The Puzzle of Cosmic Ray Generation

Michael McEllin

We can directly detect cosmic rays as high energy particles as they descend from space. We are pretty sure that all but the lowest energy particles come from outside our Solar System because they arrive uniformly from all directions.

In fact, there is no shortage of evidence that high energy particles are abundant in the universe. For example, there are many astrophysical radio sources (such as supernova remnants or the much more powerful extragalactic radio sources) which show all the characteristics of *synchrotron radiation*. (The strength of the emission varies with frequency in a way we call *non-thermal*, because it looks different from the radiation you would expect from a merely “hot” body. In addition, the strong polarisation suggests that the movements of the radiating particles are organised by a magnetic field. Synchrotron radiation is produced when high energy electrons spiral around magnetic field lines.) The mechanisms that produce high energy electron produce high energy protons as well, but being much heavier they spiral more slowly and radiate much less (but we know that they are there because of collisions with nuclei in the interstellar medium that produce detectable gamma rays). We can also detect a background radio emission from our own galaxy coming from *between* the stars. This is also synchrotron radiation and it is direct evidence of high energy particles spiralling along the magnetic field lines between the stars.

So, we *know* that they are present in very large concentrations in supernova remnants and other more exotic objects because we see them losing energy (as radio waves and sometimes as light and even X-rays) and it almost certain that they leak from these zones. We *know* that they move around the galaxy because we can detect their synchrotron radiation as they spiral round the interstellar magnetic field lines. We also *know* cosmic rays arrive at the Earth because we detect them, and the rate of arrival is consistent with what we know about their density in interstellar space. Everything suggests that the cosmic rays we detect on Earth come from at least some of the astrophysical objects which show such clear evidence of high energy particles. We cannot say that supernova remnants are responsible for all cosmic rays detected on Earth because we cannot eliminate the possibility of other types of source (and we can identify a number of real possibilities). Furthermore, the very highest energy cosmic rays may well be beyond the acceleration ability of even supernova explosions and we have to look elsewhere.

Nevertheless, we know that cosmic ray particles are accelerated somehow to very high energies because we can measure the energy with which they arrive. What we are only now coming to understand (and it is not a complete understanding) is how charged particles like protons, electrons and even some heavier nuclei are pushed up to such very high energies.

One of the most fundamental principles of physics, the second law of thermodynamics, says that all forms of energy will ultimately end up as heat energy, because it is the most disordered form of energy, basically just random motions of atoms. Fundamentally, there are very many more ways that a collection of atoms can move around in a disordered way than they can move in ordered ways. (The air molecules in a room do not suddenly all decide to rush to one end of the room. Nothing in the laws of physics prevents that happening by chance, but that chance is so remote that such an event would still be incredibly improbably in the entire history of the universe.)

Heat is, for example, the kinetic energy of the molecules in a gas (or the vibrations of molecules in liquids and solids). We can calculate the distribution of velocities of the gas molecules from first principles (the Maxwell-Bolzmann distribution: first year university physics) and if we do so we find that most molecules have close to an average velocity, with fewer at low or high velocities (and very few at very low or very high velocities). In fact, the number with high velocities (and so high energies) decreases according to a negative exponential function (which means that the number decreases very fast indeed as molecular energy increase). The fundamental reason why the molecules in a gas move towards this *thermal equilibrium,* which is the same as the state of maximum disorder, is that they collide with each other, exchange energy, and quickly "forget" any original non-random motions. Remember this point, it is important: collisions between individual particles have an important role in producing disorder. We will come back to this, because in places where cosmic rays are generated particle-particle collisions may be quite rare and it then takes a long time to establish thermal equilibrium.

If we look at the heat or the light radiated by a hot[1] body the radiation also has a very characteristic variation with frequency, which again is due to the energy distributing itself amongst light waves in a way that is controlled by randomness. This most disordered form of electromagnetic radiation is called *black body radiation*. Real hot bodies may be "coloured" in the sense that they are better at radiating at some frequencies than others, but this does not stop *thermal radiation* from being easy to recognise. In particular, just as with molecular velocities, at the higher frequencies the amount of radiation decreases very quickly, again with a negative exponential form. Stars are thermal emitters, as are planets.

Non-thermal emission is anything that is clearly not related to black body radiation and it usually means something physically interesting – even exotic - is going on. When we see *non-thermal emission* it normally means that there are more particles at high energy than there should be if everything has settled down to the greatest state of disorder that is the normal ultimate end point of all physical processes. In practice a lot of non-thermal emission turns out to have a power-law spectrum, described by formulas such as $f^{-\alpha}$ (where f is the frequency of the radiation and α is known as the spectral index). The formula says that the radiation intensity decreases as frequency increases, but much less rapidly than black body radiation. This type of radio spectrum turns up again and again when we look at the more exotic astrophysical objects such as supernova remnants and X-ray sources. It also turns up when we look at the energy distribution of cosmic rays. That is not a coincidence, because the radio spectrum is a direct product of the cosmic ray energy spectrum.

