The Quantum-Effect Device: Tomorrow's Transistor?

The components of ordinary integrated circuits can be made only so small before disruptive effects impair their function. Beyond that size limit a new species of semiconductor device could take over

by Robert T. Bate

The electronics industry and integrated circuits share an inverse destiny. The industry grows as circuits shrink, and growth will continue as long as more and more circuits can be crammed on a single chip. But common sense and careful analyses indicate that perhaps within a decade downscaling will run up against the limits of circuit technology. Even if practical limits are overcome, the physical laws that govern the behavior of circuit components set fundamental limits on the size of the components' features. In order to keep expanding, the electronics industry needs another technological revolution.

As a physicist with Texas Instruments, Incorporated, I have for many years been aware of the urgency of developing a new frontier for semiconductor devices. In 1982 my colleague Pallab K. Chatterjee published a study that heightened my concern by stressing how close the downscaling endpoint was. There is still some disagreement over that figure, with estimates of minimum feature sizes ranging between 100 and 500 billionths of a meter. While disputing the problem, many of us arrived at the same solution: that some of the very phenomena that impose size limits on ordinary circuits could be exploited in a new generation of vastly more efficient devices. The functional bases for these devices are quantum-mechanical effects that carry semiconductor technology into a realm of physics where subatomic particles behave like waves and pass through formerly impenetrable barriers. With the so-called quantum semiconductor device, I believe it will be possible to put the circuitry of a supercomputer on a single chip.

The structures for quantum devices have already been made using the same materials as today's chips: doped silicon, doped and undoped gallium arsenide, and aluminum gallium arsenide. Because they can be about 100 times smaller than the devices in present-day integrated circuits, however, designing and fabricating a viable device presents a formidable challenge. Manufacturing processes will have to become considerably more sophisticated, and new strategies for interconnection and architecture will have to be devised to cope with the special problems of size reduction.

As daunting as they are, these adjustments are worth making in order to realize the ten-thousandfold reduction in cost per function that quantum devices could bring about. They are also minor compared with the difficulty of introducing new materials for which no relevant process technology exists. And the progress that has been made at Texas Instruments as well as at other industry, government and academic laboratories around the world suggests that quantum devices just might embody the revolution the electronics industry awaits.

The motive for shrinking the components of integrated circuits is minimizing the cost and time needed to perform each circuit function. Most functions are carried out by transistors, which act essentially as switches. In a transistor the speed and precision with which switching can be controlled, as well as the power needed to produce the switching, has everything to do with the time and cost per function attained by the device. Because of its size, a transistor switch that operates on the principles of quantum mechanics would be faster and would consume less power than a conventional transistor; because of effects peculiar to quantum phenomena, it could also afford a greater degree of control.

These attributes can best be appreciated in comparison with the performance of conventional transistors. The most commonly used transistors today are field-effect transistors, or FET's. They are made from semiconducting materials doped with elements that provide carriers for electric charge. The charge carriers can be either electrons, which bear a negative charge, or positive "holes"; a semiconductor that has electrons as charge carriers is said to be negatively doped (*n*-doped) and a semiconductor that conveys charge by the movement of holes is said to be positively doped (*p*-doped). Silicon has been the traditional stuff of integrated circuits, but gallium arsenide (GaAs) transistors have been constructed that are faster.

The two types of transistor have slightly different configurations [*see illustration on page 98*]. In a typical silicon FET a region of *n*-doped silicon called the source is separated from another *n*-doped region, the drain, by a *p*-doped channel. On top of the channel there is a metal electrode called the gate, which is kept from coming in direct contact with the *p*-doped silicon by a layer of insulating silicon oxide. (This metal-oxide-semiconductor arrangement is the derivation for the common acronyms *n*-MOS, *p*-MOS and MOSFET.) A positive voltage is applied to the drain; when a weaker positive potential is also applied to the gate, electrons cluster in the silicon channel under the gate and create a bridge of negative charge carriers between the two *n*-doped regions. This bridge, called the inversion layer, enables electrons in the source to flow toward the positive voltage on the drain. The current flow can be interrupted by removing the potential on the gate, thereby dispersing the electrons in the inversion layer.

