

Intreerede Inaugural Lecture

Prof Felix Spanier

The microphysics of the high-energy universe

16 May 2014



Dit begin alles hier™



NORTH-WEST UNIVERSITY
YUNIBESITI YA BOKONE-BOPHIRIMA
NOORDWES-UNIVERSITEIT
POTCHEFSTROOMKAMPUS

®

The microphysics of the high-energy universe

Inaugural speech

Felix Spanier

In ancient times astronomy has been a matter of the observation of very distant, luminous objects. These have been identified as stars or planets in first place. A star in that sense was a luminous object that was fixed in the night sky and a planet was simply a wandering star. Over the course of many centuries astronomers were able to refine this definition and found different classes of objects. By the 18th century these were planets, stars and nebulae.

Only recently - at least in astronomical terms - it was discovered what the distances in our universe actually are. A first step was the observation of the distance between Sun and Earth, which is 1.5×10^8 km (150.000.000 km). Galileo Galilei found in 1604 that the supernova SN1604 had a distance of roughly 20.000 lightyears. This was a huge leap forward, since it was expanding the sphere of stars extremely. In the 18th century Edwin Hubble was able to determine the distance to the Andromeda galaxy (2.400.000 lightyears). And finally the cosmic microwave background gave a hint of the total size of the universe (13.700.000.000 lightyears).

These distances are incredibly huge compared to our local environment. In fact this poses a challenge to astronomy as such: The detection of objects is extremely difficult and the resolution of these objects is sometimes impossible. Even on the scales of our solar systems, it took centuries to find the planets beyond Saturn. And resolving the surface of Pluto is not possible with even the largest telescope. As I will outline, this poses a major problem in understanding the complex physics of the high-energy universe.

All phenomena of classical astronomy have one thing in common: They are observed through electromagnetic emission, which resembles a blackbody. Generally speaking a blackbody is an object, which emits radiation simply due to the temperature it has. The most common examples of blackbody radiation are the Sun and a typical light bulb. In both cases the spectrum emitted by the follows a simple function, that is only determined by the temperature of the body and its size. The temperature determines its color, while the size determines the total luminosity.

The objects that can be observed then are either emitting radiation them self (like stars) or they are radiated by other bodies (like nebulae). The main question is then: How does the heating to the observed temperature work? For stars this process is mostly

understood: Nuclear fusion is the energy source. Depending on a number of parameters, stars will burn in one of the states described in the so-called main sequence.

A prerequisite for the thermal emission is the collision of particles in the source. Frequent collision of particles yield a certain energy spectrum of particles: The so-called Maxwell-Boltzmann spectrum. This Maxwell-Boltzmann-spectrum is the key concept behind the temperature. For a given temperature the energy spectrum of the particles has the same form and this spectrum then yields the blackbody radiation. One major feature of this spectrum is that it falls off fast for high energies.

In the beginning of the 20th century Victor Hess made an important discovery: He found the so-called Cosmic Rays during a balloon experiment. At first this was only the discovery of some sort of radiation, which was ionizing and whose strength grew with rising balloon altitude. Hess established that this radiation had to come from space and he was awarded the Noble prize for his discovery.

In the course of his discovery a large number of experiments has observed cosmic rays. The main properties are, that they have energies of up to 10^{20} eV and their spectrum is very much unlike the Maxwell-Boltzmann spectrum, one would expect from thermal sources (i.e. sources, where particles are colliding all time to exchange energy). This changed the picture of the Universe dramatically: While before the discovery the universe was made up of plenty of "light bulbs", it has now another dimension of observation.

The observation of nonthermal phenomena as they are called, since they don't follow a Maxwell-Boltzmann distribution, takes place in many different ways: On the one hand there are experiments on Earth and satellites in the solar system, which observe the particles (mostly protons and electrons) directly. Well-known experiments are for example the large AUGER telescope in Argentina, which consists of hundreds of water tanks in the pampa. Experiments like this can measure the energy of the cosmic rays, but since cosmic rays arrive isotropically, there is no way of detecting single sources.

In the 1970s a different window to the nonthermal universe opened up: The COS-B satellite detected gamma-rays - extremely energetic photons - from space. Careful analysis showed, that the gamma-ray sources are associated with nonthermal particle distributions in the source. This allows for an indirect observation of the sources of cosmic rays. The picture of the sky in gamma-rays has been observed by a large number of experiments by now, the Fermi satellite and the HESS telescope being some of the most prominent. The picture which unfolds, is that of a large number of extraterrestrial sources and emission from the band of our own Galaxy.

