JOURNAL OF THE KOREAN ASTRONOMICAL SOCIETY
29: S267 ~ S270, 1996

COSMIC RAY ASTROPHYSICS

L O'C DRURY
Dublin Institute for Advanced Studies, School of Cosmic Physics, 5 Merrion Square, Dublin 2, Ireland

ABSTRACT

The problem of the origin of cosmic rays is considered in an astronomical context and the current observational situation summarised. The evidence for acceleration in supernova remnants is critically examined.

Key Words : Cosmic Rays, Supernova Remnants, Shock Waves

I. INTRODUCTION

Most astronomers regard cosmic rays as nothing more than a nuisance, an annoying and unwanted effect which disturbs their delicate photon detectors in unpredictable and irritating ways. However the origin of the cosmic rays is also one of the oldest major unsolved problems in modern astrophysics. Victor Hess discovered the existence of a very penetrating ionizing radiation coming from above the atmosphere, which we now call the cosmic rays, back in 1912 and the problem of understanding the nature and origin of this radiation has been a well-defined question since then. The only other open question which is equally specific and of similar antiquity is that of identifying the carrier(s) of the diffuse interstellar bands. Furthermore it should be noted that we are not talking about some relatively insignificant effect; the energy density of the cosmic rays in the Galaxy is comparable to that of the magnetic field, their pressure is comparable to that of the thermal gas, and they are thought to be important drivers of chemical reactions in molecular clouds.

Obviously a short article such as this cannot be an objective and detailed review of the field. Rather it should be read as a personal view of some aspects which I find fascinating. Similarly I only refer to some of the more recent interesting papers without any claim to completeness; my apologies to the many others who should have been cited. For those who want more information the best way to find out what is new in cosmic ray research is to consult the volume of rapporteur talks presented at the most recent in the biennial International Cosmic Ray Conference series (Rome, 1995; Durban, 1997; Salt Lake City, 1999; details available on the web at http://www.dias.ie/c4/icrc.html).

II. COSMIC RAY ASTRONOMY AS AN INVERSE PROBLEM

It is helpful to think of the cosmic ray origin problem as an inverse problem, where we start with what we actually observe at and near the Earth, and try to work backwards to the mysterious source or sources of the cosmic rays. There is a conventional terminology for the different processes thought to be involved along this chain which is useful even though it may be seriously misleading. At the far end, in the mysterious sources, there is an "injection" process which takes ordinary matter and feeds it into an "acceleration" process which produces a spectrum of accelerated particles over a wide range of energies. After leaving the sources the particles are subject to "propagation effects" associated with transport in the Galactic disc and, probably, an extended halo region surrounding the disc. Those which reach the neighbourhood of the solar system are then subject to "modulation" by the magnetic structures in the solar wind before finally being detected, either directly by instruments on spacecraft and balloons in the upper stratosphere, or from the ground by the secondary effects they cause in the atmosphere.

A serious difficulty in experimental cosmic ray physics is that the signal we measure can only be related to the processes we really want to study, namely those of injection and acceleration, by applying corrections for the effects of Galactic propagation and solar modulation. In addition we have virtually no directional information. In photonic astronomy one can at least be sure that photons travel along straight lines (strictly space-time geodesics; the distinction is only relevant in the context of gravitational lensing), and if one measures the arrival direction of a photon that is also the direction in which the source of that photon lies. Charged particles, in contrast, are deflected by magnetic fields so that, except at very high energies, their passage through the Galaxy resembles a random walk rather than rectilinear propagation. A charged particle detected coming from one direction may have originated in a source located in some completely different part of the sky.

III. WHAT IS OBSERVED

The cosmic rays detected at and near the Earth consist mainly of fully stripped atomic nuclei, some electrons and a few positrons and anti-protons. Concentrating on the dominant nuclear component, all species are observed to have energy spectra which are very close to power-laws of the form $N(E) \propto E^{-2.7}$ for energy per nucleon in the range 1 GeV to about 10^{14} eV. Below 1 GeV measurements show large variations which clearly correlate with solar activity. This solar modulation effectively wipes out any information about the interstellar spectrum at lower energies and in addition the measurements are contaminated by particles locally accelerated in the heliosphere.