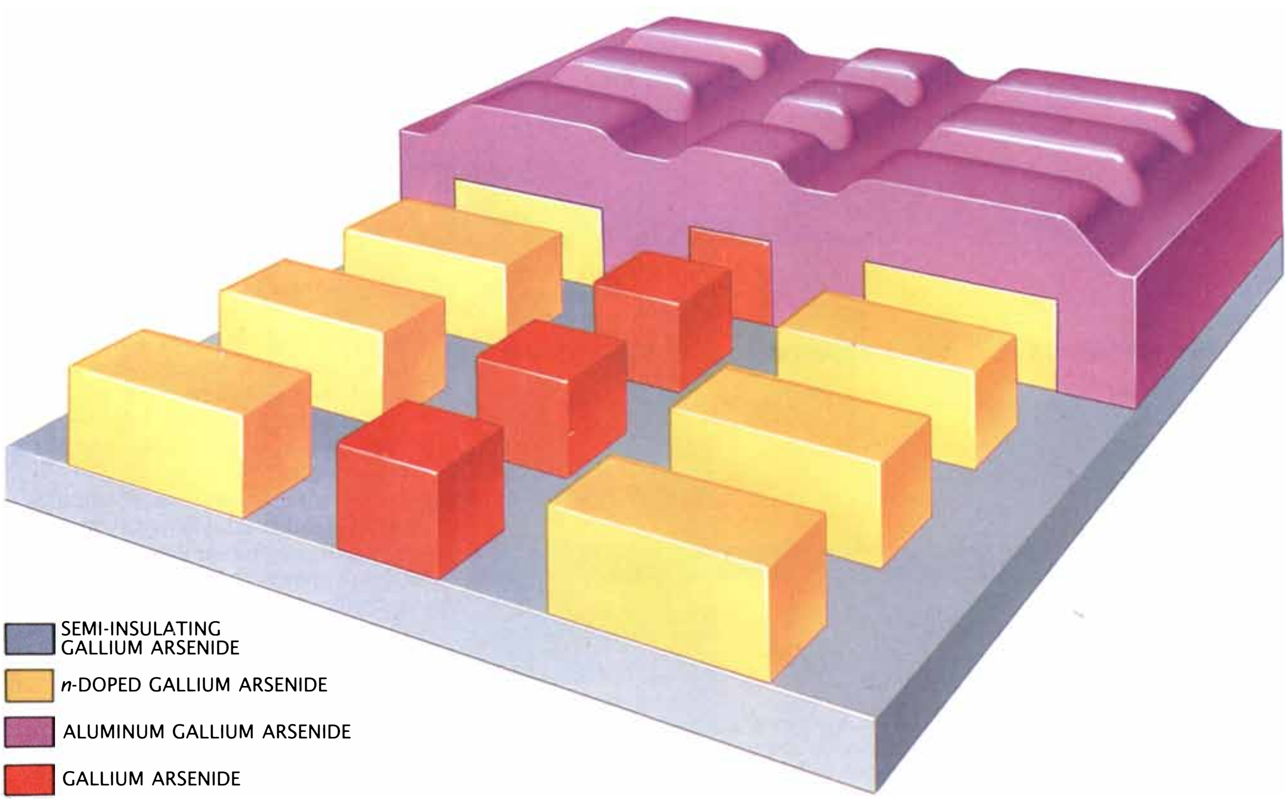
A gallium arsenide transistor also has a gate electrode and terminals that serve as source and drain, but the *n*-doped part of the substrate is not localized [see "Gallium Arsenide Transistors," by William R. Frensley; SCIENTIFIC AMERICAN, August, 1987]. When a positive potential is applied to the gate and the drain, current flows freely from the source; if the gate is given a negative voltage, it repels electrons from the area under it, blocking the path of conduction.

Both transistors are three-terminal devices, and in both of them adjusting the voltage on the gate is the most sensitive means of switching the device. Hence the transistors can be switched "on" and "off" by changing the voltage on the gate. These devices work well at present scales, but with downscaling the distinction between switching states becomes blurred. At smaller scales current leakage prevents a transistor from being truly "off"; it also causes unnecessary consumption of power. Impurities or defects in the semiconductor crystal can scatter electrons, slowing both conduction and switching. For all its usefulness, the modern FET has a problem: the smaller it gets, the worse it switches.

Because the way in which quantum semiconductor devices would function is qualitatively different, quantum devices promise more precise and efficient control of switching in a size regime ordinary transistors could never approach. This difference is manifested in the current-voltage characteristics. In particular, some quantum semiconductor devices exhibit negative differential resistance: that is, there is a voltage range in which the current *decreases* as the applied voltage is increased. On a graph of current versus voltage, this property translates into a current peak and a current valley [*see top illustration on page 100*]. The presence of negative differential resistance is often the only indication a physicist has that quantum effects are operative in an experimental device.

