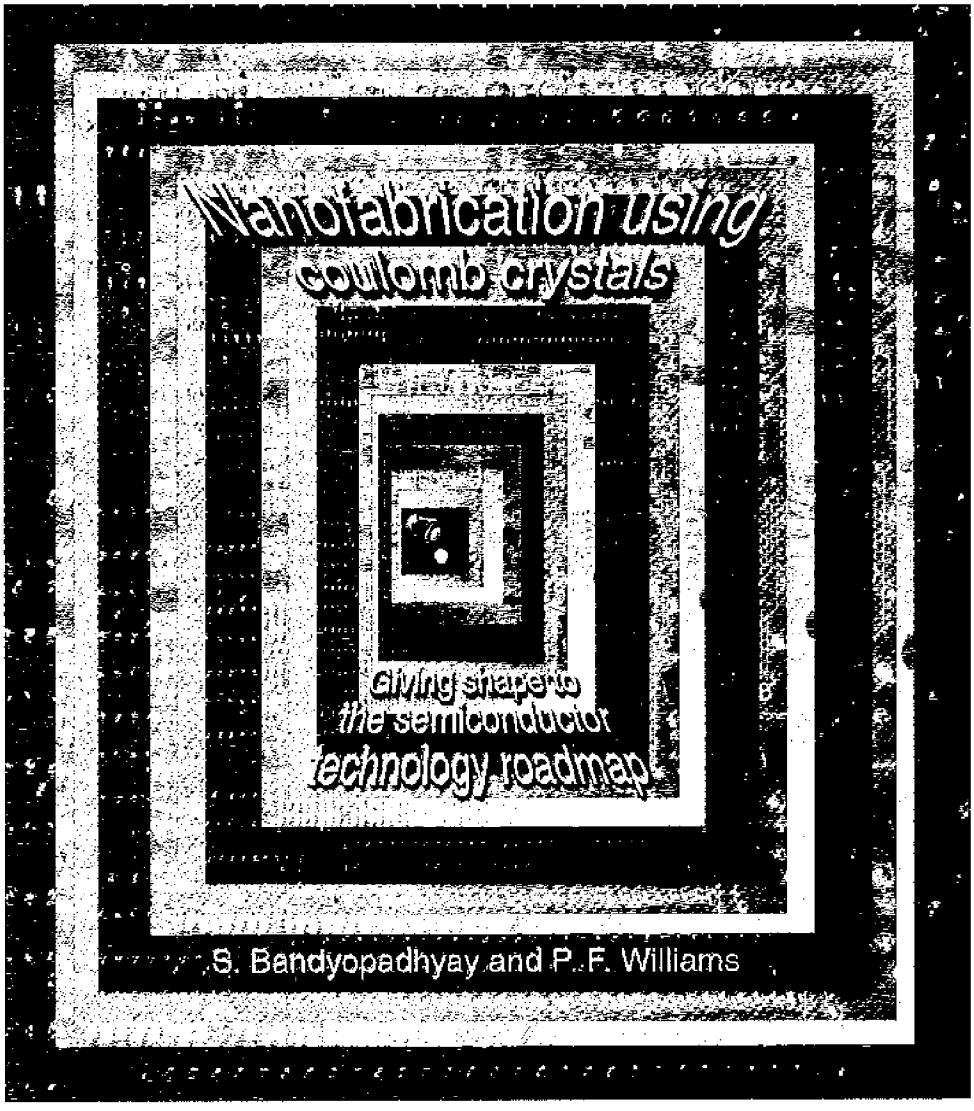
4-1-2000

Nanofabrication Using Coulomb Crystals: Giving shape to the semiconductor technology roadmap

S. Bandyopedhyay
Univeristy of Nebraska - Lincoln

P. F. Williams
University of Nebraska - Lincoln, pfw@moi.unl.edu



In 1959 Richard Feynman delivered a lecture with the title, "There is Plenty of Room at the Bottom." The audience was attendees of the American Physical Society meeting at Caltech. What Feynman, a great seer, talked about was using progressively smaller structures for useful purposes such as storing bits of information.

He even offered a $1,000 (US) prize for anyone who could take the information on the page of a book and put it in an area 1/25,000 times smaller in linear scale. In 1986, 26 years after Feynman offered the prize, it was won by a Stanford University graduate student. Tom Newman used electron beam lithography to write out the first page of "A Tale of Two Cities" by Charles Dickens. The letters were about 50 atoms wide.