Non-thermal emission means that there *has to be* some process going on with an *organised* flow of energy. In the case of a supernova, the organised flow of energy is the result of an explosion that throws the outer envelope of the star out in all directions. Until it is slowed down into turbulent disorganised motion by ploughing into the very thin interstellar medium there is a bulk flow of kinetic energy outwards which can perhaps be tapped by other physical processes to do things like make cosmic rays. In the case of supernovas we know that there is a very, very large amount of energy suddenly available (of the order of 10^{46}J – for a short time a supernova can outshine an entire galaxy of a trillion stars). Surely we can find a way to use some of it?

---

[1] "Hot" is a relative term, especially in astronomy. The general temperature of the Universe is 2.7^{o}K – that is 2.7 degrees Kelvin above absolute zero. We see the universe at this temperature by measuring the ubiquitous microwave background, which has a black body spectrum. Anything warmer than this will emit more radiation into space than it absorbs from its surroundings. Nothing is harder to detect in astronomy than a cold black body orbiting between the planets (such as small asteroids or exhausted comets). They can be really cold – maybe only 50 – 100^{o}K – but they still emit infrared radiation and are still warmer than the microwave background radiation of the universe. If we fly satellite telescope with cameras cooled in liquid helium (-2^{o}K or there about) "cold" asteroids will look "hot" against the 2.7^{o}K background. This is one of the best (sometimes the only way) to see such very difficult to detect objects.

That is very easy to say, but astrophysicists do not agree entirely on the detail of how it happens. How do we take that organised flow or energy and feed it into really high energy particles? How much of the energy can we use? That is to say, how efficiently can we tap the flow?

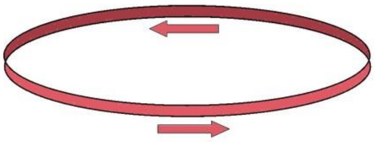
On Earth, physicists who build particle accelerator get to design things very carefully. Acceleration of high energy particles is not at all easy – some physicists say it is remarkable that particle accelerators work at all because they are so difficult to design, build and operate. However, with all these efforts they find that less than 10% of the electrical energy going into the accelerator buildings ends up in the high energy particle beam. (It tends to be less efficient the higher the energy you wish to achieve.) We might be quite surprised if nature, having to work without such careful design and organisation manages to achieve 10% efficiency, and even 1% would be considered respectable and interesting. Nature does perhaps have a few things going for it that we do not on Earth, such as lots and lots of space and lots and lots of time, and *extremely* extreme events like supernova that concentrate energy in ways entirely impossible on Earth. Still, large concentrations of energy tend to blow things apart rather than organise them in rather nice and precise ways. The puzzle of cosmic ray acceleration is not that you get *some* high energy particles in these extreme situations – we would be surprised if we did not get *some* – it is that we seem to get a lot.

How do we know that there are “a lot” of high energy particles? (That is to say, we are tapping quite a lot of the bulk energy flows to use for acceleration.) Well, this involves a fair amount of estimation, and you always have to take account of the uncertainties in the measurements.

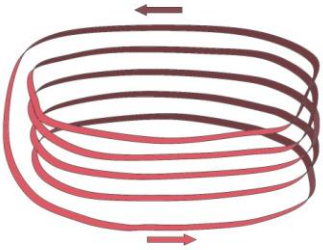
We estimate the amounts of energy involved in astrophysical events in several ways: we look at the amount of power being emitted; we look at how big things are; how dense; how fast things are moving and so on. Most of the time these estimates depend on knowing the distance to the object being studied and in astronomy that can be one of the hardest things to measure. There is no way to put a tape measure into space so we have to erect a ladder of indirect methods of estimating distance, all based on more direct methods of measuring the size of the Solar System (e.g. using radar, which is a pretty good tape measure). Sometimes these distance estimates can be quite accurate. (For an astronomer 5% is normally “accurate”!) Other times we have to accept that our distance estimates may be wrong by a factor of 2 (and numbers like energy estimates could be wrong by a factor of 10!). In other cases distance estimation is little more than guesswork. Astrophysicists get used to handing this type of uncertainty and normally explain whether they are making conservative assumptions (“It needs at least this much energy – and it is still wow!”) or being speculative (“It *could* need as much energy as this! *WOW*!”)

