

A reprint from

American Scientist

the magazine of Sigma Xi, The Scientific Research Society

This reprint is provided for personal and noncommercial use. For any other use, please send a request Brian Hayes by electronic mail to bhayes@amsci.org.

A Box of Universe

Brian Hayes

ISAAC NEWTON'S UNIVERSE was a cozy, tidy place. Gathered around the sun were six planets, a handful of moons and the occasional comet, all moving against a backdrop of stationary stars. Newton provided us with the mathematical tools needed to compute the motions of these bodies. Given initial positions and velocities, we can calculate the forces acting on each object, using Newton's law of universal gravitation. From the forces we can determine accelerations, and then update the positions and velocities for the next round of calculations. This scheme of computation is known as the *n*-body method. Perhaps Newton himself could have put it to work if he had had suitable computing machinery.

Today we have the computers. On the other hand, our universe is far larger and more intricate than Newton's. Now the solar system is merely a speck in a spiral galaxy of several hundred billion stars. Our galaxy drifts among billions of others, which form clusters and superclusters and a whole hierarchy of structures extending as far as the eye (and the telescope) can see. Those objects are getting farther away all the time because the universe is expanding, and moreover the expansion is accelerating. Strangest of all, the luminous matter of the galaxies—everything we see shining in the night sky—makes up less than one-half of 1 percent of what's out there. Most of the universe is unseen and unidentified stuff known only as "dark matter" and "dark energy."

Given this profound change in the nature and the scale of the known universe, I find it remarkable that computer simulations of cosmic evolution can

*Watch the cosmos
evolve in a cube
one billion
light-years wide*

still rely on *n*-body algorithms rooted in the principles of Newtonian mechanics. The same techniques that predict planetary motions here at home in the solar system also describe the gravitational process that assembles thousands of galaxies into filaments a hundred million light-years long.

A major new series of cosmological simulations is now beginning to release its findings. The project, known as Bolshoi, is led by Anatoly Klypin of New Mexico State University and Joel Primack of the University of California, Santa Cruz. "Bolshoi" is Russian for "big" or "grand," and the name is apt: This is a large-scale computational project, consuming six million CPU hours and producing a hundred terabytes of data. And yet, when you ponder the vast sweep of space and time being modeled, it seems a marvel that so much universe can be squeezed into such a small box.

A History of the History of the Universe

Our view of the universe—and of our own place in it—was famously upset by the Copernican revolution of the 16th century. But the past 100 years have seen even more radical upheavals in cosmology.

It began in the 1920s with the recognition that spiral "nebulae" are not dusty pinwheels scattered among the nearby stars; the nebulae, which we now call galaxies, are *made* of stars, billions of them, and they lie at immense

distances. Furthermore, we too inhabit just such a galaxy.