Feynman was a talented visionary who saw the possibility of storing information in ultra small structures. He believed the spin of a single electron could encode binary digital data. Hence, it could act as a useful device. Such miniscule devices would allow a tremendous amount of information to be stored in very small volumes of space.

Nobody has really made a useful device out of a single electron quite yet. However, a single electron performing useful device operations is no longer a pipe dream. Single electronics is an active field of research. In a single electron transistor, device action is realized from the discrete transfer of single electrons within a very small volume of material (a quantum dot). A quantum dot is still around 21 orders of magnitude larger than a single electron in volume. Nonetheless, the device's operation is dependent upon the dynamics of that miniscule entity- a single electron00whose diameter is thought to be ~10^{-15} meters.

The Semiconductor Industry Association, and its *International Technology Roadmap*, does talk about Complimentary Metal-Oxide-Semiconductor (CMOS) transistors with feature sizes approaching a quantum dot (typically 1000 atoms). However, such transistors are not expected to enter the market earlier than the year 2007 and even that projection may be somewhat optimistic. There are many serious issues to address before quantum-dot sized transistors are sold off-the-shelf. However, the potential is certainly there.

The advantages are obvious: smaller structures permit higher packing densities of functional units on chips, improved speed and lower cost. For the semiconductor industry, continuous downscaling and miniaturization of electronic devices has fueled worldwide annual sales of over $500 billion dollars (US).

Ultra small semiconductor structures, like quantum dots, with nanometer scale dimensions are wonderful entities. Everybody would like to use them for integrated circuits (ICs). But, making them is not at all easy.

The obvious manufacturing strategy would be to use nanolithography. Typically, nanolithography involves delineating the ultra small pattern on a semiconductor wafer with an electron beam, ion beam, or even a standard scanning tunneling microscope tip (a process called "direct-write"). This is followed by "chiseling" out the pattern using a combination of standard semiconductor techniques such as etching, lift-off, etc. This is exactly what Tom Newman used to meet Feynman's challenge.

Such an approach works for winning a $1,000 prize, but it has two major shortcomings for commercial IC manufacturing. First, direct-writing is a *serial* process. Each feature on each wafer is handled one at a time. This results in an excruciatingly slow throughput—maybe too slow for commercial viability. Second, the structures are damaged with the electron- or ion-beams during pattern delineation. This damage introduces spurious charges and defects on the surfaces of the semiconductor structure. The surface defects capture all the

0278-6648/00/$10.00 © 2000 IEEE

mobile charge carriers within some distance from the surface and deplete this region of carriers. The depletion layer often has several layers of atoms. It is typically 10-100 nm wide in common semiconductors.

If the structure is more than 1000 nm (1 μm) wide, then maybe less than 10% of the structure would be depleted. This is no big deal. However, if the structure is less than 100 nm wide, then surface states can deplete the *entire* structure rendering it useless. Surface damage is a serious issue in nanofabrication and must be avoided at all costs.

There are other techniques for nanofabrication. The most popular one currently is a generic approach termed *self-assembly*. Generally, self-assembly is any method where the natural physical or chemical evolution of a material system—driven to a minimal energy thermodynamic configuration—spontaneously assembles an array of nanostructures on a chosen substrate. Useful self-assembly must result in arrays where: 1) the standard deviation in the size of the nanostructures is much smaller than the average size (size selectivity), and 2) the array must have long range spatial order in all relevant dimensions (ordering).

Very few self-assembly techniques meet both requirements. The rare ones that do have been successfully used to make both quantum dot and quantum wire arrays. *Quantum wires* are quasi one-dimensional structures.

Self-assembly has emerged as a strong competitor to fine-line lithography for delineating size-selected and ordered nanostructure arrays. The reason is that it has at least two advantages: 1) it is a *parallel* process whereby several wafers can be handled simultaneously resulting in a fast throughput, and 2) it can be a "gentle" technique causing little or no processing damage to finished nanostructures.

In our laboratory, we have self-assembled essentially periodic arrays of semiconductor quantum dots with diameters of 3-4 nm. These dots (which are of a wide bandgap semiconductor) have shown optical activity. To our knowledge, dots of 10 nm diameter (of wide bandgap semiconductors) fabricated by e-beam lithography do not usually show measurable optical activity. The poor quantum efficiency is caused by surface states introduced during pattern delineation (i.e., surface damage).

