Simulating the Universe on Supercomputers

Mark Vogelsberger, `mark.vogelsbergerATemail.de`

The following text is a very brief introduction into the field of cosmological Supercomputer simulations. Those who want to dig deeper into the field should consult the references at the end.

1 The Universe

The goal of cosmological simulations is to model the growth of the structures in the Universe. In other words, these simulations allow us to compress the long times of cosmic evolution into a human lifetime and they can be considered as an experimental tool to verify theories of the origin and the evolution of our Universe.

Today we believe that this evolution started with a Big Bang. Shortly after this event small fluctuations were imprinted into the radiation and matter density field. To understand the Universe, how it looks today, we need to know how these small perturbations to an otherwise homogeneous and isotropic space evolve with time. This calculation is highly complex and can only be done numerically using large computers. Analytic methods can only be used in the linear regime but for the whole evolution of the Universe numerical methods are needed. To run such cosmological simulations one needs two main ingredients: first it is necessary to specify initial conditions, to tell the computer where it should start to calculate. On the other hand one has to tell the computer also how to calculate the evolution of the Universe. The initial conditions for the simulation can be observed. How can we do this? We get the initial conditions from the afterglow of the Big Bang. About 300.000 years after the Big Bang the radiation could decouple. This radiation is still visible today. Due to the expansion of the Universe we can observe it today at an temperature of about 2.7 Kelvin. Modern satellite missions could resolve small fluctuations in this radiation. From these fluctuations it is possible to infer the perturbations in the initial density field of the matter. Thus we know how the initial density field 300.000 years after the Big Bang looked like. This is the input of our simulation. From this initial density field we have to evolve the Universe from the starting point to today, about 13 billion years after the Big Bang.

The leading force for this evolution is gravity in an expanding space. Cosmological codes use particles to trace the density field and evolve them under their mutual gravity. As the simulation samples the smooth density field with such a finite set of particles these computer simulations are called N-body codes. The more particles you have the better the resolution you get. This is why there is a constant competition in getting the highest number of particles and the computational resources you need to run these calculations require the largest computers available today. I will focus here on the simulation of the gravity only. This is by far the most important process and also the easiest thing to simulate. Note that there is also baryonic gas in the Universe - we are for example made out of baryons. Everything you can see like stars, galaxies, planets and so on are made of baryons. Their dynamics is also influenced by hydrodynamics and complicated gas physics. This is a lot more complicated to deal with. Modern simulation codes are also able to treat the baryons and compute a Universe with galaxies.

They allow to form stars and solve the gas physics. The cosmological code Gadget (Springel, 2005) that was developed at our institute is public available and can solve both gravity and hydrodynamics. This is still quite restricted, because there are lots of processes going on that need to be taken into account to get more realistic pictures: black holes, cosmic rays, radiative transfer, magnetic fields and so. The current internal production version of the Gadget code has more than 200 options corresponding to physical processes you can turn on or off. But the main evolution of cosmic structure does not need gas physics. It can purely be calculated using the gravitational force in an expanding Universe.

The fact that we can ignore the baryons for structure formation is because they only make up four percent of the total energy content in the Universe. The largest mass component comes from what is known as Dark Matter. It is called dark, because it does not shine like stars or gas. It is invisible and therefore called dark. Today we know that about 23 percent of the Universe are made up of this Dark Matter. Dark Matter only interacts by gravitation. This is why we can indirectly observe it by its gravitational interaction on visible objects like galaxies and gas. For example, Dark Matter can act as a gravitational lens and can deflect light from visible galaxies. Besides baryons and Dark Matter the largest component of the Universe consists of Dark Energy. In Einstein's equations of general relativity this corresponds to the so called cosmological constant. Due to the small fraction of baryons in the Universe most simulations of structure formation only take into account the dark components, so Dark Matter and Dark Energy. Based on physical models and assumptions galaxies, stars and gas can be added in a post processing by so called semi-analytic codes. These codes take the output of the N-body simulations and use physical laws to infer the baryonic physics. At the moment simulations start also to explore more and more the gas physics because the relevant codes are good enough and available machines are fast enough to simulate both gas and Dark Matter within one simulation.