The elusive phenomenon at the heart of quantum effects is the wave nature of electrons. Quantum theory predicts that an electron will exhibit wavelike behavior whenever the region within which it is confined, or the barriers erected to contain it, has dimensions approaching the electron's wavelength. Hence at least one dimension of the features in a quantum device is comparable to the wavelength of an electron. In gallium arsenide at room temperature that wavelength measures just 200 angstrom units (20 billionths of a meter).

The barriers that can contain electrons are barriers of energy rather than physical barriers. All electrons possess a finite amount of energy and are said to occupy energy levels; the levels available are characteristic of a given material. A group of closely spaced levels is called a band. In most solids the energy levels in each band are so closely spaced that they are essentially continuous, and so an electron can change levels with only



QUANTUM CHIP has features 100 times smaller than those of standard chip components. Current flows from one negatively doped (*n*-doped) gallium arsenide block to another through a layer of aluminum gallium arsenide, a gallium arsenide cube and then another aluminum gallium arsenide layer. Because of certain quantum-mechanical effects that come into play in layers of this size, the current a quantum device conducts is extremely sensitive to differences in applied voltage and can therefore be closely controlled. This is an idealized model; a functioning device of such sophistication has not yet been fabricated.

an infinitesimal boost of energy.

The relative positions of energy bands determine whether electricity can be conducted across two different materials. For an electron to pass from one material to another with no change of energy, the bands of the two materials must overlap. Specifically, in the first material the average level occupied by electrons—called the Fermi level—must coincide with an energy band of the second material. If the energy band of the second material occurs at a much higher energy level than the Fermi level of the first, the second material acts as a barrier to electron movement.

For example, under ordinary circumstances aluminum gallium arsenide (AlGaAs) presents a barrier to the electrons in *n*-doped gallium arsenide. An electron cannot pass from the doped GaAs to AlGaAs because the conduction band of AlGaAs is at a much higher energy level than the Fermi level of the GaAs. Yet if the physical dimensions of the barrier are altered in such a way that the wave nature of electrons comes into play, an electron will "tunnel" through the AlGaAs that was once an obstacle. Hence when a layer of AlGaAs thinner than 200 angstroms is sandwiched between two pieces of doped GaAs, the electrons tunnel through it to the GaAs on the other side. This tunneling is one kind of quantum effect.

When barriers confine electrons within a space comparable to an electron wavelength, the electrons are subject to two other, interrelated quantum effects: size quantization and resonance. Size quantization causes the continuum of energy levels that usually exists in the conduction band of a solid to become articulated into discrete energy quanta, or states. It is most aptly described by a density-of-states graph, which shows the number of allowed discrete states of an electron within a fixed energy range [*see illustration on opposite page*].

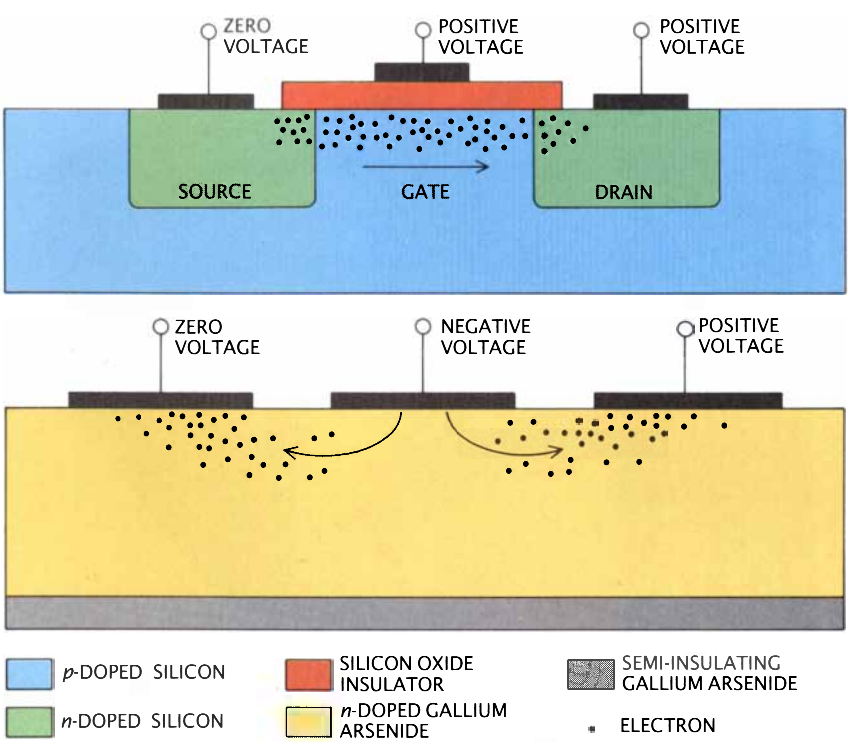
When, for example, a sliver of undoped gallium arsenide is enclosed within AlGaAs barriers, the density-of-states graph for the GaAs looks more like a ladder than a hill. The degree of quantization depends on the degree of confinement. When the electrons in GaAs are restricted in all three dimensions (a "quantum dot"), their energy levels are completely discontinuous; in one-dimensional restriction (a "quantum well") and two-dimensional restriction (a "quantum wire") the levels are still somewhat continuous.