But not only emission of cosmic rays from very distant sources is observed, but also in the direct vicinity. Satellite missions close to the Earth showed that also the solar system is filled with nonthermal particles.

The nonthermal emission - cosmic rays and gamma-rays - allow a very detailed insight into some of the most energetic events in the universe. Besides this theoretical interesting aspect, it is also worthwhile to study the interaction of cosmic rays with Earth's environment, this is the so called space weather.

In the course of the past decades a number of different sources have been identified as sources of nonthermal emission:

- Active Galactic Nuclei: Some galaxies are extremely bright and their central core emits far more energy than most other galaxies in total. It has been established, that this is due to the emission of plasma jets from the center, which is powered by the gravitation of supermassive black holes
- Pulsars: Neutron stars in our own galaxy have powerful magnetic fields and are rotating very fast. Through this process they are able to emit energetic particles and pulsating photon emission.
- The Sun: Our Sun - like most other stars - is able to accelerate particles to high energies. While the energies reached there are not comparable to the energy reached in more extreme sources like AGN and pulsars, Sun's vicinity makes it still a dominating source.

Let's come to the key problem in the understanding of nonthermal phenomena: The size. As outlined in the beginning, astronomical source are always very distant. The size of the emission regions of nonthermal particles are rather small. For the Sun the following picture is presenting itself: The typical phenomenon on the Sun, that is associated with nonthermal emission is a so called coronal mass ejection (CME). This CME starts in the magnetic loops, that build on the Sun's surface through turbulent processes. The magnetic loop has a typical size of 10^4 km. A CME starts, when this coronal loop expands and at some time breaks. When this happens, the CME itself has a size of roughly 10^6 km. This already very small compared to the distance to Earth with 1.5×10^8 km. The difference in 4 orders of magnitude makes a direct observation very difficult. For the case of the Sun one would expect to have in-situ observations (an instrument at the site of the emission), but the extreme environment of the Sun forbids any technical instrument in its vicinity.

For Active Galactic Nuclei, the numbers are even worse: The smallest identifiable unit in which energetic particles are accelerated is a so called blob. These blobs are commonly found to be 10^{11} km large. They are embedded in the plasma jet emanating from the central black hole. These jets have a length of the order of 10^{15} km. The distance to the nearest AGN is then more than 10^{21} km. These numbers clearly show, that astronomy is not able to resolve the smallest scales in AGN.

In fact, this is not the main problem: The main processes leading to the acceleration of particles are plasma processes. A plasma is - generally speaking - a gas with ionized atoms. To understand the phenomena inside an plasma it is inevitable to resolve the physics of this plasma to its smallest scale. This is the so called Debye length. Typical Debye lengths in terrestrial laboratory plasmas are of the order of millimeters. In the solar environment Debye lengths are of the order of 1 – 100 cm. The situation in AGN is rather unclear: Here Debye lengths of the order of 1 – 1000 km are expected. In any of these cases it is obvious, that the microphysical scales of physics are many orders of magnitude below the observable macroscopic objects. It is in fact impossible to observe these scales with astrophysical methods. The attempt to observe a 1 cm large plasma phenomenon over 1.5×10^8 km spans 13 orders of magnitude! This would be the same as observing a single atom over a 1 km distance.

We are now left with the problem, that we cannot observe the actual processes, we are interested in, but we may also not recreate these phenomena (due to their sheer size) in any laboratory. The astrophysical research is, therefore, required to model the processes inside these sources. The formulae governing these processes are mostly known: The Vlasov equation and Maxwell's equation determining the basics of plasma physics and electromagnetic radiation, but also quantum mechanics, quantum chromodynamics and special relativity to describe high-energy interaction between various constituents in a plasma. While the set of equation can be written down more or less easily, a solution using only pen and paper is only possible after severe simplification of the problem. The last resort for a solution is the use of computers to model problems.

The use of computers to model physical problems has been established already very early. The first computers were used to tackle simple calculations, which were just tedious, but today's generation of supercomputers allows for the simulation of rather complicated systems. While the amount of simplifications necessary to actually model any physical system is less compared to analytical calculations, other problems arise: Numerical calculations are inherently not exact.

