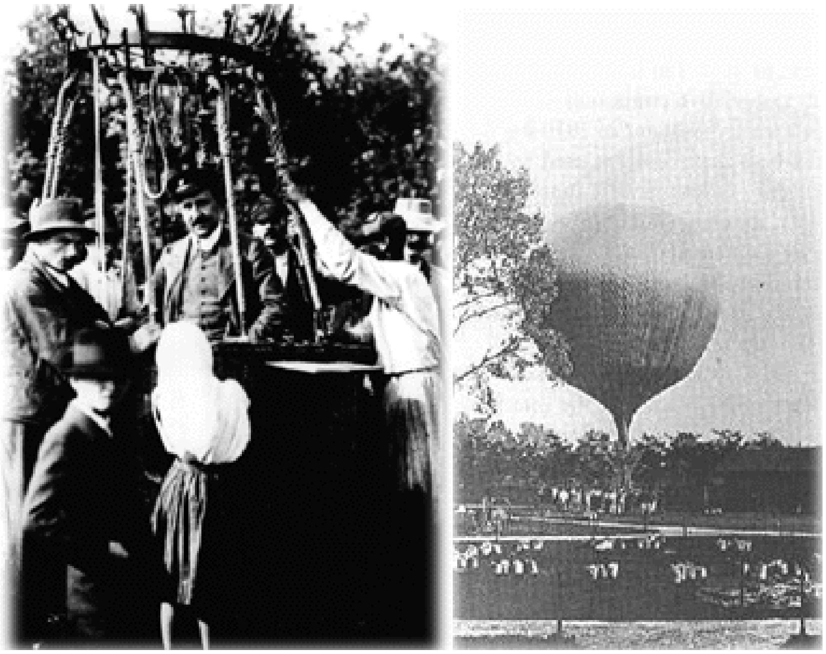
What are cosmic rays?

At the start of the 20th century scientists became very interested in a puzzling phenomena. There seemed to be rather more radiation in the environment than they could account for by the known sources of natural background radioactivity.

After much debate, the puzzle was partly solved by a daring German scientist, Victor Hess. In 1912 he took a radiation counter (he used a gold leaf electroscope) on a balloon flight. He risked his life, by travelling to 17,500 feet without oxygen, but managed to observe that the amount of radiation increased as his balloon climbed. This demonstrated that the radiation was from outer space and eventually it was dubbed "Cosmic Radiation".



Left: Victor Hess before his balloon flight, during which he observed cosmic ray intensity increasing with altitude. Right: Hess's balloon.

Since 1912 we have learnt a lot about cosmic rays. We now know that they are sub-atomic particles and possess a large range of energies (usually measured in electron-volts [eV]) from a few billion eV to more than 10^{20} eV.

The rate at which cosmic rays bombard the Earth varies enormously with their energy. Low energy cosmic rays are plentiful (many thousand per square metre every second). The highest energy cosmic rays are very rare (less than one hits a square kilometre of the Earth's surface each century). This makes detecting them very difficult.

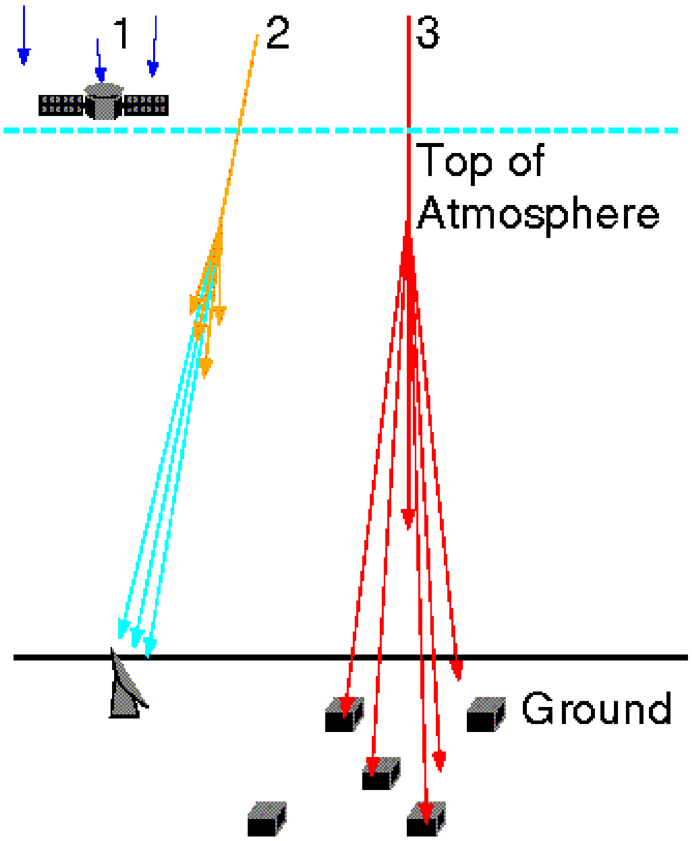
We know from measurements made on board satellites and high altitude balloons that the vast majority of cosmic rays are protons, although other heavier atomic nuclei are also present, extending all the way up to uranium nuclei. The vast majority of cosmic ray particles therefore have a positive electrical charged.

