lett. 55, 168 (1989).
- K. Ismail, T. P. Smith, W. T. Masselink, H. I. Smith, Appl. Phys. Lett. 55, 2766 (1989).
- 15. J. P. Kotthaus, Phys. Scr. T19, 120 (1987).
- A. Yen, R. Ghanbari, H. I. Smith, Microelectron. Eng. (1989), to be published.
- M. A. Reed, J. N. Randall, R. J. Aggarwall, R. J. Matyi, T. M. Moore, A. E. Wetsel, Phys. Rev. Lett. 60, 535 (1988).
- J. N. Randall, M. A. Reed, G. A. Frazier, J. Vac. Sci. Technol. B 7, 1398 (1989).
 N. Margolus, T. Toffoli, G. Vichniac, Phys. Rev. Lett. 56, 1694 (1986).
- B. J. van Wees, H. van Houten, C. W. J. Beenakker, J. G. Williamson, L. P. Kouwenhoven, D. van der Marcel, C. T. Foxon, Phys. Rev. Lett. 60, 848 (1988). D. A. Wharam, T. J. Thornton, R. Newbury, M. Pepper, H. Ahmed, J. E. F. Frost, D. G. Hasko, D. C. Peacock, D. A. Richie, G. A. C. Jones, J. Phys. C 21, L209 (1988).
- G. Timp, H. U. Baranger, P. de Vegvar, J. E. Cunningham, R. E. Howard, P. Behringer, P. M. Mankiewich, Phys. Rev. Lett. 60, 2081 (1988).
- K. Ismail, D. A. Antoniadis, H. I. Smith, Appl. Phys. Lett. 54, 1130 (1989).
- 22. H. Sakaki, Jpn. J. Appl. Phys. 19, L735 (1980).