There is a wide choice of methods to simulate physical systems. Their common denominator is the discretization of any system, but there are various possibilities on how to do this. In the following two different scenarios involving two different methods are presented, which are part of the research I pursue at North-West University.

The first example is the shock acceleration in Coronal mass ejections. This has been briefly mentioned above, when length scales in the solar environment are discussed. Since coronal mass ejections are transient events, they have a temporal evolution, which is taking place over the course of several days. This starts off with magnetic field configurations on the Sun's surface. The large magnetic loops rearrange causing plasma to be ejected out of the Sun into the interplanetary medium.

This expansion of plasma outwards is so fast, that the velocity exceeds the speed of sound in that medium. This leads to the formation of a shock front. The same effect is observed, e.g., in supersonic planes. In the climax of an CME event we find several layers of this phenomenon: The Sun at the bottom with magnetic field lines pointing outward, a plasma pile up moving outward and a shock front at the outermost boundary. This shock front can reach a length of several millions of kilometers. In very extreme events these shocks may reach the orbits of the inner planets. On the other hand the thickness of this shock is usually only as wide as the gyration radius of proton in the magnetic field. This is mostly well below a kilometer.

As outlined before, it is necessary to model the system from the smallest scale, the Debye length, upwards. The Debye length in this system is in between 1 and 100 centimeters. A full simulation of the system would, therefore, require the resolution of 11 orders of magnitude. It is obvious that this may not be accomplished by any computer simulation. While the actual limitations of computers simulations depend very much on the type of simulation performed, a rule of thumb says, that for three-dimensional problems only slightly more than three orders of magnitude can be covered.

This in turn means, that only a very tiny fraction of the problem can be simulated. For the problem set here, this would be the shock front. How can the physical behaviour

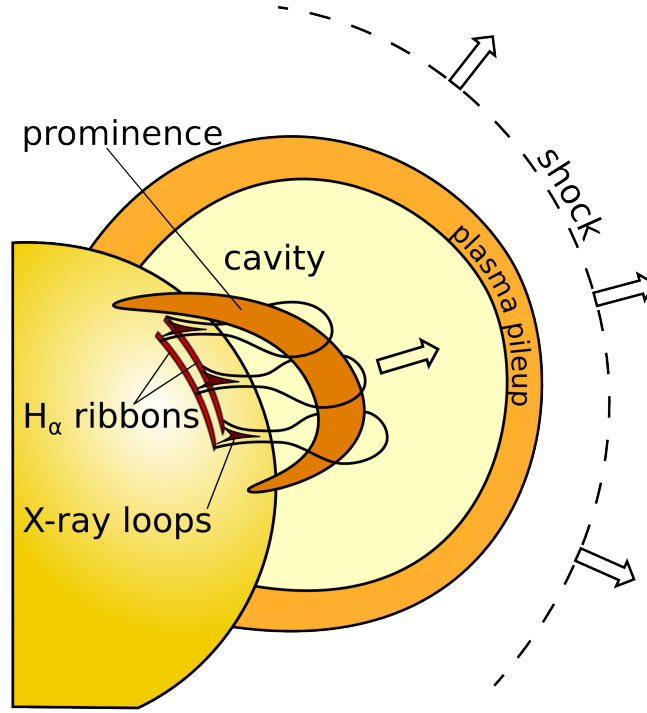


Figure 1: Sketch of a CME launch (image courtesy: Forbes, 2000)

of such a system be modeled? The most common simulation technique is the fluid simulation. This assumes, that a plasma is acting like a charged fluid. This is true for thermal phenomena, but no longer for nonthermal problems. So we have to resort to a different method.

One standard technique used to study this kind of problem is the Particle-in-Cell method. It combines the advantage of a particle-only approach (the energy distribution of particles can have any shape) with the advantage of fluid method (particles interact with each other via electromagnetic fields). The outline of such a simulation code is the following

- The simulation space is cut into cubes. In the computer's memory field quantities as the electric and magnetic field are stored for each cube
- The particles (electrons, ions, etc.) are sampled through a number of macroparticles (one macroparticle represents the properties of a large number, possibly 10^9 and more, of particles like their velocity, mass and charge)
- The motion of the particles generates charge and current, this charge and current can be deposited on the grid
- From charge and current follow the electric and magnetic field through Maxwell's equations

- Electric and magnetic fields act upon charged particles via the Lorentz force, particles are accelerated and translated accordingly
- This cycle of field update and particle motion is repeated

Typical simulations on supercomputers have a size of $1000 \times 1000 \times 1000$ cells with approx. 100 particles per cell. On a single-CPU computer one timestep would take around 100.000 seconds. This is the reason, why large supercomputers with 2000 and more CPUs are needed.

