Back in December 1959, future Nobel laureate Richard Feynman gave a visionary and now oft-quoted talk entitled “There’s Plenty of Room at the Bottom.” The occasion was an American Physical Society meeting at the California Institute of Technology, Feynman’s intellectual home then and mine today. Although he didn’t intend it, Feynman’s 7,000 words were a defining moment in nanotechnology, long before anything “nano” appeared on the horizon.

“What I want to talk about,” he said, “is the problem of manipulating and controlling things on a small scale…. What I have demonstrated is that there is room—that you can decrease the size of things in a practical way. I now want to show that there is plenty of room. I will not now discuss how we are going to do it, but only what is possible in principle…. We are not doing it simply because we haven’t yet gotten around to it.”

The breadth of Feynman’s vision is staggering. In that lecture 44 years ago he anticipated a spectrum of scientific and technical fields that are now well established, among them electron-beam and ion-beam fabrication, molecular-beam epitaxy, nanoimprint lithography, projection electron microscopy, atom-by-atom manipulation, quantum-effect electronics, spin electronics (also called spintronics) and microelectromechanical systems (MEMS). The lecture also projected what has been called the “magic” Feynman brought to everything he turned his singular intellect toward. Indeed, it has profoundly inspired my two decades of research on physics at the nanoscale.

Today there is a nanotechnology gold rush. Nearly every major funding agency for science and engineering has announced its own thrust into the field. Scores of researchers and institutions are scrambling for a piece of the action. But in all honesty, I think we have to admit that much of what invokes the hallowed prefix “nano” falls a bit short of Feynman’s mark.

We’ve only just begun to take the first steps toward his grand vision of assembling complex machines and circuits atom by atom. What can be done now is extremely rudimentary. We’re certainly nowhere near being able to commercial-
ly mass-produce nanosystems—integrated multicomponent nanodevices that have the complexity and range of functions readily provided by modern microchips. But there is a fundamental science issue here as well. It is becoming increasingly clear that we are only beginning to acquire the detailed knowledge that will be at the heart of future nanotechnology. This new science concerns the properties and behavior of aggregates of atoms and molecules, at a scale not yet large enough to be considered macroscopic but far beyond what can be called microscopic. It is the science of the mesoscale, and until we understand it, practical devices will be difficult to realize.

Scientists and engineers readily fashion nanostructures on a scale of one to a few hundred nanometers—small indeed, but much bigger than simple molecules. Matter at this mesoscale is often awkward to explore. It contains too many atoms to be easily understood by the straightforward application of quantum mechanics (although the fundamental laws still apply). Yet these systems are not so large as to be completely free of quantum effects; thus, they do not simply obey the classical physics governing the macroworld. It is precisely in this intermediate domain, the mesoworld, that unforeseen properties of collective systems emerge.

Researchers are approaching this transitional frontier using complementary top-down and bottom-up fabrication methods. Advances in top-down nanofabrication techniques, such as electron-beam lithography (used extensively by my own research group), yield almost atomic-scale precision, but achieving success, not to mention reproducibility, as we scale down to the single-digit-nanometer regime becomes problematic. Alternatively, scientists are using bottom-up techniques for self-assembly of atoms. But the advent of preprogrammed self-assembly of arbitrarily large systems—with complexity comparable to that built every day in microelectronics, in...
It is becoming increasingly clear that we are only beginning to acquire the detailed knowledge that will be at the heart of future nanotechnology.

MEMS and (of course) by Mother Nature—is nowhere on the horizon. It appears that the top-down approach will most likely remain the method of choice for building really complex devices for a good while.

Our difficulty in approaching the mesoscale from above or below reflects a basic challenge of physics. Lately, the essence of Feynman’s “Plenty of Room” talk seems to be taken as a license for laissez-faire in nanotechnology. Yet Feynman never asserted that “anything goes” at the nanoscale. He warned, for instance, that the very act of trying to “arrange the atoms one by one the way we want them” is subject to fundamental principles: “You can’t put them so that they are chemically unstable, for example.”

Accordingly, today’s scanning probe microscopes can move atoms from place to place on a prepared surface, but this ability does not immediately confer the power to build complex molecular assemblies at will. What has been accomplished so far, though impressive, is still quite limited. We will ultimately develop operational procedures to help us coax the formation of individual atomic bonds under more general conditions. But as we try to assemble complex networks of these bonds, they certainly will affect one another in ways we do not yet understand and, hence, cannot yet control.