However, the self-assembly process does have some drawbacks. It is not as well controlled as lithography. Also, it is incapable of producing arbitrary patterns. The last drawback is the more serious one.

Conventional self-assembly

At present, there are two major approaches for self-assembly: *physical* approaches and *chemical* approaches. The Stranksi-Krastanow (S-K) method (strained quantum dots on a lattice-mismatched substrate) is a popular physical approach. Here a material is deposited on a substrate using thin film deposition techniques such as molecular beam epitaxy. The lattice constants of the substrate and the material are different. This results in strain on the top material.

Under appropriate conditions, the top material "bunches up" and quantum dots nucleate spontaneously. They are connected by a so-called "wetting layer" (see Fig. 1). The dots are size controlled (less than 10% standard deviation in size), but they nucleate randomly on the surface. Thus, there is no spatial ordering. Recently, some advances for improving the spatial ordering through the effect of stacking several layers of dots vertically have been made. However, it is still far from perfect.

Chemical techniques are typically approaches where ionic reaction at an electrolyte-metal interface results in the nucleation of self-ordered nanostructures. These can be used as templates for synthesizing quantum dots or wires. Note that in this approach, the "template," rather than the dots themselves, are self-assembled. But the template can then be easily applied to create dots.

In Fig. 2, we show an atomic force microscope image of a porous film self-

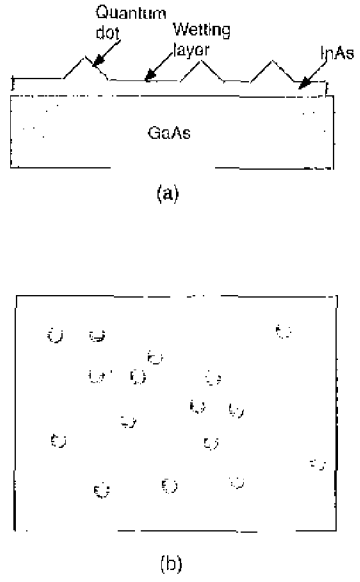


Fig. 1 a) Side view of Stranksi-Krastanow quantum dots, b) Top view

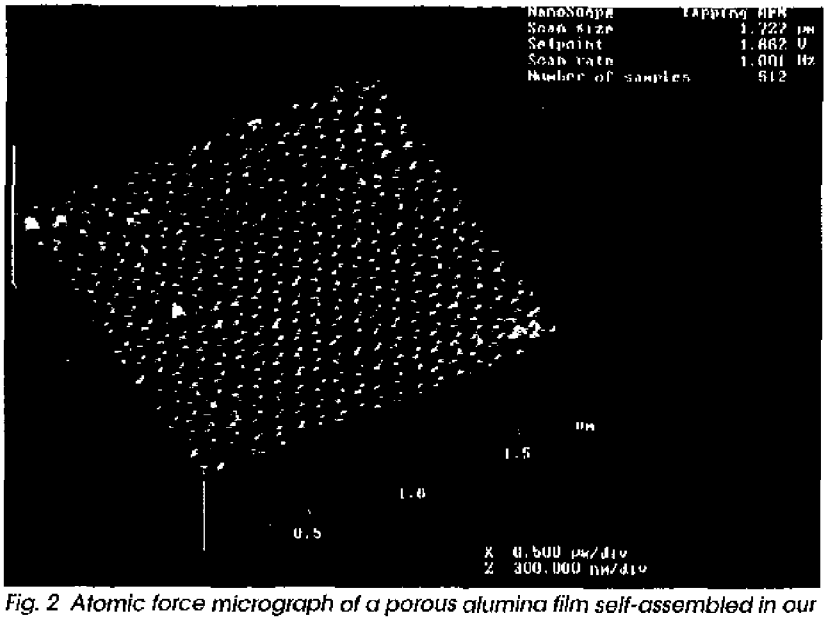


Fig. 2 Atomic force micrograph of a porous alumina film self-assembled in our laboratory using anodization of alumina in oxalic acid (an electrochemical technique). The average pore diameter is 52 nm and the standard deviation in pore size is less than 1%. Micrograph courtesy of Dr. Min Zheng.

assembled in our lab. The walls of the pores are alumina (Al_2O_3). It is a "template." One can easily fill up the pores with the material of interest using electrodeposition. This will result in a virtually periodic array of quantum dots surrounded by alumina. Thus, the self-assembled template has created a quasi-periodic array of dots.