A small fraction (0.1%) of cosmic rays are photons (in the form of gamma-rays). These gamma-ray photons are important when trying to find the origin of cosmic rays since they have no electrical charge and so arrive at the Earth undeflected by the galactic magnetic field.

http://ast.leeds.ac.uk/haverah/cosrays.shtml

How do we study cosmic rays?

Cosmic rays are studied in a variety of ways depending on how much energy they have. This is illustrated in the picture below.



How cosmic rays are detected

1. Low energy cosmic rays detected by instruments carried in satellites.
2. Higher energy cosmic rays generate a small air shower. The Cerenkov radiation emitted by the shower is detected by a large telescope on the ground.
3. Even higher energy cosmic rays generate very big air showers. The particles in the shower travel to the ground where they can be recorded by an array of detectors.

The lowest energy cosmic rays are absorbed in the upper atmosphere and so they can only be detected by equipment on board satellites and high altitude balloons.

Extensive Air Showers

When higher energy cosmic rays hit the upper atmosphere (about 20 km up) they lose about half of their energy by creating a jet of particles which carries on travelling in almost the same direction as the cosmic ray. The particles in the jet can themselves create more particles as they hit other nuclei of oxygen or nitrogen in the air. This jet is called an extensive air shower and keeps on growing until the particles in the shower run out of energy and are absorbed in the atmosphere.

We refer to the initial particle that starts the shower as the primary cosmic ray. The particles created in the air shower are known as secondary cosmic rays. Over a million of the secondary particles which are produced when primary cosmic rays hit the atmosphere pass right through your body every minute.

A single cosmic ray can generate showers with a large number of particles depending on its energy. The smaller air showers are absorbed near the top of the atmosphere and do not reach ground level. However, as the particles in the shower zip through the air they emit faint flashes of blue light known as Cerenkov radiation. Although the cosmic rays and the air showers they produce are absorbed by the atmosphere it is possible to detect the faint Cerenkov light using large telescopes but only on dark, moonless nights. The Leeds University group collaborate with scientists in the USA and Ireland at the whipple telescope in Arizona and use this technique to observe high energy gamma rays from dead stars such as the Crab nebula and the centres of very active galaxies.

Air Shower Arrays

At even higher energies the air showers contain vast numbers of secondary particles, numbering in the billions for the most energetic cosmic rays. The particles in these showers are of such high energy that they can travel all the way from the top of the atmosphere (about 20 kilometres up) down to the ground where they can be detected directly with particle detectors.

For example: A primary cosmic ray enters our atmosphere. At an altitude of ~20 km it collides with molecules in the air and generates a shower of secondary particles. These also generate further particles which travel, at almost light speed, towards the ground where some are detected by an air shower array. In this example, suppose the shower hits the detectors to the left before those on the right. This helps us to determine the direction of the primary cosmic ray. There are more particles at the centre, or core, of the shower. Most of the secondary particles are absorbed in the ground, but some of the higher energy particles in the core can penetrate many kilometres below ground where they can be detected by experiments such as AMANDA.

The detectors are usually arranged in a grid formation (or array) on the ground allowing measurements of each shower to be made at several points. Information from the detectors tell us how many particles struck the detector and the time that they hit. By adding up the number of particles recorded by each of the detectors we can estimate how many particles were in the shower and from that we can make a good guess as to the energy of the cosmic ray that started the shower. We can use the time that each detector was struck to measure the direction the cosmic ray was travelling when it hit the Earth's atmosphere.

Its important to realise that when we measure extensive air showers, we do not "see" the primary cosmic ray. Rather we measure the secondary particles that were generated as the cosmic ray travelled through our atmosphere.