Resonance, the other consequence of quantum confinement, occurs only when some degree of size quantization has been achieved. Electron waves that enter, say, a quantum well are reflected off the far wall of the well; the waves essentially bounce back and forth within the quantum chamber [*see bottom illustration on page 100*]. In doing so they increase the tunneling current substantially—they resonate. Both size quantization and resonance result from the constructive interference of the forward and backward waves. It is difficult to separate the current enhancement that can be attributed to resonance from the enhancement that results from the increased density of states at a given energy level.

As it happens, that distinction is not crucial for transistor operation. What does matter is that in a quantum-effect device two slightly different voltages can evoke profoundly different responses. The differences should be most pronounced in the most confined structure, the quantum dot, because it exhibits the highest degree of quantization. At voltages where tunneling occurs, current is enhanced by the high density of states and by resonance effects to create a peak; at other voltages, the total absence of states at energies intermediate between quantum levels ensures that very little tunneling occurs, and a valley in the current is thus created.

To visualize how these quantum effects could come in handy in a transistor, imagine two slabs of *n*-doped GaAs separated by an AlGaAs-GaAs quantum dot. Electrons trying to pass from one slab of doped GaAs to the other must tunnel through a layer of AlGaAs into the quantum dot and then through another stretch of AlGaAs. They cannot enter the quantum dot, however, unless one of the energy levels in the dot is on a par with the Fermi level of the doped gallium arsenide from which the electrons are emitted.