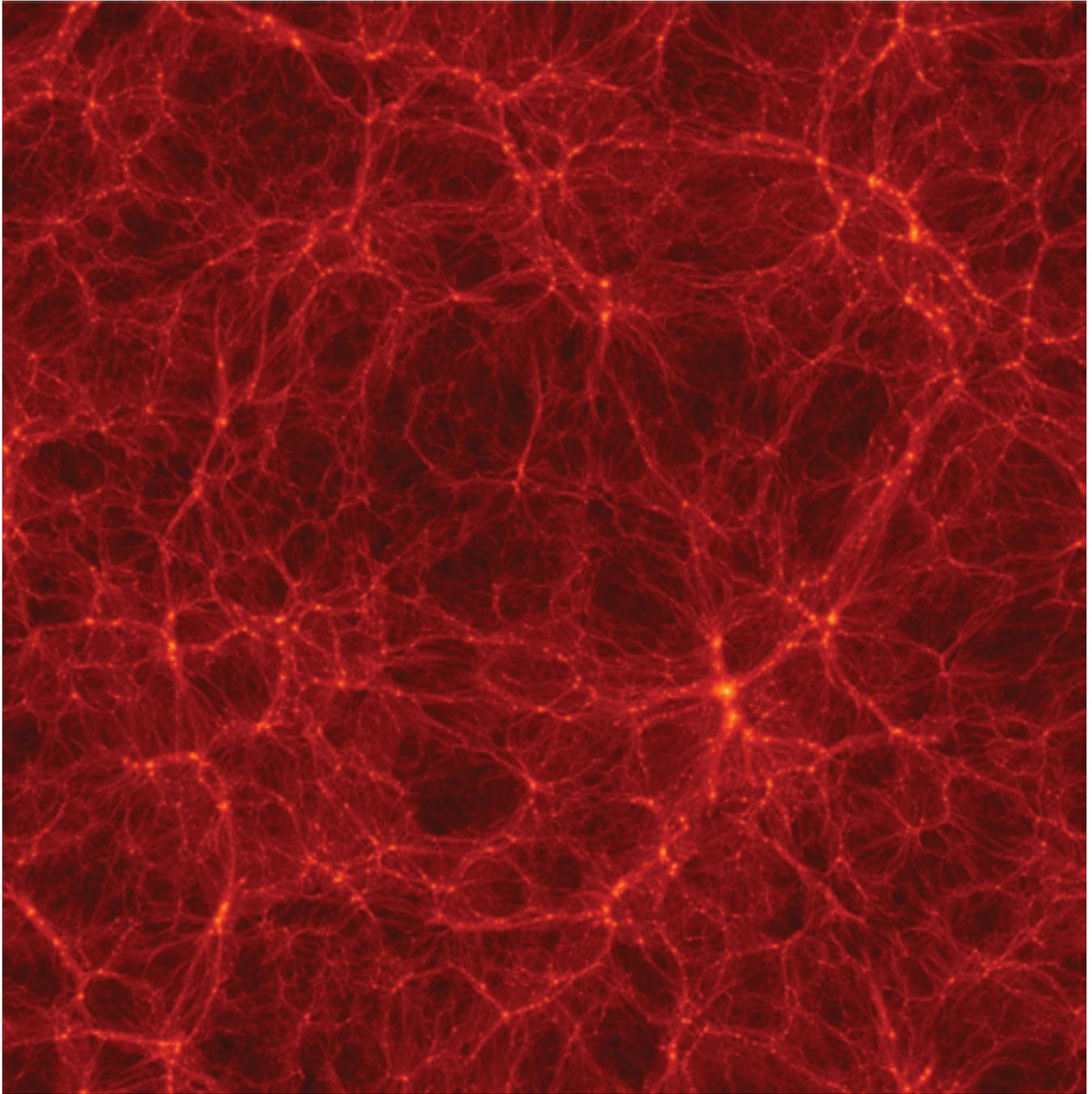
Edwin P. Hubble, examining the spectra of galaxies, found that almost all of them are shifted toward the red end of the spectrum. The redshifts indicate that the galaxies are moving away from us, and Hubble showed that their velocity is proportional to their distance. The whole universe is expanding.

The expansion could be traced backward to a moment of origin, the Big Bang, which meant the universe must have a definite age (now estimated at 13.7 billion years). The clinching evidence for the Big Bang model was the discovery in 1965 of cosmic background radiation—a relic of the early, incandescent universe, now cooled to a temperature of 2.7 kelvins, with peak emission in the microwave region of the electromagnetic spectrum.

By the 1970s the Big Bang model was firmly established, but nagging problems remained. Looking around us, we see a well-stirred porridge—the same in all directions—and yet the universe is too young to be so thoroughly mixed. There are regions that have never communicated because speed-of-light signals have not yet had time to pass between them. How did these disconnected pieces all come to have the same temperature? Explaining the uniformity called for another dramatic revision in the history of the early universe: a split-second episode of "cosmic inflation" so rapid that a tiny (and therefore nearly uniform) patch of space expanded to become what is now our entire observable universe.

Then there was the mystery of the missing mass. The orbital velocities of stars within a galaxy depend on the total mass of the galaxy and how that mass is distributed. On a larger scale, the motions of galaxies within a cluster also depend on the cluster's mass. Measurements of both kinds of velocities sug-

Brian Hayes is senior writer for American Scientist. Additional material related to the Computing Science column appears at <http://bit-player.org>. Address: 11 Chandler St. #2, Somerville, MA 02144. E-mail: brian@bit-player.org



A thin slice of simulated universe reveals a tangle of filaments and membranes on an immense scale. The dimmer red spots are the size of entire galaxies, and the brightest nodes along the major filaments are large enough to envelop clusters of hundreds or thousands of galaxies. The slice is taken from a cubic volume roughly one billion light-years on a side; the thickness (looking into the page) is about 35 million light-years. In computer simulations the lacy structures arise spontaneously when matter is allowed to collapse under its own gravitational attraction; similar forms have been observed in surveys of the real sky. The data were generated in the Bolshoi cosmological simulation, which examines the distribution of cold dark matter, a component of the universe distinct from the ordinary “baryonic” matter of stars and planets. The illustration was created by Stefan Gottlöber of the Leibniz Institute for Astrophysics in Potsdam, Germany.

gested there is more matter out there than meets the eye. Galaxies must be enveloped in huge halos of “dark matter,” unseen because it neither emits nor absorbs electromagnetic radiation. The nature of this substance remains uncertain, but it apparently constitutes 80 percent of all the matter in the universe.