The air showers recorded by the SPASE-2 array at the South Pole have diameters of 10's of metres at ground level and so the detectors in these arrays are spaced between 30 and 50 metres apart. The very highest energy cosmic rays produce air showers which cover many square kilometres. For this reason the planned Pierre Auger Observatory will have 1600 detectors spaced 1.5 km apart.

`http://ast.leeds.ac.uk/haverah/dets.shtml`

How do our detectors work?

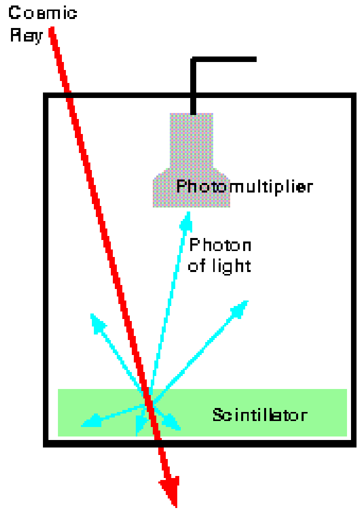
In Leeds we use two kinds of detectors:

Scintillator detectors.
Water Cerenkov detectors.

Both types of detectors are relatively simple. Here you can find out how they work.

Scintillator Detectors

How Scintillator Detectors work: The scintillator detector is made up of a special piece of plastic called a `scintillator'. When fast moving, charged particles, such as cosmic rays pass through the scintillator they excite the atoms in the plastic by giving them some energy (the cosmic ray then slows down a little). The excited atoms then lose this energy by emitting some photons of light. The light is detected by a sensitive piece of equipment called a "photomultiplier".

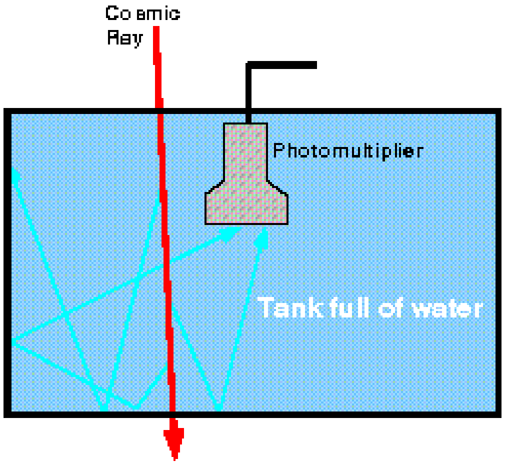


A photomultiplier tube

The photomultiplier, as its name suggests, multiplies the small flash of light into a large electrical signal that can be measured. From the size of the electronic signal we can tell how many particles passed through the scintillator. The scintillator and the photomultiplier are housed in a dark box so that the only light detected is caused by cosmic rays. This kind of detector is illustrated in the diagram below. These kind of detectors are used in the SPASE-2 array.

Water Cerenkov Detectors

Water Cerenkov detectors are similar to the scintillator detectors except that the dark box contains no scintillator but is filled with pure, clear water. When cosmic rays pass through the water they emit faint flashes of blue light known as Cerenkov radiation. The sides of the water tank are lined with reflective material and some of this light is reflected onto a photomultiplier which produces an electronic signal. The size of the signal can be used to find out how many cosmic rays passed through the detector. This kind of detector was used in the 12 square kilometre array at Haverah Park. A diagram of this kind of detector is shown here.



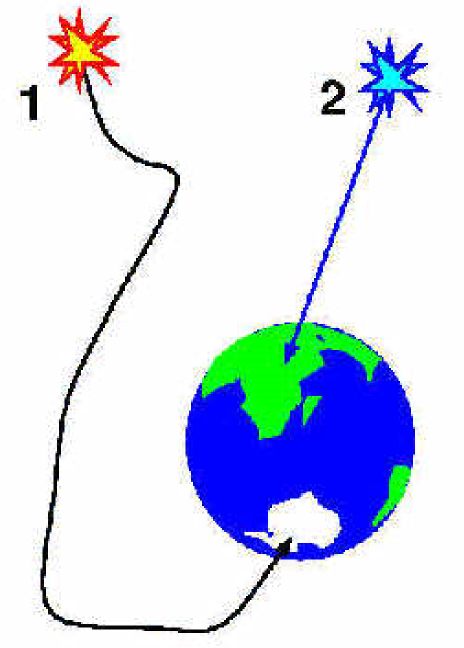
How Water Cerenkov detectors work: When cosmic rays pass through the detector, photons of Cerenkov light are emitted. These reflect off the sides of the water tank and some hit the photomultiplier, which creates an electronic signal.

http://ast.leeds.ac.uk/haverah/detector.shtml

Where do Cosmic Rays come from?