The advantages of the chemical techniques are twofold: 1) they produce more *ordered* arrangements of dots than does Stranksi-Krastanow growth, and 2) chemical techniques are much cheaper in terms of capital investment. The lack of ordering in Stranksi-Krastanow dots has restricted their applications to optical devices (e.g., quantum dot lasers), or to discrete electronic devices (e.g., transistors). (Several dots together constitute a single device.)

In contrast, electrochemical self-assembly produces at least quasi-periodic arrays of quantum dots. They are synergistic with neural or cellular automata type computing architectures. (A single dot can act as a circuit node or an active device.) This can lead to very powerful, ultra dense and extremely fast large-scale circuits.

Unfortunately, electrochemical self-assembly techniques also have some drawbacks. The material quality of the finished nanostructures is often poor. The dots are nucleated in a chemical solution rather than in a high vacuum environment.

What's more, these techniques are not immediately compatible with conventional lithographic techniques for Ultra-Large Scale Integration (ULSI) chip fabrication. Hence, they are not yet attractive to commercial manufacturers.

Finally, most chemical techniques share a common feature; they all involve wet chemistry insofar as the electrolytes invariably contain water. Consequently, they are unsuitable for materials such as high T_c superconductors, chemical catalysts such as K, Na, Mo, etc., and some organic non-linear optical materials. All these materials react adversely with water. Thus, one cannot make quantum dots using them with ordinary electrochemical synthesis. There are nonaqueous techniques for electrochemistry that could be adapted for self-assembly. They may be able to handle these materials. However, there is no standard approach yet. Also, much research needs to be done before it can emerge as a useful tool for nanosynthesis.

Coulomb crystals

There is a novel "dry" self-assembly technique based on an entity called a Coulomb crystal. It is "dry" because the materials used for the quantum dots never come into contact with moisture. Additionally, this new technique is somewhat flexible and may be integrated with lithography and standard semiconductor processing for seamless IC manufacturing.

A Coulomb crystal is a more-or-less regularly spaced array of small particles produced in plasma. Plasma is a gas containing free electrons and positive ions in nearly equal concentrations. Plasmas can be produced placing two electrodes in a low-pressure (typically about 0.001 of atmospheric pressure) gas. Then a voltage sufficiently high to cause the gas to break down into an electrical discharge is applied. A "neon" sign is a common application of such discharges (see box).

Figure 3 shows a schematic diagram of a typical experimental setup for producing Coulomb crystals. Two electrodes are placed in a vacuum cell that can be filled with the desired gas (such as argon) at the desired pressure (typically about 1 Torr). A Radio Frequency (R.F.) voltage (usually 13.56 MHz) is applied to the electrodes through a capacitor as indicated. The result is a mostly-uniform glow discharge between the two electrodes.

The particles are introduced into the plasma through a "salt shaker," a small cup with a porous membrane on the bottom. The "salt-shaker" is mechanically vibrated or "shaken" to release particles into the plasma through the membrane. The pores in the membrane have a uniform diameter. This way the sieving action introduces an upper limit on the size of the particles introduced into the plasma.

After the introduction, the particles fall through the plasma acquiring a negative charge. They are levitated in a thin layer just above the lower electrode. Because they are negatively charged, they repel each other.

To form a Coulomb crystal, we "fence" the particles to prevent them from escaping from the top of the electrode. In this case, the lowest energy configuration is one in which the particles are arranged on a regularly spaced lattice. That is a Coulomb crystal (see Fig. 4).

Particle charging

Consider a particle in plasma. The plasma consists of equal densities of free electrons and positive ions. The mass of the electrons is about 1/10,000 that of the ions. If both particles had the same temperature (kinetic energy), the electrons would be moving about 100 times faster than the ions.

Power is supplied to the plasma from the applied electric field. (That is how the ionization and light emission are maintained.) This power is supplied primarily to the free electrons. For that reason, the electrons are typically about 100 times hotter than the ions. This makes the electrons move about 10

Neon signs

A neon sign consists of a glass tube bent to form whatever message is desired. It contains two electrodes and a suitable gas (often not neon because it's expensive) at low pressure. Applying a voltage of several kV to the two electrodes causes a discharge to form and the gas to glow. Free electrons are accelerated by the electric field, and collide with neutral atoms, exciting them and occasionally knocking off another electron. The atoms in excited states then relax to lower-lying states by emitting photons with energy equal to the difference in energy of the two levels involved. That is the source of the glow.—*SB & PFW*