Feynman’s original vision was clearly intended to be inspirational. Were he observing now, he would surely be alarmed when people take his projections as some sort of gospel. He delivered his musings with characteristic playfulness as well as deep insight. Sadly for us, the field that would be called nanotechnology was just one of many that intrigued him. He never really continued with it, returning to give but one redux of his original lecture, at the Jet Propulsion Laboratory in 1983.

New Laws Prevail

In 1959, and even in 1983, the complete physical picture of the nanoscale was far from clear. The good news for researchers is that, by and large, it still is! Much exotic territory awaits exploration. As we delve into it, we will uncover a panoply of phenomena that we must understand before practical nanotechnology will become possible. The past two decades have seen the elucidation of entirely new, fundamental physical principles that govern behavior at the mesoscale. Let’s consider three important examples.

In the fall of 1987 graduate student Bart J. van Wees of the Delft University of Technology and Henk van Houten of the Philips Research Laboratories (both in the Netherlands) and their collaborators were studying the flow of electric current through what are now called quantum-point contacts. These are narrow conducting paths within a semiconductor, along which electrons are forced to flow [see illustration on page 96]. Late one evening van Wees’s undergraduate assistant, Leo Kouwenhoven, was measuring the conductance through the constriction as he varied its width systematically. The research team was expecting to see only subtle conductance effects against an otherwise smooth and unremarkable background response. Instead there appeared a very pronounced, and now characteristic, staircase pattern. Further analysis that night revealed that plateaus were occurring at regular, precise intervals.

David Wharam and Michael Pepper of the University of Cambridge observed similar results. The two discoveries represented the first robust demonstrations of the quantization of electrical conductance. This is a basic property of small conductors that occurs when the wavelike properties of electrons are coherently maintained from the “source” to the “drain”—the input to the output—of a nanoelectronic device.

Feynman anticipated, in part, such odd behavior: “I have thought about some of the problems of building electric circuits on a small scale, and the problem of resistance is serious…. ” But the experimental discoveries pointed out something truly new and fundamental: quantum mechanics can completely govern the behavior of small electrical devices.

Direct manifestations of quantum mechanics in such devices were envisioned back in 1957 by Rolf Landauer, a theoretician at IBM who pioneered ideas in nanoscale electronics and in the physics of computation. But only in the
Nanobridge device allowed Caltech physicists to first observe the quantization of thermal conductance—a fundamental limit to heat flow in minute objects. Four holes (black) etched into a silicon nitride membrane defined an isolated thermal reservoir (central green square) suspended by four narrow bridges. One gold transducer (yellow) electrically heated this reservoir; the second measured its temperature. Thin superconducting films (blue) on top of the bridges electrically connected the transducers to off-chip instrumentation but carried no heat. The reservoir therefore cooled only through the silicon nitride bridges, which were so narrow that they passed only the lowest-energy heat waves.

Mid-1980s did control over materials and nanofabrication begin to provide access to this regime in the laboratory. The 1987 discoveries heralded the heyday of “mesoscopia.”

A second significant example of newly uncovered mesoscale laws that have led to nascent nanotechnology was first postulated in 1985 by Konstantin Likharev, a young physics professor at Moscow State University working with postdoctoral student Alexander Zorin and undergraduate Dmitri Averin. They anticipated that scientists would be able to control the movement of single electrons on and off a “coulomb island,” a conductor weakly coupled to the rest of a nanocircuit. This could form the basis for an entirely new type of device, called a single-electron transistor. The physical effects that arise when putting a single electron on a coulomb island become more robust as the island is scaled downward. In very small devices, these single-electron charging effects can completely dominate the current flow.

Such considerations are becoming increasingly important technologically. Projections from the International Technology Roadmap for Semiconductors, prepared by long-range thinkers in the industry, indicate that by 2014 the minimum feature size for transistors in computer chips will decrease to 20 nanometers. At this dimension, each switching event will involve the equivalent of only about eight electrons. Designs that properly account for single-electron charging will become crucial.