At around 1 GeV per nucleon it is relatively straightforward to determine the detailed chemical composition of the cosmic rays; it turns out to be quite similar to the "local galactic abundances" of the elements, but with one major difference. The deep valleys which occur in the standard galactic abundances, most notably those associated with the light elements Li, Be and B but also the region just below Fe, are partially filled in in the cosmic ray abundances. This can be easily explained as a propagation effect. In travelling through the interstellar medium (ISM) the cosmic ray nuclei will occasionally collide with the nucleus of an ISM atom. Many of these collisions result in what are called spallation reactions in which a fragment is chipped or spalled off the incident nucleus. In this way a relatively abundant C, N or O nucleus can be converted to a rare Li, Be or B nucleus and the sub-Iron region can be partially filled in by spallation products from the Iron peak.

In addition to providing a simple explanation for the gross features of the cosmic ray composition this idea is important because it enables us to experimentally determine some of the effects of Galactic propagation. The primary cosmic ray flux in the solar vicinity can be measured, the interaction cross-sections are in principle known from nuclear physics, and the mean density of the ISM and its composition is known; from these the production spectrum of spallation products can be calculated. By comparing the observed spectrum of secondary nuclei with this deduced source spectrum we can determine the effects of Galactic propagation on the secondary nuclei, and thus by inference also on the primary nuclei.

In the region around a few GeV per nucleon where we have sufficiently accurate composition data for this comparison, the observed energy spectrum is substantially steeper than the production spectrum. The natural explanation for this is in terms of energy-dependent escape from the Galaxy. At low energies the particles are more tightly coupled to the magnetic field and take longer to diffuse out of the Galaxy, whereas at higher energies the confinement time in the Galactic field is less. The compositional data are consistent with a confinement time of order 10^8 yr at around a GeV per nucleon falling as $E^{-0.6}$. This in turn implies that to produce an observed primary spectrum of $E^{-2.7}$ we need a production process which produces a harder spectrum more like $E^{-2.1}$.

Strictly this argument can only be made over a rather small energy range, say 1–30 GeV, but the fact that the observed primary spectrum is a very close approximation to a single power-law out to 10^5 GeV suggests that the same relationship between production spectrum, observed spectrum and confinement time holds over this extended energy range. However it must be admitted that not only is this a very dangerous extrapolation, but also there are then some difficulties in understanding the very small anisotropy (of order 10^{-4}) of the cosmic rays at 10^{13} eV.

There is a feature, usually called the "knee" in the energy spectrum at about 10^{15} eV beyond which the observed spectrum steepens to roughly E^{-3} and there may be a flattening (predictably called the "ankle") at around $10^{18.5}$ eV. Unfortunately it has not been possible to determine the detailed composition at these energies and thus almost nothing is known about the propagation characteristics. In particular it is not known whether the "knee" reflects a change in the production spectrum or a change in propagation. The very high energy cosmic rays, from say 10^{16} eV to at least 10^{20} eV are a fascinating topic; one interesting model for their origin is presented at this meeting by H Kang.

Returning to more modest, but still relativistic, energies there is one further important piece of evidence which has to be discussed. This comes from gamma-ray astronomy. In addition to spallation reactions nuclear collisions at energies above a GeV typically produce large numbers of pions. The neutral π^0 particles then decay rapidly into two gamma-rays which, as uncharged photons, do travel in straight lines. Thus observations of diffuse gamma-ray emission above 100 MeV can be used to trace the product of the cosmic ray intensity and the ISM gas density. Compton Gamma-Ray Observatory EGRET, and also earlier COS-B and SAS-2 observations, indicate that the cosmic rays are distributed rather smoothly throughout the Galactic disc, but with a radial gradient from higher values near the centre to lower values in the outer Galaxy. This is a clear indication that the cosmic rays, at least in this energy range, are of Galactic origin. Further confirmation comes from observations of the Magellanic clouds which show that the cosmic ray intensity in the small cloud cannot be more than one quarter of that observed in the solar neighbourhood.