A little more than 10 years ago came yet another big surprise. Two groups of astronomers had set out to measure changes in the rate of the cosmic expansion over the history of the universe. Gravitational attraction acts to slow the expansion, and might even cause an eventual reversal and collapse. But the

measurements showed that the expansion is actually accelerating: Something is pushing space apart. That something has been given the name “dark energy,” and it appears to be the majority constituent of the universe.

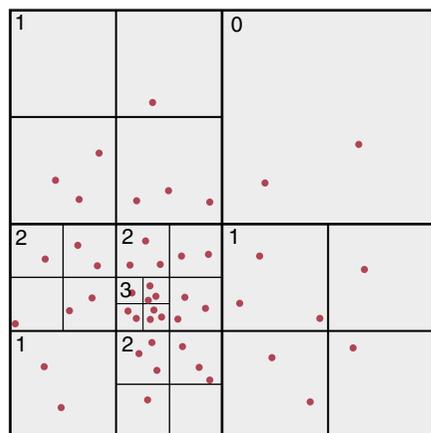
All of these ideas are elements of the current consensus view of cosmology

called the Λ CDM model. The Greek letter Λ (lambda) is the symbol Einstein chose for his “cosmological constant,” which offers one way of understanding the dark-energy phenomenon. CDM stands for cold dark matter. The matter is cold in the sense that it consists of massive particles moving at modest speeds, rather than a high-temperature gas of nearly massless particles.

In the Λ CDM model, almost 73 percent of the substance of the universe is dark energy, and about 23 percent is dark matter. Thus more than 95 percent of the universe is stuff that wasn’t even known to exist a mere 30 years ago. Our own substance, called “baryonic” matter after the class of particles that includes the proton and the neutron, makes up 4.6 percent of the total. Most of that is hydrogen gas in the intergalactic medium. The stars in galaxies account for 0.4 percent.

Observation, Theory, Simulation

All of these head-spinning changes in our understanding of the cosmos are anchored in scientific knowledge. They may seem like one wild idea after another, but they are supported by data from observations, by physical theory and by simulations that connect theory with observation.



Adaptive refinement tree allows efficient calculation of gravitational interactions among billions of masses. Rather than compute pairwise forces between individual particles, the algorithm partitions space into cubical cells and then evaluates interactions between cells. Each cell is subdivided until the number of particles inside falls below a threshold; adjacent cells may not differ by more than one subdivision level. In the two-dimensional example shown here the threshold is three particles. The Bolshoi simulation began with 16.8 million level-0 cells and allowed 10 levels of subdivision.

In the past 20 years astronomy has become a data-rich, statistical science. For example, the Sloan Digital Sky Survey has catalogued 500 million objects and recorded almost two million spectra. The spectra allow measurements of redshifts and thus of distance along the line of sight. The data yield a three-dimensional map that covers about a third of the sky and goes back in time more than a billion years.

Another ongoing endeavor aims to extract information from the oldest photons in the universe. A satellite called the Wilkinson Microwave Anisotropy Probe (WMAP) has measured tiny spatial variations in the microwave background radiation, giving us a glimpse of the distribution of matter and energy at an early epoch—the initial conditions for the visible universe.

All these data are interpreted in the context of Einstein’s general theory of relativity, which describes the interactions of matter, energy, space and time in an expanding universe. Quantum mechanics also has a role, for example in explaining the spectrum of fluctuations in the microwave background. But these theoretical principles are not enough to predict the shapes and sizes and other properties of the structures that emerge as the universe evolves. Computer simulation is the best tool for this purpose. Starting from plausible initial conditions and known physical laws, we can compare the output of the simulation with astronomical data. Along the way, we get to watch a movie of the universe unfolding.

As early as 1941 the Swedish astronomer Erik Holmberg simulated the clustering of galaxies with a remarkable analog computer he made out of light bulbs and photocells. By the early 1960s digital computers were the instrument of choice. Sverre J. Aarseth of the University of Cambridge studied galaxy clusters having 25 to 100 members.