By 1987 advances in nanofabrication allowed Theodore A. Fulton and Gerald J. Dolan of Bell Laboratories to construct the first single-electron transistor [see illustration on page 98]. The single-electron charging they observed, now called the coulomb blockade, has since been seen in a wide array of structures. As experimental devices get smaller, the coulomb blockade phenomenon is becoming the rule, rather than the exception, in weakly coupled nanoscale devices. This is especially true in experiments in which electric currents are passed through individual molecules. These molecules can act like coulomb islands by virtue of their weak coupling to electrodes leading back to the macro-world. Using this effect to advantage and obtaining robust, reproducible coupling to small molecules (in ways that can actually be engineered) are among the important challenges in the new field of molecular electronics.

In 1990, against this backdrop, I was at Bell Communications Research studying electron transport in mesoscopic semiconductors. In a side project, my colleagues Larry M. Schiavone and Axel Scherer and I began developing techniques that we hoped would elucidate the quantum nature of heat flow. The work required much more sophisticated nanostructures than the planar devices used to investigate mesoscopic electronics. We needed freely suspended devices, structures possessing full three-dimensional relief. Ignorance was bliss; I had no idea the experiments would be so involved that they would take almost a decade to realize.

The first big strides were made after I moved to Caltech in 1992, in a collaboration with John M. Worlock of the University of Utah and two successive postdocs in my group. Thomas S. Tighe developed the methods and devices that generated the first direct measurements of heat flow in nanostructures. Subsequently, Keith C. Schwab revised the design of the suspended nanostructures and put in place ultrasensitive superconducting instrumentation to interrogate them at ultralow temperatures, at which the effects could be seen most clearly.

In the late summer of 1999 Schwab and a group studying nanoscale systems. Among the holy grails his team is chasing are a billionfold improvement in present-day calorimetry, which would allow observation of the individual heat quanta being exchanged as nanodevices cool, and a quadrillionfold increase in the sensitivity of magnetic resonance imaging, which would enable complex biomolecules to be visualized with three-dimensional atomic resolution.
Continually to increase the density of transistors and the clock rates (frequencies) of microprocessors, the problem of keeping microchips cool to avoid complete system failure is becoming monumental. This will only become further exacerbated in nanotechnology.

Considering even this complexity, Feynman said, “Let the bearings run dry; they won’t run hot because the heat escapes away from such a small device very, very rapidly.” But our experiments indicate that nature is a little more restrictive. The thermal conductance quantum can place limits on how effectively a very small device can dissipate heat. What Feynman envisioned can be correct only if the nanoengineer designs a structure so as to take these limits into account.

From the three examples above, we can arrive at just one conclusion: we are only starting to unveil the complex and wonderfully different ways that nanoscale systems behave. The discovery of the electrical and thermal conductance quanta and the observation of the coulomb blockade are true discontinuities—abrupt changes in our understanding. Today we are not accustomed to calling our discoveries “laws.” Yet I have no doubt that electrical and thermal conductance quantization and single-electron-charging phenomena are indeed among the universal rules of nanodesign. They are new laws of the nanoworld. They do not contravene but augment and clarify some of Feynman’s original vision. Indeed, he seemed to have anticipated their emergence: “At the atomic level, we have new kinds of forces and new kinds of possibilities, new kinds of effects. The problems of manufacture and reproduction of materials will be quite different.”

We will encounter many more such discontinuities on the path to true nanotechnology. These welcome windfalls will occur in direct synchrony with advances in our ability to observe, probe and control nanoscale structures. It would seem wise, therefore, to be rather modest and circumspect about forecasting nanotechnology.

The Boon and Bane of Nano

The Nanoworld is often portrayed by novelists, futurists and the popular press as a place of infinite possibilities. But as you’ve been reading, this domain is not some ultraminiature version of the Wild West. Not everything goes down there; there are laws. Two concrete illustrations come from the field of nanoelectromechanical systems (NEMS), in which I am active.

Part of my research is directed toward harnessing small mechanical devices for sensing applications. Nanoscale structures appear to offer revolutionary potential; the smaller a device, the more susceptible its physical properties to alteration. One example is resonant detectors, which are frequently used for sensing mass. The vibrations of a tiny mechanical element, such as a small cantilever, are intimately linked to the element’s mass, so the addition of a minute
The difficulties in communication between the nanoworld and the macroworld represent a central issue in the development of nanotechnology.

amount of foreign material (the “sample” being weighed) will shift the resonant frequency. Work in my lab by then postdoc Kamil Ekinci shows that nanoscale devices can be made so sensitive that “weighing” individual atoms and molecules becomes feasible.