The Fermi level of the GaAs "emit-



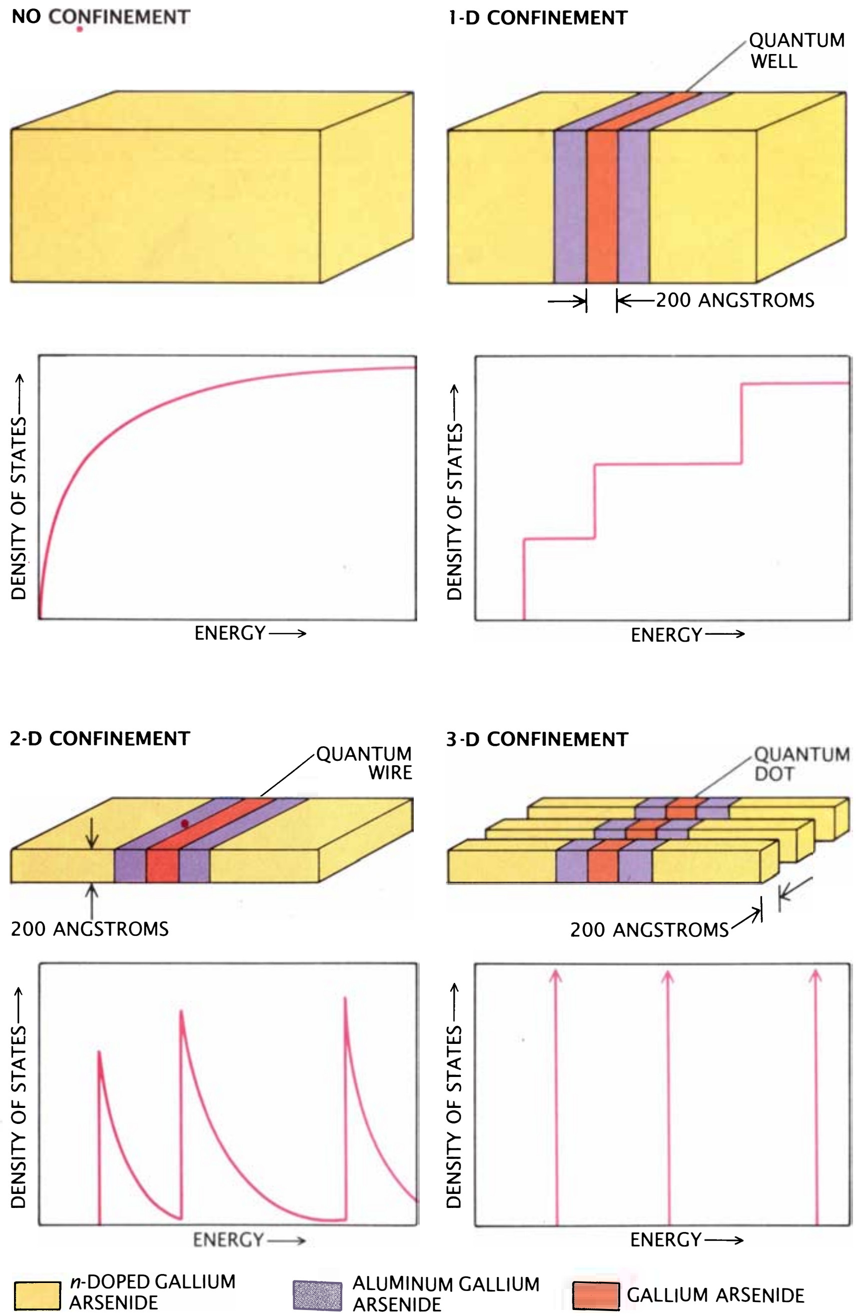
FIELD-EFFECT TRANSISTORS make up the majority of integrated-circuit components today and operate according to the laws of classical physics. In the silicon transistor (*top*) electrons flow between the source and the positively biased drain when a positive voltage is applied to the gate. The gate potential creates a kind of electron bridge between two *n*-doped regions; without it the electrons in the positively doped (*p*-doped) silicon channel disperse and the channel becomes impassable. In contrast, the gallium arsenide transistor (*bottom*) conducts when there is no potential on the gate, but the application of a negative voltage disrupts the flow of electrons from source to drain.

ter" can be raised with respect to the rest of the structure by applying a positive voltage to the doped GaAs on the opposite side of the dot—the "collector." At some voltage the Fermi level of the emitter will attain the same energy as one of the energy levels in the dot, and electrons will move into and resonate within the dot. There is a single voltage at which this occurs; the conduction that takes place at other voltages owing to thermal excitation and to leakage and scattering is negligible. Here, then, is a way to control precisely the switching of a semiconductor device.

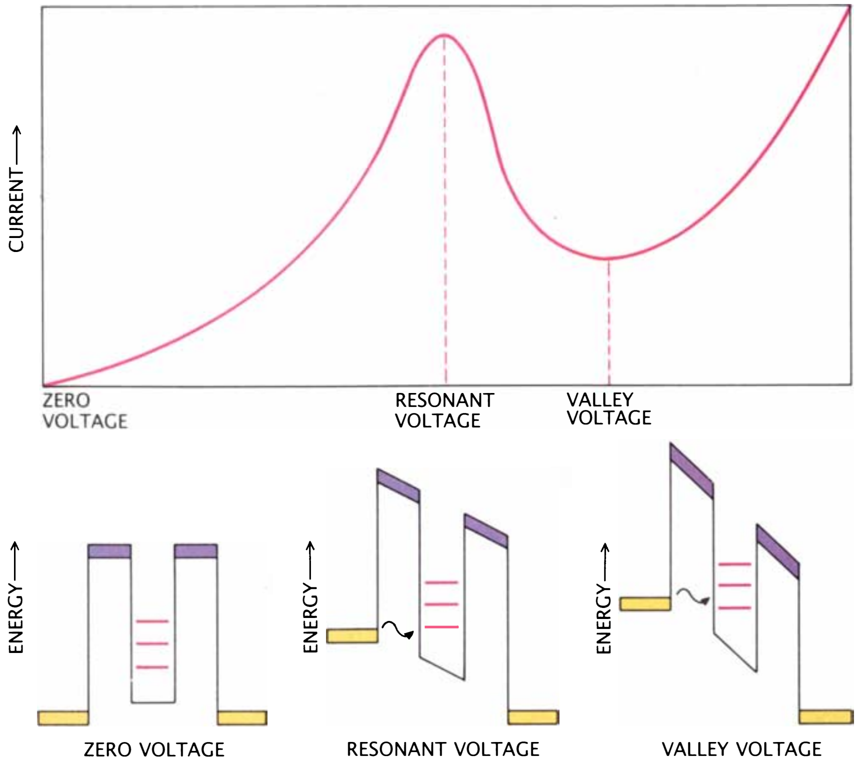
Although from this description the incorporation of a quantum-dot structure in a so-called quantum coupled device may seem like a remote possibility, actually the realization of such a device may not be too many years away. Indeed, the exploitation of quantum effects in semiconductor devices dates from the 1950's. The Esaki tunnel diode, named for its inventor, Leo Esaki, now at the IBM Corporation's Thomas J. Watson Research Center in Yorktown Heights, N.Y., was the first quantum semiconductor device. In this diode *n*- and *p*-doped semiconductors were juxtaposed to create a layer having no charge carriers at all. When the doping was extremely high, the so-called depletion layer became thin enough for electrons to tunnel through. The diode never had widespread appeal, however, because the three-terminal devices that were coming of age at the time proved to be more efficient and convenient.