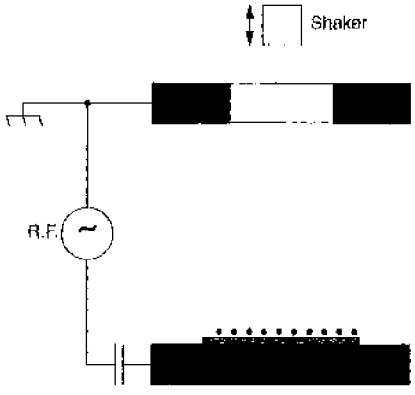


Fig. 3 Schematic diagram of a typical experimental setup for producing Coulomb crystals

times faster yet, for a total speed ratio of about 1000.

Both electrons and positive ions strike the particle. When an electron hits, there is a good probability that it will be captured. When an ion hits, there is a good probability that it will acquire one of the particle's electrons, thereby neutralizing it.

Because the densities of electrons and ions are about the same and the electrons are moving about 1000 times faster, the rate of electron collisions with a neutral particle is much higher than that of ion collision. Thus, the particles introduced into the plasma will charge negatively.

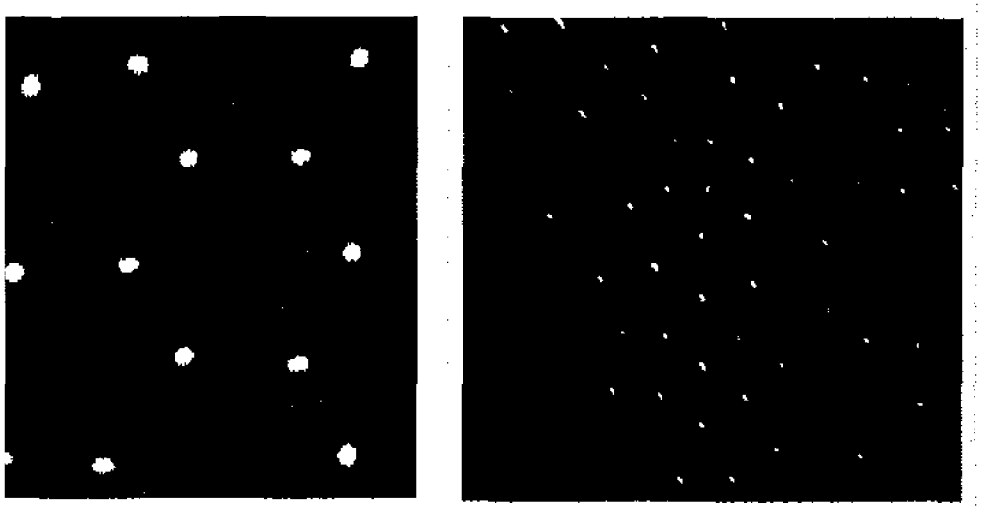


Fig. 4 Photographs of Coulomb crystals. The photo in a) was published by Chu and Lin, and b) was obtained in our laboratory.

As the particle becomes negative, it repels electrons reducing the collision rate. Steady state occurs when the charge has risen to the point that the collision rate of electrons and positive ions is the same. This can be viewed as a zero current condition, because the charge's rate of change on the particle is zero. (The positive charge acquired from ions is balanced by the negative charge acquired from electrons.)

Particle levitation

The lower electrode of the cell shown in Fig. 3 also acquires a negative charge. The reason for the charge is also related to the relative mass of the electrons and the ions. A sinusoidal voltage is applied to the electrode by the R.F. generator. When the electrode is uncharged, the potential is positive and negative for equal amounts of time. While the electrode is positive, it attracts electrons; and, while it is negative, it attracts positive ions. Because the electrons are much lighter, the total current of electrons during the positive cycle is larger than the total current of positive ions during the negative cycle.

The electrode, therefore, acquires a negative charge driving the potential down. As the negative charge increases, the average potential decreases. This decreases the fraction of the cycle that the electrode potential is positive. In steady state, the charge is such that the total negative charge transferred to it during the time it is positive is just equal to the net positive charge transferred while it is negative. An accurate estimate of the average electrode potential at this time is difficult, but it would be about the zero-to-peak R.F. voltage.