But there is a dark side. Gaseous atoms and molecules constantly adsorb and desorb from a device’s surfaces. If the device is macroscopic, the resulting fractional change in its mass is negligible. But the change can be significant for nanoscale structures. Gases impinging on a resonant detector can change the resonant frequency randomly. Apparently, the smaller the device, the less stable it will be. This instability may pose a real disadvantage for various types of futuristic electromechanical signal-processing applications. Scientists might be able to work around the problem by, for example, using arrays of nanomechanical devices to average out fluctuations. But for individual elements, the problem seems inescapable.

A second example of how “not everything goes” in the nanoworld relates more to economics. It arises from the intrinsically ultralow power levels at which nanomechanical devices operate. Physics sets a fundamental threshold for the minimum operating power: the ubiquitous, random thermal vibrations of a mechanical device impose a “noise floor” below which real signals become increasingly hard to discern. In practical use, nanomechanical devices are optimally excited by signal levels 1,000-fold or a million-fold greater than this threshold. But such levels are still a millionth to a billionth of the amount of power used for conventional transistors.

The advantage, in some future nanomechanical signal-processing system or computer, is that even a million nanomechanical elements would dissipate only a millionth of a watt, on average. Such ultralow power systems could lead to wide proliferation and distribution of cheap, ultraminiature “smart” sensors that could continuously monitor all of the important functions in hospitals, in manufacturing plants, on aircraft, and so on. The idea of ultraminiature devices that drain their batteries extremely slowly, especially ones with sufficient computational power to function autonomously, has great appeal. But here, too, there is a dark side. The regime of ultralow power is quite foreign to present-day electronics. Nanoscale devices will require entirely new system architectures that are compatible with amazing low power thresholds. This prospect is not likely to be received happily by the computer industry, with its overwhelming investment in current devices and methodology. A new semiconductor processing plant today costs more than $1 billion, and it would probably have to be retooled to be useful. But I am certain that the revolutionary prospects of nanoscale devices will eventually compel such changes.

Monumental Challenges

CERTAINLY A HOST of looming issues will have to be addressed before we can realize the potential of nanoscale devices. Although each research area has its own concerns, some general themes emerge. Two challenges fundamental to my current work on nanomechanical systems, for instance, are relevant to nanotechnology in general.

Challenge I: Communication between the macroworld and the nanoworld. Nanoscale devices are incredibly small, yet their motion can be far smaller. For example, a nanoscale beam clamped on both ends vibrates with minimal harmonic distortion when its vibration amplitude is kept below a small fraction of its thickness. For a 10-nanometer-thick beam, this amplitude is only a few nanometers. Building the requisite, highly efficient transducers to transfer information from such a device to the macroworld involves reading out information with even greater precision.

Compounding this problem, the natural frequency of the vibration increases as the size of the beam is decreased. So to track the device’s vibrations usefully, the ideal NEMS transducer must be capable of resolving extremely small displacements, in the picometer-to-femtometer (trillionth to quadrillionth of a meter) range, across very large bandwidths, extending into the microwave range. These twin requirements pose a truly monumental challenge, one much more significant than those faced so far in MEMS work. A further complication is that most of the methodologies from MEMS are inapplicable; they simply don’t scale down well to nanometer dimensions.

Richard Feynman predicted the rise of nanotechnology in a landmark 1959 talk at Caltech. “The principles of physics,” he said, “do not speak against the possibility of maneuvering things atom by atom.” But he also anticipated that unique laws would prevail; they are finally being discovered today.
These difficulties in communication between the nanoworld and the macro-world represent a generic issue in the development of nanotechnology. Ultimately, the technology will depend on robust, well-engineered information transfer pathways from what are, in essence, individual macromolecules. Although the grand vision of futurists may involve self-programmed nanobots that need direction from the macroworld only when they are first wound up and set in motion, it seems more likely that most nanotechnological applications realizable in our lifetimes will entail some form of reporting up to the macroworld and feedback and control back down. The communication problem will remain central.