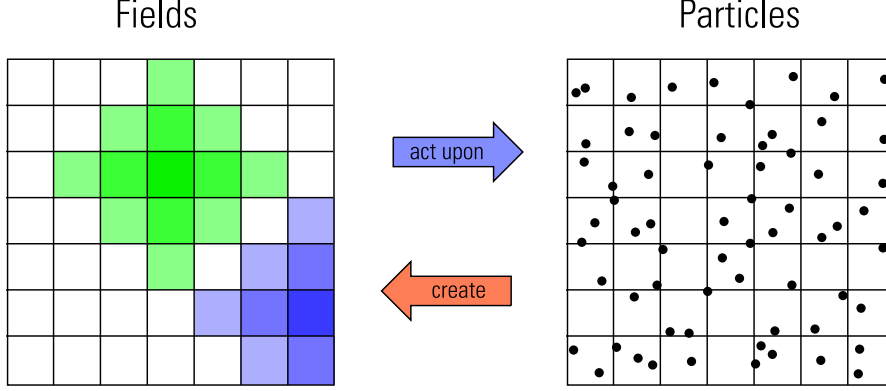


Figure 2: Schematic of a particle-in-cell code

This is especially true, since the physical size of grid cell is limited by the Debye length and the duration of a time step is limited by the light crossing time. That yields typical simulation lengths of several thousand timesteps.

With the basic simulation technique in place, we can now return to actually simulating the formation of a shock in the interplanetary medium. In the real world this just happens, when a supersonic plasma flow is moving outwards. In computer simulations this by far more complicated. One common technique to reproduce the real physics is by using a box with a reflecting wall. Then one can simulate a plasma streaming from the other side. It will be reflected and then eventually produce a shock front moving through the plasma.

This setting seems like a very simple approach, but this is mainly an artifact of our experience on Earth: In the thermal and collision dominated on Earth, a gas in this situation would just build up a shock front. In the solar setting, which is not dominated by collisions, the plasma streams may partially flow through each other. One observes, that the behaviour along such shock fronts exhibits very complicated patterns. In first place filament of currents show up, then plasma waves are excited. This leads on the one hand to a strongly modified shock wave. On the other hand the complex shock pattern allows for the acceleration of charged particles in the electric and magnetic fields of the shock wave.

The simulations performed so far in this field are only scratching the surface: We are limited in the total extent of the physical realm, the temperature of the plasma or the

magnetic field. And even worse: We are not even able to simulate the real mass ratio of protons and electrons. But the results that can be achieved now are promising enough to keep up the work on this kind of project. Processes that may never will be accessible to observations can be far better understood using this simulation type.

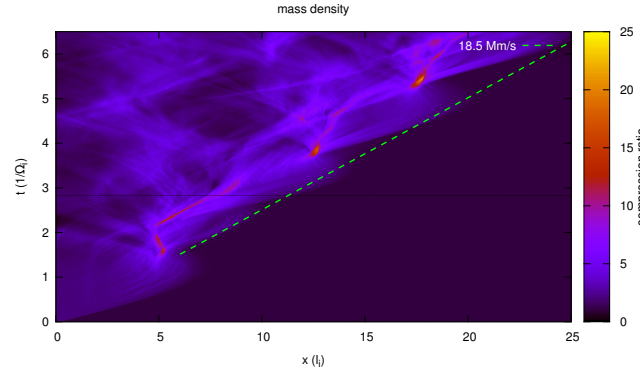


Figure 3: The mass density in a shock simulated with the PiC-technique (image courtesy: Kilian, 2013)

Another example is the gamma-ray emission from AGN. The previous discussion of length scales for this setting shows, that this is an even more extreme example than the simulation of coronal mass ejections. There approximately ten to twelve orders of magnitude in between the smallest plasma physical scale and the length of a jet. Again on has to ask the question, which length scales can be considered. Rather than going to the smallest scales, like in the previous example, the setting demands that the so-called blobs as smallest unit of gamma-ray emission have to be resolved (this assumes that all the physics beyond the physics is more or less isotropic). This blob as smallest unit is then used to model a longer jet.