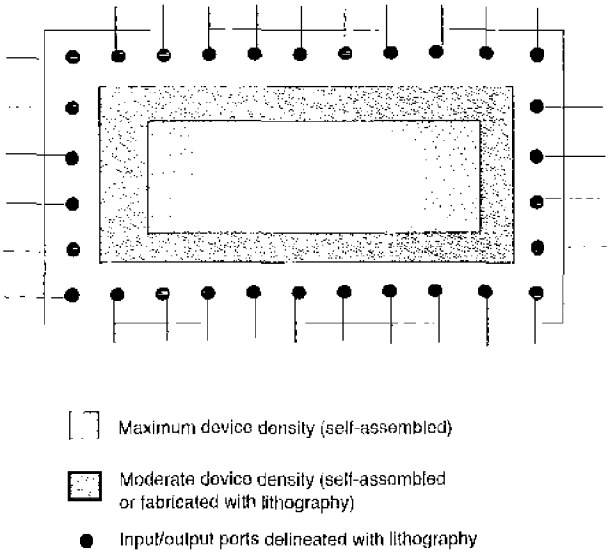


Fig. 5 A schematic depiction of an ideal hypothetical chip. The device density is maximum at the interior and falls off towards the edges. The devices in the interior of the chip are self-assembled and the I/O connections on the periphery are made by nanolithography.

The presence of a negative charge on the lower electrode and a negative charge on the particles explains why they float above the electrode. But, there is one more consideration involved: the range of the electric field from the charged electrode into the plasma.

The negative charge on the electrode repels free electrons from the region above it. A positive charge layer (the now-uncompensated positive ions) is left. This positive charge tends to cancel the negative charge on the electrode, reducing the electric field to nearly zero a short distance above the electrode.

Thus, as the negative particle falls toward the lower electrode, it encounters little upward-directed electric force, qE, until it enters this region. Then the force rises rapidly. The equilibrium height is that position where the qE force is equal but opposite to the downward-directed force. The downward force is produced by gravity, mg, and also by the ion "wind" near the electrode.

Fencing the particles in

The negatively charged particles levitated just above the lower electrode will repel each other and flow off the surface. To form a Coulomb crystal, they must be confined to a region just above the electrode. There are several ways to build a fence to do that.

A popular method is to machine a shallow well in the electrode or to put a metal ring on

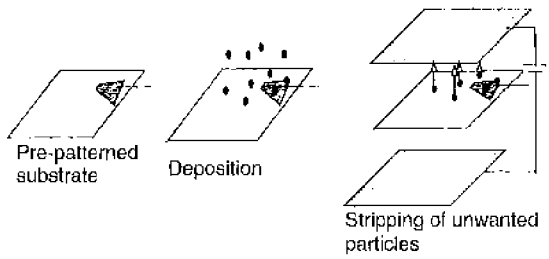


Fig. 6 A schematic depiction of "nanoxerography," a scheme for selective area transfer of a Coulomb crystal.

International technology roadmap for semiconductors

The International Technology Roadmap for Semiconductors (ITRS) is an assessment of what is needed to face the technological challenges over the next 15 years in the semiconductor industry. It prescribes a list of objectives to be met in order to meet future challenges. ITRS is sponsored by the US Semiconductor Industry Association (SIA), the European Electronic Component Association (EECA), the Electronic Industry Association of Japan (EIAJ), the Korean Semiconductor Industry Association (KSIA) and Taiwan Semiconductor Industry Association (TSIA). The ITRS can be found at the website address <http://www.itrs.net/ntrs/publntrs.nsf>.—*SB & PFW*

the electrode. Instead, we put a glass ring on top of the electrode. The glass charges negatively in the plasma for the same reason that the particles do. As a result, it repels the particles that stray too close.

Inter-particle separation

Probably the most serious obstacle for using Coulomb crystal technology to self-assemble nanostructures is the distance between neighboring particles in the crystal. The crystals in most reports have a lattice constant of about 100 μm. That means that the average distance between particles is 1/10th of a millimeter, which is clearly too large. For most applications of nanostructures, this separation should be reduced substantially to 1 μm or less.

What physical mechanisms determine the separation? The answer is not well understood, but it is known that electric field shielding plays an important role. The negatively charged particle repels electrons from the immediate neighborhood. A shell (a *sheath*) of uncompensated positive ions is left. Like the region near the lower electrode, this sheath acts to shield the plasma from the field produced by the particle.