The local energy density in the cosmic rays is of order 0.4 eV cm^{-3} and the associated pressure 0.17 eV cm^{-3} comparable to other energy densities and pressures in the ISM. However the number density of the cosmic rays is very low, of order 5×10^{-10} cm^{-3}. Assuming a Galactic origin, and using the information obtained from the compositional data to constrain propagation, the power needed to maintain the cosmic rays at this level is about 10^{41} erg s^{-1}.

IV. WHERE ARE THE SOURCES?

There is one main piece of circumstantial evidence suggesting a rôle for supernovae (SNe) in the acceleration of cosmic rays. The mechanical energy input into the Galaxy from SNe is estimated as 10^{42} erg s^{-1} and it is hard to think of anything else which could plausibly drive the cosmic ray accelerator. In addition we have a theoretical picture of how the acceleration (and injection) might actually work; this is through a version of Fermi acceleration operating at the strong shock waves associated with the supernova remnants (SNRs) produced by the SN explosions. It is worth noting that the so-called "adiabatic loss problem" rules out any model in which the cosmic rays are generated directly in the SN explosion itself; thus if there is a link to SNe it must be indirect and through the resulting remnants. As the SNRs can be thought of as the dynamical structures which process the mechanical energy of the explosion into other forms they are also the most natural places to locate an acceleration process. Quite detailed models of cosmic ray acceleration in supernova remnants have been calculated by several groups; probably the best are those described in Berezhko et al (1996) which also contains references to earlier work. Remarkably it has turned out that it is actually easier to understand time-dependent expanding spherical shocks than steady planar ones! The latter are plagued with various mathematical pathologies which turn out to be much less serious in the real world (Drury et al, 1995).

The hypothesis that cosmic rays are produced by this process operating at the strong shocks associated with supernova remnants works quite well in terms of the general composition (even accounting for the so-called FIP biases when dust sputtering is incorporated), total power and energy spectrum up to the "knee" region. Unfortunately it fails catastrophically thereafter. Without going into detail the maximum energy that can be reached by shock acceleration in a remnant of radius R, velocity R' and magnetic field strength B is of order (as might be guessed on dimensional grounds) charge times BRR'. If we substitute $R \approx 10\,\mathrm{pc}$, $R' \approx 10^3\,\mathrm{km\,s^{-1}}$ and $B \approx 3\,\mu\mathrm{G}$ we find that protons cannot be accelerated to more than $10^{15}\,\mathrm{eV}$ whereas the observed cosmic ray spectrum goes on for at least another five decades! Furthermore the transition is extremely smooth although there is a change of slope (the "knee" in the spectrum). Although most unsatisfactory, it seems that if we wish to use SNRs to generate the particles below the "knee" we must then invoke a second source for the particles above the "knee". But the circumstantial evidence is quite strong, the mechanism is plausible, and there is no other rival theory, so we should look for direct observational evidence.

V. OBSERVATIONAL EVIDENCE FOR ACCELERATION IN SNRS

There is actually good evidence for the acceleration of *electrons* in SNRs. The most convincing case comes from radio and X-ray studies of young remnants like those of Tycho's SN and SN1006 which have simple spherical morphology (presumably reflecting a Type Ia supernova origin). These have non-radiative shocks, therefore modest compression, and thus the radio emission cannot be simple van der Laan compression but must result from relativistic electron acceleration. The extreme sharpness of the radio rims (Achterberg et al, 1994) is a strong indication of locally generated wave turbulence, small diffusion coefficients and ongoing injection and acceleration in the shock front. The mainly radial magnetic field structure revealed by polarization studies also indicates strong turbulence. The spectral indices seen in the radio can be interpreted in terms of acceleration at shocks modified by an accelerated ion component, and this may be weak evidence for ion acceleration; otherwise the radio observations, while convincing proof of electron acceleration to GeV energies, say nothing about acceleration of ions. The exciting new result in this field is the detection (Koyama et al 1995) by the ASCA satellite of nonthermal X-ray emission in parts of the remnant of SN1006 which is interpreted as synchrotron emission from TeV electrons (Reynolds, 1996; Mastichiadis and de Jager, 1996).