The origin of cosmic rays is still not known and is the burning question in high energy astrophysics research.

Experiments such as the SPASE array were designed to measure the direction that cosmic rays are travelling in when they hit the Earth. Unfortunately this does not tell us where the cosmic ray came from. The problem is that cosmic rays carry electric charge and do not travel in straight lines. Their trajectories are bent by the magnetic fields that are known to exist between stars and galaxies.



Cosmic Ray trajectories. The diagram illustrates the trajectories of cosmic rays and gamma rays from their point of origin to the Earth.

1. Electrically charged cosmic rays are bent by interstellar magnetic fields and do not travel in straight lines. When we measure their trajectory at the Earth we cannot tell where they came from.
2. Gamma rays are neutral particles and so travel in straight lines. If we can measure their trajectory when they hit the Earth, then we can see where they came from.

Unlike cosmic rays, gamma rays carry no electric charge and so are not deflected by magnetic fields. The telescope at the whipple observatory used by the Leeds group has seen gamma rays coming from the Crab Nebula and more exotic objects such as Active Galactic Nuclei. It was hoped that the SPASE array might also detect gamma ray emission from objects in the sky but at the higher energy at which SPASE operates no gamma ray sources were detected.

The very highest energy cosmic rays may come from outside our galaxy, and are deflected much less by the magnetic fields due to their high momentum. For this reason there is great interest in detecting large numbers of these particles in the hope of discovering where they come from. Unfortunately these very high energy cosmic rays are very rare (at 10^{20} eV only 1 cosmic ray hits each square kilometre of the Earth's surface per century!) and so a giant air shower array must be built in order to detect them. The proposed Pierre Auger Observatory would be just such an array.

http://ast.leeds.ac.uk/haverah/origin.shtml

The Very Highest Energy Cosmic Rays

The very highest energy cosmic rays are of particular interest for various reasons. They may provide a useful tool for finding the origin of cosmic rays because they are deflected very little by the galactic and interstellar magnetic fields that permeate space. Therefore the direction in which they are travelling when they arrive at Earth should point back to the area of space where they came from.

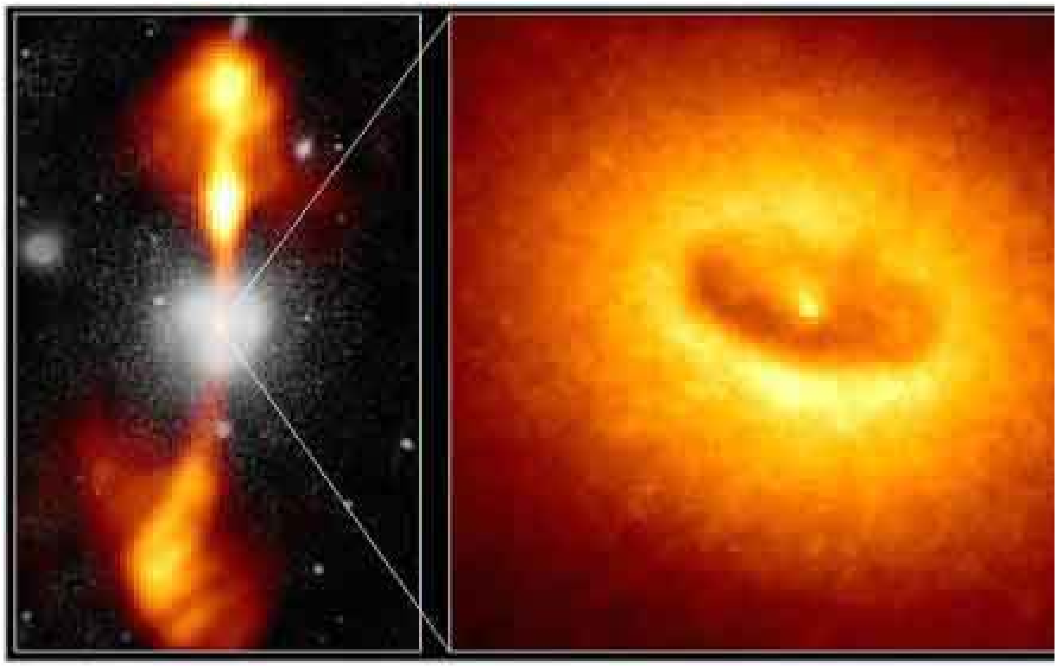
There are many unanswered questions regarding their production:-

How are they produced?