In the 1960's workers at the Watson Research Center verified that quantum confinement in one dimension takes place in the inversion layer of silicon MOSFET's. Because the influence of quantum effects on device characteristics was so small, that discovery had little impact on transistor development. Subsequent work by Nick Holonyak, Jr., of the University of Illinois at Urbana-Champaign made quantum wells standard ingredients in lasers. In the 1970's Esaki, along with Leroy L. Chang of the Watson Research Center and Raphael Tsu, now at North Carolina Agricultural and Technical State University, carried out the earliest experiments on resonant tunneling through wells. Quantum effects were not deliberately induced in transistors until recently, in the so-called modulation-doped FET's. The quantum wells in these devices, however, serve only to improve the mobility of electrons that otherwise act as they do in conventional transistors.

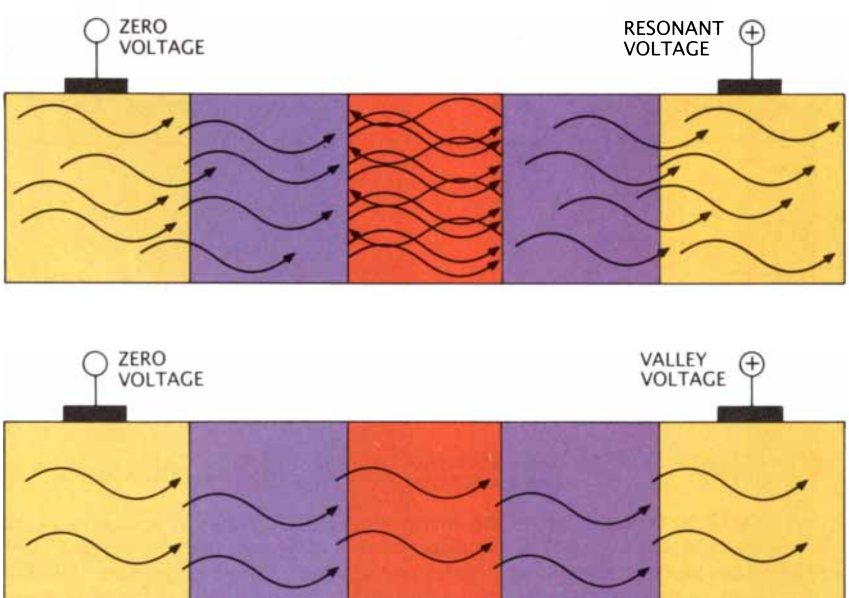
While seemingly tangential, these developments helped to advance the techniques required to make quantum semiconductor devices, so that the technology for constructing experimental structures was at hand when interest in the field finally blossomed. For the past four years the realization of zero-dimensional quantum structures has been the focus of attention for workers around the



QUANTUM CONFINEMENT alters the energy states an electron can occupy in a conducting material. For example, in an ordinary piece of *n*-doped gallium arsenide (*top left*) electrons move freely among a continuum of states, but when barriers of aluminum gallium arsenide are erected in one dimension around a gallium arsenide quantum well the width of an electron wavelength (200 angstrom units), the density of energy states in the well becomes quantized, or discontinuous (*top right*). Restricting the height of the well gives rise to a quantum wire (*bottom left*). The degree of quantization depends on the degree of confinement; true quantization is realized only when gallium arsenide is confined in three dimensions in the quantum-dot structure (*bottom right*).