Modern computing machinery allows much larger models. In 2005 Volker Springel of the Max Planck Institute for Astrophysics in Garching, Germany, and 16 colleagues published the results of a landmark simulation called the Millennium Run. It was based on the Λ CDM model; it included more than 10 billion particles of dark matter in a cubical volume more than 2 billion light years on a side; it covered a span of time from 12 million years after the Big Bang up to the present. Unfortunately, some of the parameters that defined the initial con-

ditions were based on early results from the WMAP satellite mission, which were later significantly revised.

The Bolshoi simulation (described in more detail below) is slightly smaller than the Millennium Run but attains higher resolution in measuring positions, masses and forces. It is also based on the updated WMAP parameters.

2-Body, 3-Body, N-Body

Even though cosmology is rooted in the general theory of relativity, the simulated universes of Millennium and Bolshoi operate on Newtonian principles. Here gravity is not a warping of space-time; it’s a force between particles. How can we get away with this retro vision of physics? Relativistic methods would be crucial if objects were moving near the speed of light, but galaxies travel at a stately few hundred kilometers per second. Newton’s equations are perfectly adequate.

Newton solved the two-body problem: Given the positions, velocities and masses of two pointlike particles interacting only with each other, he could describe the particles’ trajectories for all times, both past and future. But the three-body problem was beyond the reach of Newton’s methods. As for the n -body problem (with n larger than three), he speculated that “to define these motions by exact laws allowing of convenient calculation exceeds, unless I am mistaken, the forces of the entire human intellect.”

Newton was not mistaken: For the n -body problem, exact solutions are indeed unthinkable. No mathematical expression defines the positions and velocities of the particles for all time. However, with enough computer power we can have a “marching” solution, which takes the state of the system at one moment and walks forward a short distance along all the trajectories to yield a new state a moment later. The result is an approximation, but if the time steps are short enough, it can be highly accurate.

The inputs to an n -body calculation are the masses of the n particles (which never change) and their spatial coordinates and velocities at some initial time t . Each particle feels a gravitational attraction to every other particle. For particles i and j the magnitude of this force is given by Newton’s equation $F = Gm_i m_j / r^2$, where m_i and m_j are the particle masses and r is the distance between them. (The constant G merely

reconciles units of mass and length.) The total force acting on a given particle is the vector sum of all the pairwise forces.

Another Newtonian equation relates force to acceleration: $F = ma$. Once we know a particle's acceleration, we can assign it a new velocity and then calculate its motion over some brief interval Δt . This process yields a new set of particle positions and velocities at time $t + \Delta t$, which serve as inputs to the next round of computations.

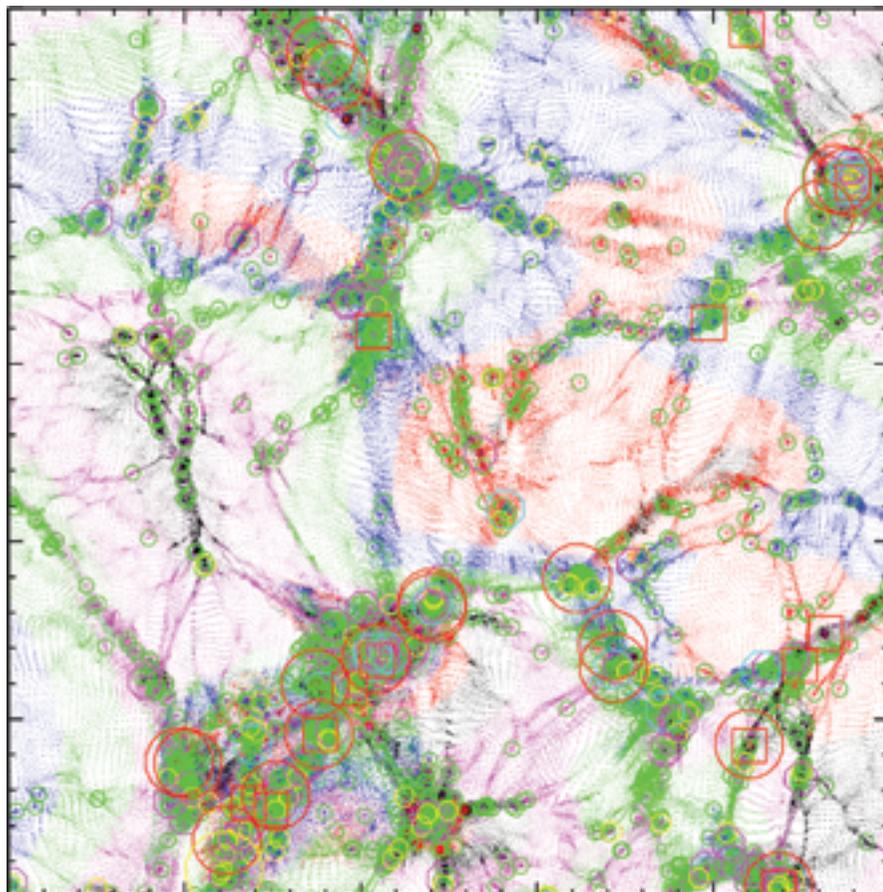
In cosmological n -body simulations, n can be a number in the billions, which imposes a formidable computational burden. There are $n(n-1)/2$ pairs of particles requiring a force calculation at each time step. With 10^9 particles, that's 5×10^{17} evaluations of the force equation, far beyond the limits of computer capacity, not to mention human patience. Thus the straightforward algorithm described above is not much use in practice.