Of course one can always argue that if electrons are being accelerated to GeV and TeV energies, then the ions should also be accelerated. But it is desirable to have some more direct evidence than hints from the radio spectral indices. One obvious possibility is to use gamma-ray observations. If SNRs are filled with freshly accelerated protons (and other ions) these will interact with the ejecta and the sweptup interstellar medium to produce, among other products, π^0 particles which then decay to produce gamma-rays. It is relatively straightforward to estimate the expected fluxes of gamma-rays from SNRs in the GeV and TeV regions with the result that at least some of the nearby SNRs should indeed be detectable sources (Dorfi, 1990, 1991; Drury et al, 1994; Naito and Takahara, 1994). However while there do appear to be real detections in the GeV region from the EGRET instrument on the Compton gamma-ray observatory satellite (Esposito et al, 1996) searches in the TeV region and above have so far only yielded upper limits (G E Allen et al, 1995; W H Allen et al, 1995; Lessard et al, 1995; Prosch et al, 1995).

In conclusion there is very good evidence for electron acceleration in SNRs, and some evidence for ion acceleration although the absence of TeV gamma-ray detections is beginning to be worrying. However at best this is only half the story. We still need to explain the remaining six decades or so of the spectrum beyond the "knee". Pulsars appear to be the only half-plausible location for acceleration from the "knee" region to the "ankle" in the Galaxy. However the smoothness of the spectrum in the "knee" region, as revealed in the latest results from the Tibet ASγ collaboration (Amenomori et al, 1996) is a real problem. If, as seems probable, the ultra-high energy cosmic rays are of extragalactic origin it is interesting to note that they are the only sample of extra-galactic matter which is ever likely to reach the solar system.

ACKNOWLEDGEMENTS

This review is substantially the same as one prepared for the Spring meeting of the Hungarian Astronomical Society in Visegrad although with slightly changed emphases. I would like to record my appreciation of the hospitality of the local organisers in Pusan and to thank

them for inviting me to give this review.

REFERENCES

Achterberg A., Blandford R.D., Reynolds S.P., 1994, A&A, 281, 220

Allen G. E., et al. 1995, ApJ, 448, L25

Allen W. H. et al, 1995, Proc 24th ICRC (Rome), 2, 447

Amenomori M et al, 1996, ApJ, 461, 408

Berezhko E G, Elshin V K and Ksenofontov L T, 1996, Zh. Eksp. Teor. Fiz., 109, 3 (English translation in JETP, 82, 1)

Dorfi E A, 1990, A&A, 234, 419

Dorfi E A, 1991, A&A, 251, 597

Drury L O'C, Aharonian F and Völk H J, 1994, A&A, 267, 959

Drury L O'C, Völk H J and Berezhko E G, 1995, A&A, 299, 222

Esposito J A, Hunter S D, Kanbach G and Sreekumar P, 1996, ApJ, 461, 820

Koyama K, Petre R, Gotthelf E V, Matsuura M, Ozaki M and Holt S S, 1995, Nature, 378, 255

Lessard R. W. et al., 1995 Proceedings of the 24th. Int. Cosmic Ray Conf, Rome, 2, 475

Mastichiadis A and de Jager O C, 1996, A&A, 311, 5

Naito T and Takahara F, 1994, J Phys G, 20, 477

Prosch C et al., 1995, Proc 24th ICRC (Rome), 2, 405

Reynolds, S. P., 1996, ApJ, 459, L13