CURRENT-VOLTAGE CHARACTERISTICS of a quantum-well device reflect the quantization of energy states in the gallium arsenide well. Such devices show a range of voltage in which the current conducted by the device decreases as the voltage applied to one of the *n*-doped gallium arsenide contacts increases. This happens because at one voltage (the resonant voltage) the average energy of electrons in the *n*-doped substance (*top of yellow band*) shifts to a level that coincides with one of the quantum states (*red*) in the well, but beyond that voltage the energy band of the doped gallium arsenide occurs between quantum states. Hence at the resonant voltage an electron (*arrow*) can tunnel through the aluminum gallium arsenide energy barrier (*purple*) into the well, whereas at the valley voltage there are no states for the electron to tunnel into.



TUNNELING ELECTRONS (*arrows*) resonate in a gallium arsenide quantum well (*red*) when a positive bias called the resonant voltage is applied to one of the contacts (*top*). The electron waves bounce back and forth inside the well, enhancing the current to give rise to the peak on the graph at the top of this page. At the valley voltage (*bottom*) little tunneling or resonance takes place, consequently the current dips dramatically.

world. At the AT&T Bell Laboratories, IBM, the Massachusetts Institute of Technology, the University of Cambridge and the Philips Research Laboratories, size quantization in quantum wires has been demonstrated in silicon and gallium arsenide devices alike; quantum dots have been fabricated at AT&T, Bell Communications Research, the Hughes Research Laboratories and the University of Glasgow as well as at Texas Instruments, where the clearest indication of size quantization in dots has been found.

An operational semiconductor device has yet to be constructed from a quantum-dot structure, but a prototype should be available within one or two years. One of the objectives of current research is the conversion of quantum devices, which are most readily constructed as diodes, to three-terminal devices with a third contact directly modulating the potential of the quantum structure. Such a connection would yield the most compact device, and one that would most closely approach the maximum switching speed afforded by tunneling. Devising a technology to manufacture reliable and nondestructive contacts for such thin layers, however, will require a great deal of ingenuity.

By placing quantum dots in close proximity, electrons might also be enabled to tunnel from one dot to another—from one quantized state to another. This arrangement would provide the ultimate in circuit control because the energy states the electrons could assume at both the point of departure and the point of arrival would be strictly dictated. Again, the challenge lies in the formidable task of fabricating structures hundreds of times smaller than any of the features in current semiconductor products. And that degree of downscaling will in turn bring about problems with interconnections and architecture that industry will have to solve before quantum semiconductor devices can be regarded as marketable entities.

The commitment of so many research teams to a problematic technology attests to the tremendous potential of these devices and to the faith that they will take the lead in the next semiconductor revolution. The costs and risks involved must be borne in order to revitalize a rapidly maturing electronics industry; the results can only benefit a society that has learned to depend on integrated circuits in many ways.

No other system of keeping up can compare with ours.

YOUR SYSTEM: a time-consuming, futile struggle to keep up with the information explosion.

The classic texts are convenient references—but the information they contain is obsolete before publication.

Like many physicians, you probably rely on the texts you first used in medical school. But even using the most recent editions, you find material that no longer reflects current clinical thinking—and they lack the latest information on such topics as herpes, oncogenes, AIDS, and photon imaging.

Reading stacks of journals alerts you to recent developments—but can't give you quick answers on patient management.

Struggling through the hundreds of journal pages published each month—even on only the really significant advances in the field—is arduous and memory-taxing. And it's a task that costs physicians valuable time—their most precious resource.

Review courses cover clinical advances—but, months later, do you recall the details of a new procedure or unfamiliar drug?

Seminars can also be costly and make you lose valuable time away from your practice—expenses that may amount to several thousand dollars. And, the speaker's skill often determines how much you learn.

OUR SYSTEM: a rewarding, efficient way to keep yourself up-to-date—and save hundreds of hours of your time for patient care.

A comprehensive, 2,300-page text in two loose-leaf volumes, incorporating the latest advances in medical practice as of the month you subscribe.