The key to solving this problem is to lump together groups of nearby particles, summing their contribution to the gravitational field. This coarsening of the system has a small cost in accuracy but a major benefit in speed. Whereas the direct particle-particle method requires on the order of n^2 calculations, a hierarchy of lumped masses offers hope of reducing the computation time to $n \log n$. For the billion-particle case, that's the difference between 10^{17} evaluations and 10^{10} .

The Bolshoi simulation employed an algorithm called the adaptive refinement tree, devised in the 1990s by Andrey V. Kratsov, now of the University of Chicago, Klypin and Alexei M. Khokhlov of the U.S. Naval Research Laboratory. The procedure starts by dividing the cubical simulation volume into a grid of smaller cubical cells. Each of these level-0 cells can be further divided into eight level-1 cells, each of which can be split into eight level-2 cells, and so on. The splitting process continues until the number of particles in a cell falls below a threshold. The subdivision level can vary across the grid, so that costly high-resolution computations are done only where needed; but adjacent cells are allowed to differ by no more than one level, to avoid sharp discontinuities.

The Bolshoi Intergalactic Ballet

The domain of the Bolshoi simulation is a cube roughly one billion light-years on a side. For comparison, the



Halo-finding algorithm identifies dense concentrations of particles that are gravitationally bound to one another. Dark-matter halos are the sites where galaxies live; they are called halos because they extend beyond the perimeter of the galaxy. Since the Bolshoi simulation includes only dark matter, the galaxies themselves are not explicitly present. In this thin slice through the final state of the Bolshoi simulation, halos with various properties are marked by circles and polygons. Just as individual particles fall together to form halos, the halos also tend to merge, forming a nested hierarchy. The image was created by Stefan Gottlöber.

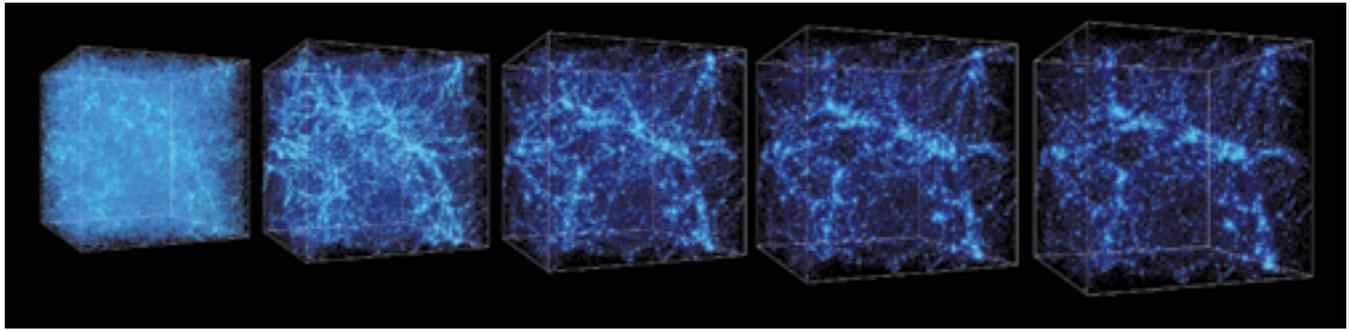
diameter of the luminous part of a large galaxy might be 100,000 light years, and the diameter of the observable universe is about 90 billion light-years. Thus the simulation volume is large enough to contain millions of galaxies, but it still represents only a tiny fraction of the visible universe.