Although we are very sure that there is Dark Energy and Dark Matter, we actually do not know what these main components of the Universe are made of. Dark Energy is very mysterious and for Dark Matter we have some particle candidates that are well motivated from particle physics. These are particles that are beyond the Standard Model of particle physics, like supersymmetric particles.

The fact that lots of structure formation simulations only take into account the dark components means, that the simulation particles represent the Dark Matter density field. Dark Matter behaves as a collisionless fluid and one needs to take some care to model this correctly. Therefore every particle in the simulation is not treated like a point source of a gravitational potential. The force is softened to avoid what is called two-body relaxation. This is needed to preserve the collisionless character of the Dark Matter fluid. One has to take into account one very important fact when representing the Dark Matter density distribution by a discrete set of particles. These particles are not real Dark Matter particles. Typical masses for some proposed Dark Matter particles are in the range of 100 GeV. The mass of the particles in the simulation are in the range of thousands of solar masses. It is totally impossible to simulate each Dark Matter particle on its own. So to speak the particle distribution of the Dark Matter fluid is only a Monte-Carlo representation.

After running the simulation its output can be statistically compared to observations. The important point is that both statistics show very good agreement. An agreement of those statistics then proves that our model of structure formation that we have put into the computer simulation is correct.

2 Some details

Gravity is the dominant force at large scales. At the beginning of the Universe there were small density perturbations. These were magnified by gravity during the evolution of the Universe. The main gravitational effect comes from Dark Matter, only at smaller galaxy like scales baryonic physics has to be taken into account. To simulate the Dark Matter one has to solve the equations for gravity in an expanding Universe. Normally the expansion is taken into account by a tricky time integration scheme and the coordinates in the simulation are so called comoving coordinates. These are the physical coordinates rescaled by the current size of the Universe. The main challenge for the force calculation lies in the long range $1/r^2$ character of the gravitational force. The long range character implies that every particle in the simulation feels every other particle. This results in N^2 force interactions. Typical particle numbers for cosmological simulations that are required, are too high to solve this N^2 problem. Without clever techniques to reduce the N^2 for these so called Particle-Particle methods (PP) it is therefore impossible to run such a simulation. The PP method only works for quite low number of particles. With special hardware it can also be used for higher number of particles. So called GRAPE chips are specially designed to calculate the gravitational force with an extreme speed. Using special hardware like this it is possible to use PP methods also with higher number of particles. But this is still by far not enough for cosmological structure formation applications.

A very common method to solve this problem is the Tree method. The idea is that the force of a distant group of particles can be approximated by the force of the center of mass force of that group. This approximation reduces the scaling of the number of calculations from N^2 to a lot better $N\log(N)$. The question is how to arrange the particles in an efficient way. A good way is the so called Tree method. For that the simulation volume is divided into smaller cubes with $1/8$ the volume each at every stage till the smallest cells have only one particle in them. The question for the force calculation is then whether to open a cell, or whether it is fine to take a whole group for the force calculation. Cells that are far away from the point of force evaluation do not have to be opened. Nearby groups need to be opened. To decide on whether to open or not is given by a so called acceptance criterion. This criterion in the end determines the force accuracy you get.

Another very popular method to calculate the gravitational forces are so called Particle-Mesh (PM) methods. In fact they were the first methods used to run larger cosmological simulations. These methods use the fact that the Poisson equation relevant for the gravitational forces is a simple algebraic equation in Fourier space. With a Fast Fourier Transformation (FFT) the forces can be calculated very fast. The FFT requires sampling functions at uniformly spaced points. A grid/mesh is used for this. In the simulation particles are used for representing the density and velocity field. This means that the density field at the mesh points has to be interpolated. The fact that both particles and meshes are used in the simulation gives this technique its name. The Fourier method has some advantages: it automatically implies periodic boundary conditions, softens the forces at small scales because of the mesh resolution and the FFT can easily be parallelised. These points are very important for cosmological simulations. But PM methods have also very critical disadvantages: the softening on mesh scales is very fine because softening is needed to simulate the collisionless Dark Matter fluid, but this also means the the PM code cannot resolve scales below the mesh scale. This is a very serious limitation of the dynamical range of PM simulations. An extension of classical PM methods are so called Adaptive Mesh Refinement (AMR) codes. In these methods the grid is refined in higher density regions. This way the resolution is increased where it is needed.