Mechanisms to accelerate particles up to energies of 10^{15} eV have been proposed and generally consist of binary star systems (two stars in orbit around each other) or supernova remnants (the turbulent shell of gas left behind after a star has exploded). However the acceleration mechanisms involved in producing the highest energy cosmic rays are still unknown. There may even be new physics involved. One possibility is that they are generated by very massive particles produced at the beginning of time.

Where are they produced?

The places in the universe where cosmic rays of $>10^{18}$ eV are produced must either have very large magnetic fields or be of enormous size. If the highest energy cosmic rays come from within our galaxy the production sites would be expected to be relatively close to Earth because the galaxy cannot trap such energetic particles within its magnetic field (unlike lower energy cosmic rays) and they would readily escape after travelling a short distance. If this were the case, we would expect to see more high energy cosmic rays coming from the direction of the galactic plane than from elsewhere.



Do the highest energy Cosmic Rays come from objects such as this? On the left is a radio telescope image of the active galaxy NGC-4261. The width of the image covers a region of space 88,000 light years across. To the right is a close up image from the Hubble Space Telescope covering just 400 light years. It shows a doughnut shaped ring at the centre. This ring is thought to orbit a giant black hole with a mass more than a hundred million times that of our sun.

However, if the highest energy cosmic rays come from outside our own galaxy they would not be able to travel for more than about 150 million light years. This may sound like a long way but it is only the distance to some of our neighbouring galaxies. The reason the high energy cosmic rays can not travel further is that as they race through space they occasionally bump into photons of microwave radiation. This radiation is the light left over from the big-bang explosion that created the universe. When the high energy cosmic rays hit microwave photons, they lose some of their energy. This effect is known as the "GKZ cut-off" and because of this many physicists believed no cosmic rays with energies above about $4x10^{19}$ eV existed. However experiments such as that at Haverah Park, the Fly's Eye(USA), and AGASA(Japan) have shown that particles above the GKZ cut-off do exist.

What is the highest energy cosmic ray ever detected?

In 1993 the "Fly's Eye" experiment in Utah detected a cosmic ray with an energy of $3x10^{20}$ eV. So far this is the highest energy particle ever detected. This particle had a kinetic energy similar to that carried by a tennis ball travelling at 180 mph! Cosmic Rays are 10^{14} times smaller than tennis balls so the energy is packed into an incredibly small volume.

How can we study the highest energy cosmic rays?

Cosmic rays with energies above 10^{20} eV are very rare. On average one such particle hits each square kilometre of the Earth only once a century. So to detect a large number of them and study them in detail we need a huge detector. The Pierre Auger Observatory will eventually consist of two observatories, one in the northern hemisphere and one in the south. Each will have 1600 water Cerenkov detectors spread over 3000 km^2 and 3 fluorescence light detectors. The first part of the observatory will be built in Argentina. Construction commenced on March 17th 1999. When this observatory is completed sometime in the next century we may finally be able to answer the question. "Where do cosmic rays come from?".

3f

Cosmic rays reveal their roots 25 April 2002 Physicsweb.org

Two groups of scientists have shed light on the origin of the streams of high-energy particles known as cosmic rays that continually bombard the Earth. Ryoji Enomoto of the University of Tokyo and co-workers have found the first strong evidence that cosmic rays with energies up to 10^{15} eV are produced by remnants of supernovas (R Enomoto *et al* 2002 *Nature* 416 823). Meanwhile, a team of researchers from NASA and Princeton University has proposed that cosmic rays with energies of over 10^{20} eV are made by black holes in ancient quasar galaxies.

Cosmic rays were first detected in 1912 but there is still no consensus on where they are produced or how they are accelerated to such high energies. Scientists have speculated that supernovas – the huge explosions produced by collapsing stars – could be responsible. This is because the combined energy of cosmic rays in our galaxy is a significant fraction of the total energy released by galactic supernovae. In addition, the mechanism by which this energy could be transferred – through the shock waves generated by supernovae – can account for the observed energy distribution of the cosmic rays that reach the Earth.