Orchestrating such communication immediately invokes the very real possibility of collateral damage. Quantum theory tells us that the process of measuring a quantum system nearly always perturbs it. This can hold true even when we scale up from atoms and molecules to nanosystems comprising millions or billions of atoms. Coupling a nanosystem to probes that report back to the macroworld always changes the nanosystem’s properties to some degree, rendering it less than ideal. Introducing the transducers required for communication will do more than just increase the nanosystem’s size and complexity. They will also necessarily extract some energy to perform their measurements and can degrade the nanosystem’s performance. Measurement always has its price.

**Challenge II: Surfaces.**

As we shrink MEMS to NEMS, device physics becomes increasingly dominated by the surfaces. Much of the foundation of solid-state physics rests on the premise that the surface-to-volume ratio of objects is infinitesimal, meaning that physical properties are always dominated by the physics of the bulk. Nanoscale systems are so small that this assumption breaks down completely.

For example, mechanical devices patterned from single-crystal, ultrapure materials can contain very few (even zero) crystallographic defects and impurities. My initial hope was that, as a result, there would be only very weak damping of mechanical vibrations in monocrystalline NEMS. But as we shrink mechanical devices, we repeatedly find that acoustic energy loss seems to increase in proportion to the increasing surface-to-volume ratio. This result clearly impli-
Cates surfaces in the devices’ vibrational energy-loss processes. In a state-of-the-art silicon beam measuring 10 nanometers wide and 100 nanometers long, more than 10 percent of the atoms are at or next to the surface. It is evident that these atoms will play a central role, but understanding precisely how will require a major, sustained effort.

In this context, nanotube structures, which have been heralded lately, look ideal. A nanotube is a crystalline, rodlike material perfect for building the miniature vibrating structures of interest to us. And because it has no chemical groups projecting outward along its length, one might expect that interaction with “foreign” materials at its surfaces would be minimal. Apparently not. Although nanotubes exhibit ideal characteristics when shrouded within pristine, ultrahigh vacuum environments, samples in more ordinary conditions, where they are exposed to air or water vapor, evince electronic properties that are markedly different. Mechanical properties are likely to show similar sensitivity. So surfaces definitely do matter. It would seem there is no panacea.

Payoff in the Glitches

FUTURISTIC THINKING is crucial to making the big leaps. It gives us some wild and crazy goals—a holy grail to chase. And the hope of glory propels us onward. Yet the 19th-century chemist Friedrich August Kekulé once said, “Let us learn to dream, gentlemen, then perhaps we shall find the truth.... But let us beware of publishing our dreams before they have been put to the proof by the waking understanding.”

This certainly holds for nanoscience. While we keep our futuristic dreams alive, we also need to keep our expectations realistic. It seems that every time we gain access to a regime that is a factor of 10 different—and presumably “better”—two things happen. First, some wonderful, unanticipated scientific phenomenon emerges. But then a thorny host of underlying, equally unanticipated new problems appear. This pattern has held true as we have pushed to decreased size, enhanced sensitivity, greater spatial resolution, higher magnetic and electric fields, lower pressure and temperature, and so on. It is at the heart of why projecting forward too many orders of magnitude is usually perilous. And it is what should imbue us with a sense of humility and proportion at this, the beginning of our journey. Nature has already set the rules for us. We are out to understand and employ her secrets.

Once we head out on the quest, nature will frequently hand us what initially seems to be nonsensical, disappointing, random gibberish. But the science in the glitches often turns out to be even more significant than the grail motivating the quest. And being proved the fool in this way can truly be the joy of doing science. If we had the power to extrapolate everything correctly from the outset, the pursuit of science would be utterly dry and mechanistic. The delightful truth is that, for complex systems, we do not, and ultimately probably cannot, know everything that is important.

Complex systems are often exquisitely sensitive to a myriad of parameters beyond our ability to sense and record—much less control—with sufficient regularity and precision. Scientists have studied, and in large part already understand, matter down to the fundamental particles that make up the neutrons, protons and electrons that are of crucial importance to chemists, physicists and engineers. But we still cannot deterministically predict how arbitrarily complex assemblages of these three elemental components will finally behave en masse. For this reason, I firmly believe that it is on the foundation of the experimental science under way, in intimate collaboration with theory, that we will build the road to true nanotechnology. Let’s keep our eyes open for surprises along the way!

MORE TO EXPLORE


The author’s group: www.its.caltech.edu/~nano

Richard Feynman’s original lecture “There’s Plenty of Room at the Bottom” can be found at www.its.caltech.edu/~feynman