Having said all that, and made all the excuses, it does look like places such as supernova remnants must manage to feed a surprisingly high proportion of the energy of the explosion into high energy particles and this takes some explaining.

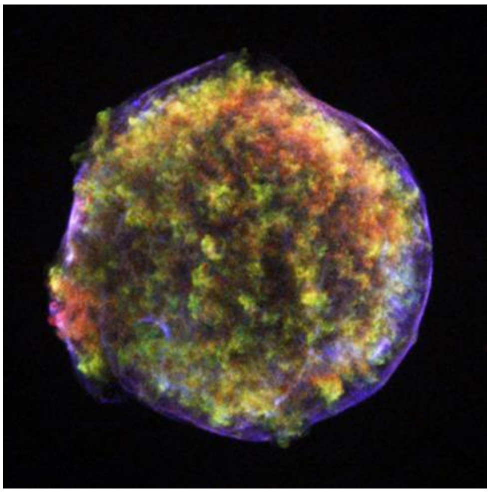
Firstly, we have to remember that all the interesting physical processes are going on in a *plasma* – the fourth state of matter – in which electrons and proton (and some helium nuclei) all move independently. Plasmas carry electrical currents and astrophysical plasmas are actually very good conductors, because the charged particles are so far apart that they hardly ever collide. That means that they do not lose their energy so easily. (In a metal, electrons suffer collisions at high frequency and give up their energy of motion – the current – at a relatively high rate.) In a plasma, the current also means a magnetic field, and a magnetic field means that the charged particles spiral around the field lines, and the particles cannot leave their field line without a collision. Hence, magnetic field is tied to the charges and the charges to the field. (The field is “frozen” in.)



Where does the magnetic field (and the currents) come from in the first place? You have to start with a very small initial "seed" field (and there are several ways of getting this, which, however, would take too much time to explain here – please accept it). The seed field is frozen into the plasma and must move with it, so we can twist parts of the plasma around and around and the same field lines are wound over and over themselves. More field lines mean a stronger field! It is, of course, in truth rather more complicated than this, but the essential points are correct: if you stir up a conducting fluid threaded by a small magnetic field some of the energy in the turbulent motions of the fluid gets transferred to energy of twisted magnetic field lines. (This type of process is also responsible for the Earth's magnetic field.)

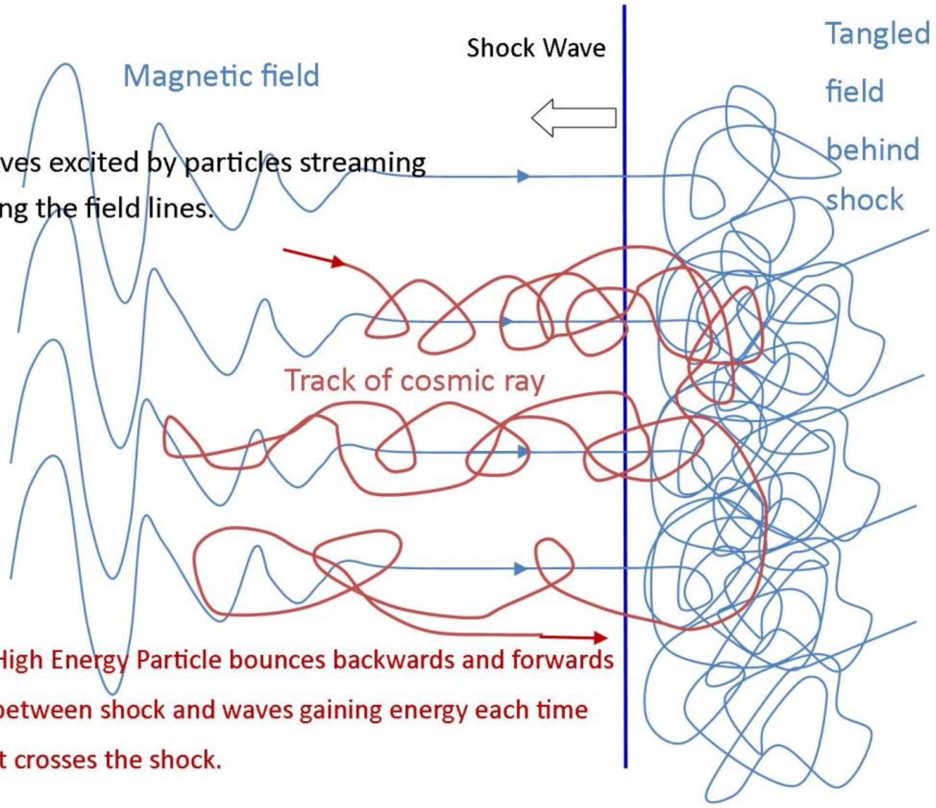


In a supernova explosion the outer shell of the star is ejected at very high speed and ploughs into the interstellar medium. The material from the star contains its own magnetic field and the interstellar medium has its own field. The two magnetic fields cannot penetrate each other, because the charged particles supporting the currents do not collide with each other. The supernova ejecta therefore pushes the interstellar medium before it like a snow plough.