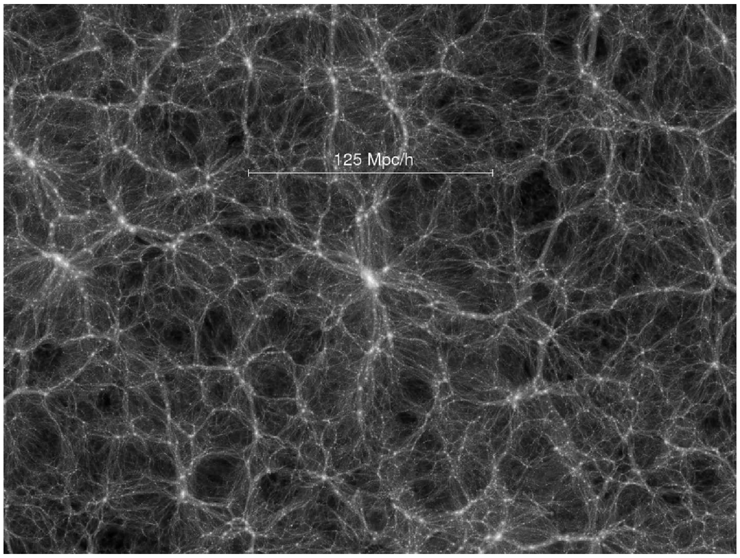

Figure 1: Dark Matter density field. This is a slice through the Millennium Simulation (see references). One can clearly see that the Dark Matter shows a filament like structure. There are also very dense and under dense regions. These under dense regions correspond to very large voids in the Universe.

Another possibility to get rid of the low resolution on mesh scales is to combine the mesh method with a particle based method. This means that the “bad” forces of the mesh on small scales are corrected by a summation of the direct particle forces for close neighbors. These methods are called $\mathrm{PP} + \mathrm{PM} = \mathrm{P}^3\mathrm{M}$ methods (Particle-Particle plus Particle-Mesh). The direct summation of the PP part can also be replaced by a Tree based method. These codes are then called hybrid codes. A very efficient hybrid method is the TreePM method. It uses a force splitting between short and long range force. The short range force is calculated with a Tree whereas the long range part uses the PM method to calculate the forces.

The algorithm for the force calculation is only one problem in simulations. Another important issue is the so called domain decomposition strategy to divide the work between lots processors. Cosmological simulations are often run with a number of processors of the order of 1000. The goal is to reach optimal load and memory balance. There are different schemes around. The cosmological code Gadget uses a fractal space-filling Peano-Hilbert curve as decomposition scheme.

Once all the forces are calculated the simulation can be advanced one time step. The time integration algorithm that is mostly used is a quasi-symplectic leapfrog.

Cosmological simulations have to face lots of other technical issues like for example I/O issues, because the data needs to be stored in parallel, because the typical snapshot size is extremely large.

3 The Millennium Simulation

The Millennium Simulation is a project of the VIRGO consortium, a group of scientists from Germany, UK, Canada, Japan and the USA. The focus of this international team is to run large cosmological simulation and answer important questions by analyzing the output of these runs. The Millennium Simulation was running for about a month on a 512 CPU cluster. After finishing the simulation lots of scientists started to analyze it and they still do until today. The amount of data is very large and the simulation gives us a perfect tool to test our models and see whether they are correct or not. The simulation was done with the Gadget code. Fig. 1 shows one output of the simulation. It is the Dark Matter density field of a slice through the simulation box.

4 Further reading

1. **How to simulate the Universe in a Computer** (Alexander Knebe)
   `http://arxiv.org/abs/astro-ph/0412565`
2. **Cosmological N-Body Simulations** (J.S. Bagla, T. Padmanabhan)
   `http://arxiv.org/abs/astro-ph/0411730`
3. **Cosmological N-Body simulation: Techniques, Scope and Status** (J.S. Bagla)
   `http://arxiv.org/abs/astro-ph/0411043`
4. **Millennium Simulation** (Springel et al)
   `http://www.mpa-garching.mpg.de/galform/press/`