In the past decades a simple model for the emission from AGN has been established: An isotropic sphere filled with homogenous plasma in equilibrium is assumed and the emission through various processes is calculated. This kind of model leaves a large number of questions open: How can the time-dependency of AGN be described self-consistently? How can the radio emission of extended areas of the jet be understood? Which processes do contribute to the emission of the AGN?

The model my group has developed, is assuming a set of blobs with an underlying fluid velocity, that can resemble shocks, as they are seen in AGN. The acceleration of particles is reproduced by the so-called Fermi-I process, a scattering process over shock fronts. This produces power-law like particle spectra for electrons and protons. These particles can radiate through a number of processes starting with simple synchrotron radiation in magnetic fields over Compton scattering to photo-meson production.

While a first sketch of the underlying physics looks very simple, the set of equations involved is rather complex and it involves more than just plasma physics. Since at the high energies observed in AGN quantum mechanics and particle physics also play a role,

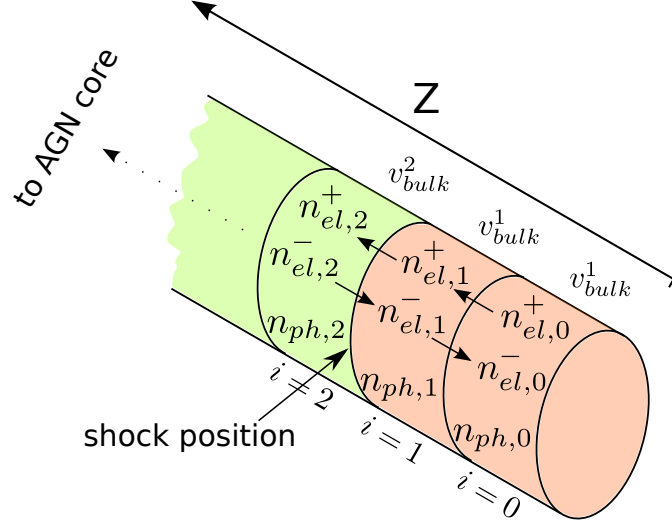


Figure 4: Schematic of the resolved AGN simulation code (image courtesy: Richter, 2013)

their effect have to be incorporated.

Using such a model yields spectra and lightcurves for AGN under different physical conditions. With the growing number of observations of AGN, an endeavour also supported by the South African government and North-West University, this will allow for narrowing down the parameter space in which AGN exist. Unfortunately we are facing a problem here: The more detailed models are more and more able to understand the physics of AGN, but on the other hand it is necessary to test many different parameter sets to accommodate the observations. And this again requires enormous computational resources.

There are still many open questions, that have to be answered by astrophysics:

- What are the accelerators of cosmic rays? We cannot observe them directly by now. We have to understand the physics of these sources by modeling their interior.
- By which process are particles actually accelerated to high energies?
- What is the mechanism of gamma-ray emission in AGN?

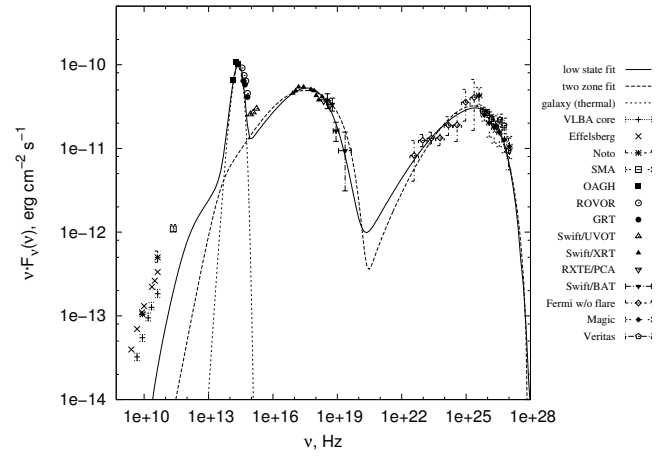


Figure 5: AGN spectrum of Markarian 501 simulated with the above mentioned code (image courtesy: Rcither, 2013)