We therefore get shock waves moving out into the interstellar medium and everything behind the shock gets very turbulent (i.e. a prime site for winding up magnetic fields!). In fact, we have bundles of plasma held together by magnetic fields bumping in to each other at very high speeds. In the image on the left (Tycho's Supernova), which is an overlay of from images from several X-ray bands, the red and green is the hot and turbulent material ejected from the star. The blue is the radiation from extremely high energy electrons accelerated in the shock wave moving ahead of the expanding shell of ejecta.

The next bit is complicated and really only makes complete sense if you write down mathematical equations. In essence what goes on is rather like a ball bouncing between a couple of moving tennis rackets. A particle already moving somewhat higher than normal thermal energies might meet an approaching shock wave and it bounces back moving very much faster (because it cannot penetrate the shock's magnetic field, and it adds the shock's velocity to its own reversed velocity. One bounce is not enough, but it turns out that the accelerated particle cannot get very far up stream of the shock wave before it gets turned back and bounces again off the advancing shock. Magnetic field lines in a plasma are a bit like elastic bands or stretched strings: twang them and waves can move up and down (but being a plasma, there are many different types of wave). Now let us suppose that high energy particles are streaming along a field line ahead of a shock wave after their first bounce, and they encounter a small kink in the field (not at all unlikely). They get slightly deflected and exert a small push on the kink which starts a wave running along the field line. That wave, however, cannot move nearly as fast as the accelerated particles, so the high energy particles continue to push on the moving kink, feeding it energy, and it gets bigger and bigger, until it grows to a point where the particles run up the front of the wave and get turned back towards the still advancing shock wave.



This is like two tennis players running towards the net and exchanging a ball that gets faster and faster until one of them misses a shot and the particle escapes from the acceleration zone.

That explanation is a gross oversimplification (for a start we have taken no account of relativity) but it does capture a couple of important points. The whole thing is possible because the high energy particles gyrating round the magnetic fields rarely collide with each other (and hence randomising the energy) so they have the possibility of *accumulating* energy from many interactions with shock waves. The other point is that we have a very light particle (say a proton or an electron) interacting with a bundle of plasma held together by magnetic fields, which looks to the proton like a very heavy object (because it is a collective motion of a very large numbers of particles bound together). When you share energy in collisions of this sort, the lighter particles end up moving *very* much faster than the heavier particles. Given enough collisions and enough time it appears that this energy sharing can feed at least a few percent of the energy in the turbulent motions into magnetic fields and from there into high energy particles. The most optimistic estimates say that as much as half the kinetic energy of the shock eventually goes into magnetic fields. (Others say <1%. The evidence is not yet conclusive.) We also have reasons for suspecting that the energy ultimately gets shared fairly equally between magnetic fields and high energy particles. Since a fair fraction of the energy of the supernova explosion ends up in these turbulent motions that means quite a bit of the original explosion energy ends up in cosmic rays.

That is just a flavour of what might be going on in the regions where the shock waves from the supernova propagate through the interstellar medium. If you want to know the full tale you will have to study astrophysics at university.

This is not the only acceleration mechanism that we know about. For example, a rapidly rotating neutron star with a strong magnetic field is in essence a very powerful dynamo capable of creating very strong electric fields. (The Crab neutron star is heavier than our Sun, but only 10km across and it rotates at about 30 times per second. Its magnetic fields may be billions of times more powerful than the most powerful fields ever generated on Earth. They create electric potential differences thousands of times bigger than the effective potential – 14×10^{12} volts - of the Large Hadron Collider.)

It is certainly the case that the pulsar in the Crab Nebula is accelerating high energy electrons into the remnant because we can see their effects close to the location of the pulsar. (The image on the left, which was made in X-rays, shows streams of ultra-hot particles coming from the Crab pulsar.) I am not sure that I can explain how this works. Correction! I *know* I cannot explain how it works. The electrodynamics of relativistic rotation of ultra-strong fields is fearsome and complex computer calculations are only now beginning to produce results that look like the image above. However, all the publications that I have looked at emphasise how much is still not known about how all this works.