The level-0 mesh of the adaptive refinement tree has 256^3 (or 16.8 million) cells. These cells can be subdivided to a maximum of 10 levels, at which point the side length of the smallest cells is about 4,000 light years.

What happens at the edges of this big box of universe? Rather than create artificial barriers or let particles leak out, the simulation adopts periodic boundary conditions. If a particle leaves the box on one side, it immediately reenters through the opposite side. This policy avoids disruptive edge effects, but it also introduces another hazard: a sort of gravitational

narcissism, where an object is attracted to its own image a billion light-years away. The remedy is to impose a long-range cutoff, suppressing gravitational forces beyond about half this distance. A short-range force cutoff is also needed, to prevent spurious strong scattering if two particles happen to come close together.

The Bolshoi cube is populated with $2,048^3$ (equal to 8,589,934,592) identical particles. This is a great many objects to be managed in a computer program, but on the other hand it is a very coarse partitioning of the substance in the universe. Each particle represents about 200 million solar masses. Because this quantum of mass is so large, there's no point in trying to distinguish between baryonic matter and dark matter. All of the particles in the simulation are dark matter. The primary structures formed as the universe evolves are the dark-matter halos in



Snapshots from the evolution of a simulated universe show the gradual emergence of large-scale structures as particles come together under their mutual gravitational attraction. In the leftmost panel the initial state of the system, representing a moment roughly 100 million years after the Big Bang, has particles distributed almost uniformly, with only slight fluctuations in density. Over the subsequent 13 billion years the particle population condenses in filaments, some of which extend across the entire volume. The rightmost panel shows the final configuration. The illustration is not from the Bolshoi project but from an earlier simulation carried out at the National Center for Supercomputer Applications by Andrey V. Kravtsov of the University of Chicago and Anatoly Klypin of New Mexico State University. The visualization was created by Kravtsov.

which galaxies are embedded, but the galaxies themselves are not explicitly represented.

At the start of the run, the eight billion particles are distributed almost uniformly in the cube, with only slight fluctuations in density. This configuration represents the state of the universe post-inflation, in an era not too long after the cosmic background radiation was emitted. The simulation begins when the universe is about 23 million years old and progresses to the present in some 400,000 time steps. Snapshots of intermediate states are saved at intervals of 40 to 80 million years.

The simulation was run on a computer called Pleiades at the NASA Ames Research Center in California. Pleiades is currently ranked seventh in the listings of the top 500 supercomputers worldwide. Bolshoi made use of 13,824 processor cores and 13 terabytes of memory.

When a simulation run ends, the computing is not yet finished. The first stage in analyzing the data is to run a program called a halo finder, which identifies regions of elevated density where particles are bound together by gravitational attraction. Finding density peaks is not difficult, but identifying bound groups of particles is a subtle problem. The particles in a halo move like bees in a swarm, so it's not obvious which particles are members and which are merely passing through.

Evolution of a Universe

The first results from the Bolshoi simulation appear in a series of papers by Klypin, Primack and their colleagues. (The first two papers have been published in *Astrophysical Journal* and are also available as preprints at the Bol-

shoi web site, <http://hipacc.ucsc.edu/Bolshoi>.) The data generated by the simulations, including the snapshots, catalogs of halos, and "merger trees" tracing the provenance of halos, are also being made publicly available through a repository called the MultiDark database.

What does the Bolshoi universe look like, and how does it compare with the world we see when we look out the window? Early in the course of the simulation, some five million halos evolved, and the number later grew to 12 million as additional particles clumped together under their own gravitational attraction. But the halos too experience these forces, and so they tend to merge, leaving fewer but larger halos. At the end of the run, there were about 10 million halos left (many of which are classified as "sub-halos," existing inside other halos but retaining their identity). Halo sizes range from 10^{10} solar masses (comparable to a small galaxy) up to 10^{15} solar masses (equivalent to a large cluster of galaxies).

Statistics on the halos and the larger structures they form suggest opportunities for comparing the simulation with reality. This work is just beginning. An example concerns a survey of satellite galaxies. The Milky Way has two prominent satellites, the Large and Small Magellanic Clouds of the Southern Hemisphere. Observational data from the Sloan Digital Sky Survey suggest that this situation is somewhat rare; only about 10 percent of galaxies in the mass range of the Milky Way have two such companions. A study of the Bolshoi results by Michael T. Busha and others yields a similar probability.