This superbly designed, heavily illustrated resource, called "the best written of all [the internal medicine] books" by *JAMA* (251:807, 1984), provides a practical, comprehensive description of patient care in 15 subspecialties. And, because the text is updated each month, the clinical recommendations reflect all the current findings. A practice-oriented index and bibliography of recent articles further enhance the efficiency of the text.

Each month, six to nine replacement chapters to update your text *plus* new references, a four- to five-page news bulletin, and a completely new index.

You'd have to read hundreds of journal pages each month—and memorize the contents—to get the same information SCIENTIFIC AMERICAN *Medicine* contains. With our updated text, you read only the information you really need. Our authors, largely from Harvard and Stanford, sort through the literature and monitor developments, incorporating the significant advances into our chapters.

At no additional cost, a 32-credit CME program, to save you valuable patient-care time and the expense of attending review courses.

Earn 32 Category 1 or prescribed credits per year with our convenient self-study patient management problems; each simulates a real-life clinical situation. Choose either the complimentary printed version or, at a modest extra charge, the floppy-disk version for your IBM® PC/PC Jr., Macintosh,™ or Apple® IIe or II+ (or compatible).

☐ **Yes,** I'd like to try the SCIENTIFIC AMERICAN *Medicine* system. Please enter my subscription at a first-year price of US$245.*
☐ Also enroll me in the CME program. I'd prefer:
  ☐ the printed version at **no additional charge.**
  ☐ the floppy-disk version at $US97* additional. Computer type or compatible:
    ☐ IBM® PC/PC Jr. (128K RAM, DOS 2.0/2.1) ☐ Macintosh™ (512K RAM)
    Apple® II+ (64K RAM) 80-col. card by: ☐ Apple® ☐ Videx® ☐ Apple® IIe (64K RAM)
☐ Enroll me in the floppy-disk CME **only** at US$152.* (Note model above.)
☐ Check enclosed* ☐ Bill me Signature ____________
☐ VISA ☐ MasterCard Exp. Date ________ Account No. ____________

Name ____________ Specialty ____________
Address ____________
City ____________ State ________ Zip Code ________

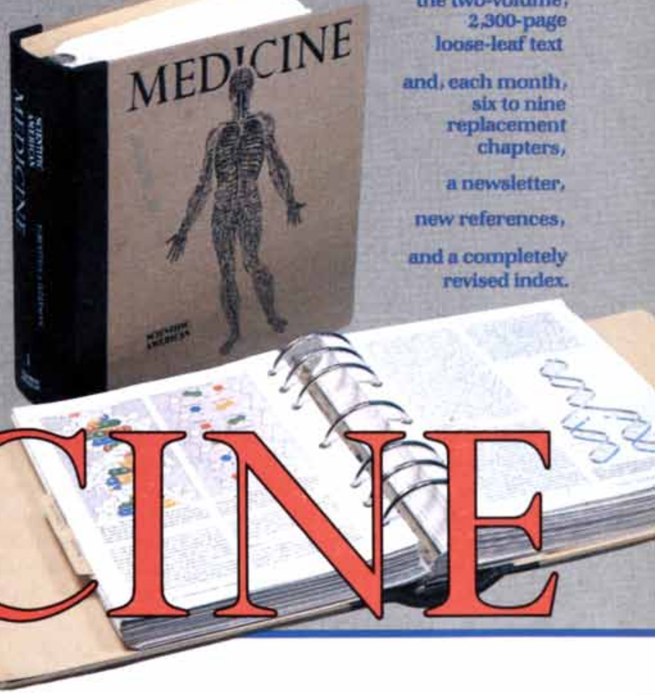
Or, call toll-free: 1-800-345-8112 (in Penn. 1-800-662-2444). 7R

*Add sales tax if for Calif., Ill., Mich., or N.Y. Allow 8 weeks for delivery. Add US$10 for shipping to Canada. IBM is a registered trademark of International Business Machines Corporation. Apple is a registered trademark of Apple Computer, Inc. Videx is a registered trademark of Videx, Inc. Macintosh™ is a trademark of McIntosh Laboratory, Inc. and is used by Apple Computer, Inc. with its express permission.

Your subscription includes:

- the two-volume, 2,300-page loose-leaf text
- and, each month, six to nine replacement chapters,
- a newsletter,
- new references,
- and a completely revised index.

SCIENTIFIC AMERICAN
415 MADISON AVENUE
NEW YORK, N.Y. 10017



MEDICINE