The distance over which this shielding occurs is called the *Debye length*, λ_D, or the *screening length* and is given by

$$\lambda_D = \sqrt{\frac{\varepsilon_0}{n_0 q_e^2} kT} \qquad (1)$$

Here n_0 is the plasma density (density of electrons or positive ions), and T is the "parallel" combination of the ion and electronic temperatures,

$$\frac{1}{T} = \frac{1}{T_i} + \frac{1}{T_e}$$

Typical plasmas utilized for creating Coulomb crystals use rather low R.F. power. The plasma density is about $10^8 cm^{-3}$. Using that density, and assuming room temperature, λ_D is about 100 μm.

Returning to the question of lattice constant, consider the force between two particles vs. the separation between them. Because of the exponential fall-off of the field, there is little force between them for separations exceeding a few Debye lengths. The repulsive force rises rapidly, however, as the particles come closer together. Thus, we would expect that the lattice constant in a Coulomb crystal would be of the order of one or two Debye lengths. This is what, in fact, is found experimentally—at least for plasmas with density of about 10^8 cm^{-3}.

Reducing the lattice constant

If it is agreed that the lattice constant of a Coulomb crystal is about λ_D, then we must reduce λ_D if we are to reduce the lattice constant. Looking at Eq. (1), it appears that the only parameters we can control are n_0, and T. Reducing T much below room temperature is not very promising, so we must increase the plasma density, n_0. Because of the dependence on the square root, to reduce λ_D from 100 to 1 μm would require an increase in the plasma density by about 10^4. That means a plasma density of about 10^{12} cm^{-3}. Is that possible? The answer is yes, but doing so introduces some other issues.

For a given plasma cell, what factors influence the plasma density? The answer to that question is not simple. If the power to plasma is turned off, the density of electrons and positive ions will decay to zero due to recombination. The recombination rate is proportional to the density of electrons and of positive ions. Thus, the higher the density of the plasma, the higher the recombination rate, and the more power must be supplied to maintain it. This power mostly winds up as heat. Thus, the plasma density can be increased with turning up the power, but doing so results in a hotter plasma.

For a fixed plasma density, the recombination rate is typically proportional to the pressure. The reason is that recombination tends to be a three-body process. When an electron and a positive ion collide it is unlikely that the ion will capture the electron because it is difficult to satisfy both momentum and energy conservation. Thus, recombination is unlikely in such an event. Except at very low pressures, recombination usually involves a third body, typically a neutral atom. At a fixed plasma density, such a recombination rate is proportional to the neutral atom density and, hence, to the pressure. Thus, for a fixed input power we can increase the plasma density by reducing the pressure.

Practice bears out these expectations. High-density plasmas (density exceeding 10^{12} cm^{-3}) can be produced at pressures near atmospheric using high power excitation. Such plasmas are very hot and, in fact, are often called *plasma torches*.

On the other hand, low-temperature plasmas with densities of 10^{12} cm^{-3} or a bit more can be created at pressures between 1 and 10 mTorr. At such low pressures, the problem is more about getting the electrons to collide frequently enough with the neutrals to produce the needed ionization, rather than the power level required match the recombination rate.

Transferring the pattern

Once the Coulomb crystal has been created in the plasma just above the substrate, the pattern must be transferred to the substrate. One technique involves simply turning the R.F. excitation off and allowing the particles to fall onto the substrate. In that case, some

method must be used to fix the particles in place. In our laboratory, we usually pre-coat the substrate with a sticky substance. This way the particles stick where they land.

An alternative method (which we don't think has been shown yet), involves applying a fast-rising, positive voltage pulse to the lower electrode. The resulting net qE force will be downward, and will pull the particles toward the substrate. This is the so-called "electrostatic chuck" method that has found applications in other areas. With this method, it should be possible to transfer a Coulomb crystal to a substrate without significantly disrupting the spatial order of its particles.

The perfect match

A holy grail in the world of nanofabrication is the "marriage" of self-assembly with lithography. An ideal chip will have the configuration shown in Fig. 5.

All the densely packed device elements will be in the interior of the chip and produced by self-assembly. All the input/output ports for connection to the world will be on the periphery of the chip. Connections to the input/output (I/O) ports must be made with lithography since these are arbitrary connections. If the device elements are smaller than 10 nm, it is very difficult, if not impossible, to align the connecting wires to the I/O devices by lithography.

In this case there is only one solution– the I/O devices will have to be aligned with the connecting wires and not the other way around. Hence, the contacts are first delineated on the chip by lithography and then the I/O device is self-assembled right on the contact. This accomplishes some sort of a happy marriage between lithography and self-assembly.