The model and real-world data also agree about the probability of zero satellites or one satellite.

But not every cross-check between model and reality comes out so neatly. Sebastian Trujillo-Gomez and his colleagues have compared a sample of real galaxies with simulated dark-matter halos classified according to rotational velocity, a property that is closely correlated with total mass and luminosity. Over most of the range, the two samples are in accord. But Bolshoi predicts a few too many galaxies at the highest rotational velocities. More worrisome is a larger discrepancy at the other end of the scale: The simulation predicts an overabundance of dwarf galaxies by a factor of almost 10.

For now this conflict remains unresolved. There are several candidate explanations. The cause might turn out to be a problem in the simulation. On the other hand, the observational survey could have some unrecognized selection bias, and the Bolshoi result could be correct. Or the cause might lie at a deeper theoretical level. For example, the difference might be explained if the dark matter pervading the universe is not cold but lukewarm. If this last possibility begins to look promising, then we could be on the cusp of yet another major shift in cosmological thinking.

Bruno's Legacy

Four hundred years ago, the idea that the Earth goes around the Sun rather than vice versa was not just a scientific breakthrough but also a cultural bombshell. People were asked to re-imagine the world they were living in. Not everyone welcomed the op-

portunity. Books were burned. In the case of Giordano Bruno, an author was burned.

In the modern world, cosmological revolutions seem to cause hardly a ripple in public consciousness. Inflation, dark matter, dark energy—these ideas also call for a reimagining of the world we live in, but they have provoked very little fuss outside the community of science. It's certainly a relief that no one will be burned at the stake over matters of cosmological doctrine. But are we really more liberal and open-minded, or just not paying attention?

Bibliography

- Aarseth, S. J. 1963. Dynamical evolution of clusters of galaxies, I. *Monthly Notices of the Royal Astronomical Society* 126:223–255.
- Behroozi, P. S., et al. 2011 preprint. Gravitationally consistent halo catalogs and merger trees for precision cosmology. <http://arxiv.org/abs/1110.4370>.
- Bertschinger, E., and J. M. Gelb. 1991. Cosmological N-body simulations. *Computers in Physics* 5:164–179.
- Blumenthal, G. R., et al. 1984. Formation of galaxies and large-scale structure with cold dark matter. *Nature* 311:517–525.
- Busha, M. T., et al. 2010 preprint. Statistics of satellite galaxies around Milky Way-like hosts. <http://arxiv.org/abs/1011.6373>.
- Callan, C., R. H. Dicke and P. J. E. Peebles. 1965. Cosmology and Newtonian mechanics. *American Journal of Physics* 33(2):105–108.
- Gottlöber, S., and A. A. Klypin. 2009. The ART of cosmological simulations. In *High Performance Computing in Science and Engineering*, edited by S. Wagner et al., pp. 29–44. Berlin: Springer.
- Holmberg, Erik. 1941. On the clustering tendencies among the nebulae. II. A study of encounters between laboratory models of stellar systems by a new integration procedure. *Astrophysical Journal* 94:385–395.
- Klypin, A. 2002. Numerical simulations in cosmology. In *Modern Cosmology*, edited by S. Bonometto, V. Gorini and U. Moschella, pp. 420–473. Philadelphia: Institute of Physics.
- Klypin, A. A., S. Trujillo-Gomez and J. Primack. 2011. Dark matter halos in the standard cosmological model: Results from the Bolshoi simulation. *Astrophysical Journal* 740:102.
- Kravtsov, A. V., A. A. Klypin and A. M. Khokhlov. 1997. Adaptive Refinement Tree: A new high-resolution *n*-body code for cosmological simulations. *Astrophysical Journal Supplement* 111:73–94.
- Riebe, K., et al. 2011 preprint. The MultiDark database: Release of the Bolshoi and MultiDark cosmological simulations. <http://arxiv.org/abs/1109.0003>.
- Springel, V., et al. 2005. Simulations of the formation, evolution and clustering of galaxies and quasars. *Nature* 435:629–636.
- Trujillo-Gomez, S., et al. 2011. Galaxies in Λ CDM with halo abundance matching: Luminosity-velocity relation, baryonic mass-velocity relation, velocity function, and clustering. *Astrophysical Journal* 742:16.