How does one self-assemble a nano-sized device right on the contact? One can charge the contact temporarily to a positive voltage that will attract the negatively charged particles, levitating in the plasma, directly to the contact. Thus, the particles will land selectively, or at least preferentially, on the contact.

The charge on the contacts will neutralize the negative charge on the particles that land on them. The remaining particles that do not land on the contact, but land elsewhere on the substrate, are still negatively charged. They can be removed by passing the wafer under a positively charged stripping plate. This leaves particles on the contact areas and nowhere else.

What we have described is a selective area transfer of the Coulomb crystal. This is the first step towards a hybrid nanofabrication scheme employing both self-assembly and lithography. We have coined a term for this process; we call it "nanoxerography" since that is more or less how a Xerox photocopier machine works. Figure 6 shows a schematic diagram of the nanoxerography process. Such techniques can ultimately lead to the marriage of self-assembly with nanolithography.

Conclusion

Electrical engineers have only begun to explore techniques for fabricating ultra small, high-performance chips. These are exciting times for solid state devices and electronics.

Read more about it

• R.P. Feynman, "There is Plenty of Room at the Bottom," in *Miniaturization*, H.D. Gilbert, ed. (Reinhold Publishing Corp., New York, 1961), pp. 282-296.

• S. Bandyopadhyay, V.P. Roychowdhury and D.B. Janes, "Chemically Self Assembled Nanoelectronic Computing Networks," in *Quantum Based Electronic Devices and Systems*, M. Dutta and M.A. Stroscio, eds., *Selected Topics in Electronics and Systems*, Vol. 14 (World Scientific, Singapore, 1998), Chapter 1, pp. 1-35.

• J.B. Pieper, J. Goree and R.A. Quinn, "Experimental Studies of Two-Dimensional and Three-Dimensional Structures in a Crystallized Dusty Plasma," *J. Vac. Sci. Technol.* A, 14, 519-24 (1996).

• S. Bandyopadhyay and V. P. Roychowdhury, "Granular Nanoelectronics: The Logical Gateway to the 21st Century," *IEEE Potentials*, April/May issue, 8-11 (1996).

About the authors

Supriyo Bandyopadhyay is a Professor of Electrical Engineering at the University of Nebraska-Lincoln where he tinkers with various self-assembly techniques for making nanostructures. He directs the Quantum Device Laboratory where some important strides have been made towards making novel electronic, optical and magnetic devices based on self-assembled nanostructures. He collaborates with Professor Williams and others at the Lincoln, Omaha and Kearney campuses of the University of Nebraska in developing Coulomb crystallization into a viable self-assembly technique for making nanoscale devices.

Frazer Williams is the Lott College Professor of Electrical Engineering at the University of Nebraska-Lincoln. His current research interests include plasma processing of semiconductors, atmospheric-pressure glow-like discharges, and numerical modeling of discharges. In 1996, he directed a NATO Advanced Studies Institute on the topic of plasma processing of semiconductors. [*Plasma Processing of Semiconductors*, P.F. Williams, ed. (Kluwer, Dordrecht, 1997).] He also regularly tortures his neighbors in an apparently hopeless attempt at learning to play the piano.

Richard Feynman

Richard Feynman (1918-1988) was an American physicist who made major contributions in such diverse areas of physics as color vision, and the behavior of liquid helium. He won the Nobel Prize in 1965 for his work on Quantum Electrodynamics. Gleick describes him as follows.

Architect of quantum theories, brash young group leader on the atomic bomb project, inventor of the ubiquitous Feynman diagram, ebullient bongo player and storyteller, Richard Phillips Feynman was the most brilliant, iconoclastic, and influential physicist of modern times. He took the half-made conceptions of waves and particles in the 1940s and shaped them into tools that ordinary physicists could use and understand. He had a lightning ability to see into the heart of the problem's nature posed. Within the community of physicists, an organized, tradition-bound culture that needs heroes as much as it sometimes mistrusts them, his name took on a special luster. It was permitted in connection with Feynman to use the word *genius*. He took center stage and remained there for 40 years, dominating the science of the post-war era—40 years that turned the study of matter and energy down an unexpectedly dark and spectral road.

Perhaps his character is best described by his last words. As he lay dying he is supposed to have said, "I'd hate to die twice. It's so boring."—*SB & PFW*

(Source: James Gleick, *Genius*, Vintage Books, New York, 1992)