The observations by Enomoto and colleagues support this theory. Using the CANGAROO telescope in Australia, they detected showers of optical photons resulting from gamma-rays hitting the Earth's upper atmosphere with energies of about 10^{12} eV (1 TeV), from the direction of the supernova remnant RX J1713.7-3946. Such gamma rays could result from the decay of short-lived particles called pions, which are produced by the interaction of protons – the main constituent of cosmic rays – with the interstellar gas surrounding a supernova remnant.

Gamma rays with energies of the order of 1 TeV have previously been detected from two other supernova remnants. But in these cases the gamma rays could have been produced by high-energy electrons that scattered and energized photons from the microwave radiation left over from the big bang, the so-called cosmic microwave background. In contrast, the energy spectrum of the gamma rays detected by Enomoto and colleagues closely matches that expected from the radiation produced by protons rather than electrons.

In a related discovery, Diego Torres of Princeton University and Elihu Boldt and colleagues at NASA's Goddard Space Flight Center have found that four elliptical galaxies relatively close to Earth may be responsible for cosmic rays with energies of at least 10^{20} eV. These ultra-high-energy cosmic rays must originate from within 200 million light years of Earth, otherwise their energy would be diminished by interactions with the cosmic microwave background. At a press conference earlier this week the scientists announced that these cosmic rays appear to arrive on Earth from the direction of these galaxies.

But in order to generate cosmic rays, the supermassive black holes known to exist at the cores of these galaxies must spin. Torres and colleagues admit that they do not know if this is the case, but point out that at least one supermassive black hole in the universe is known to spin.

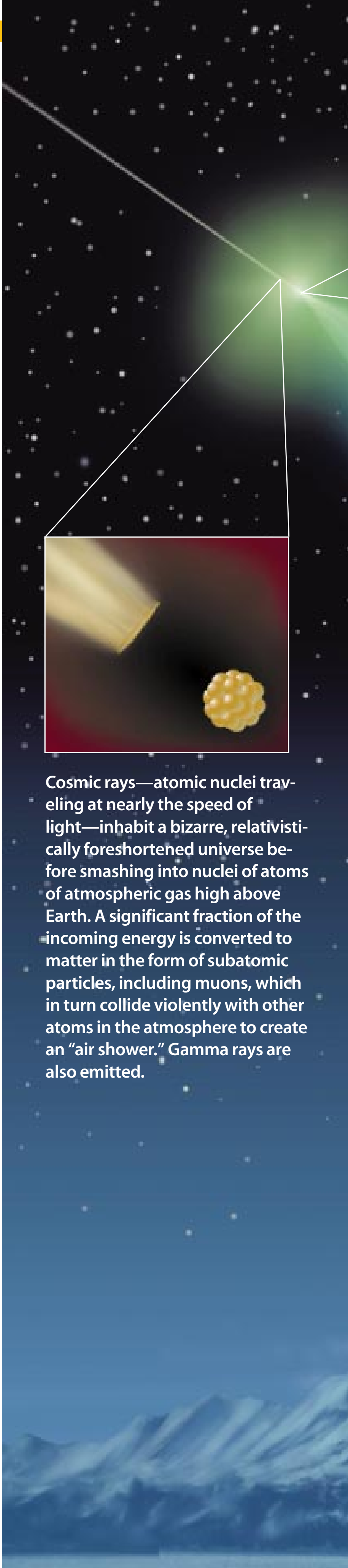
About the author

Edwin Cartlidge is News Editor of *Physics World*

Cosmic Rays at the Energy Frontier

These particles carry more energy than any others in the universe. Their origin is unknown but may be relatively nearby

by James W. Cronin, Thomas K. Gaisser and Simon P. Swordy



Cosmic rays—atomic nuclei traveling at nearly the speed of light—inhabit a bizarre, relativistically foreshortened universe before smashing into nuclei of atoms of atmospheric gas high above Earth. A significant fraction of the incoming energy is converted to matter in the form of subatomic particles, including muons, which in turn collide violently with other atoms in the atmosphere to create an "air shower." Gamma rays are also emitted.

MICHAEL GOODMAN

Roughly once a second, a subatomic particle enters Earth's atmosphere carrying as much energy as a well-thrown rock. Somewhere in the universe, that fact implies, there are forces that can impart to a single proton 100 million times the energy achievable by the most powerful Earthbound accelerators. Where and how?

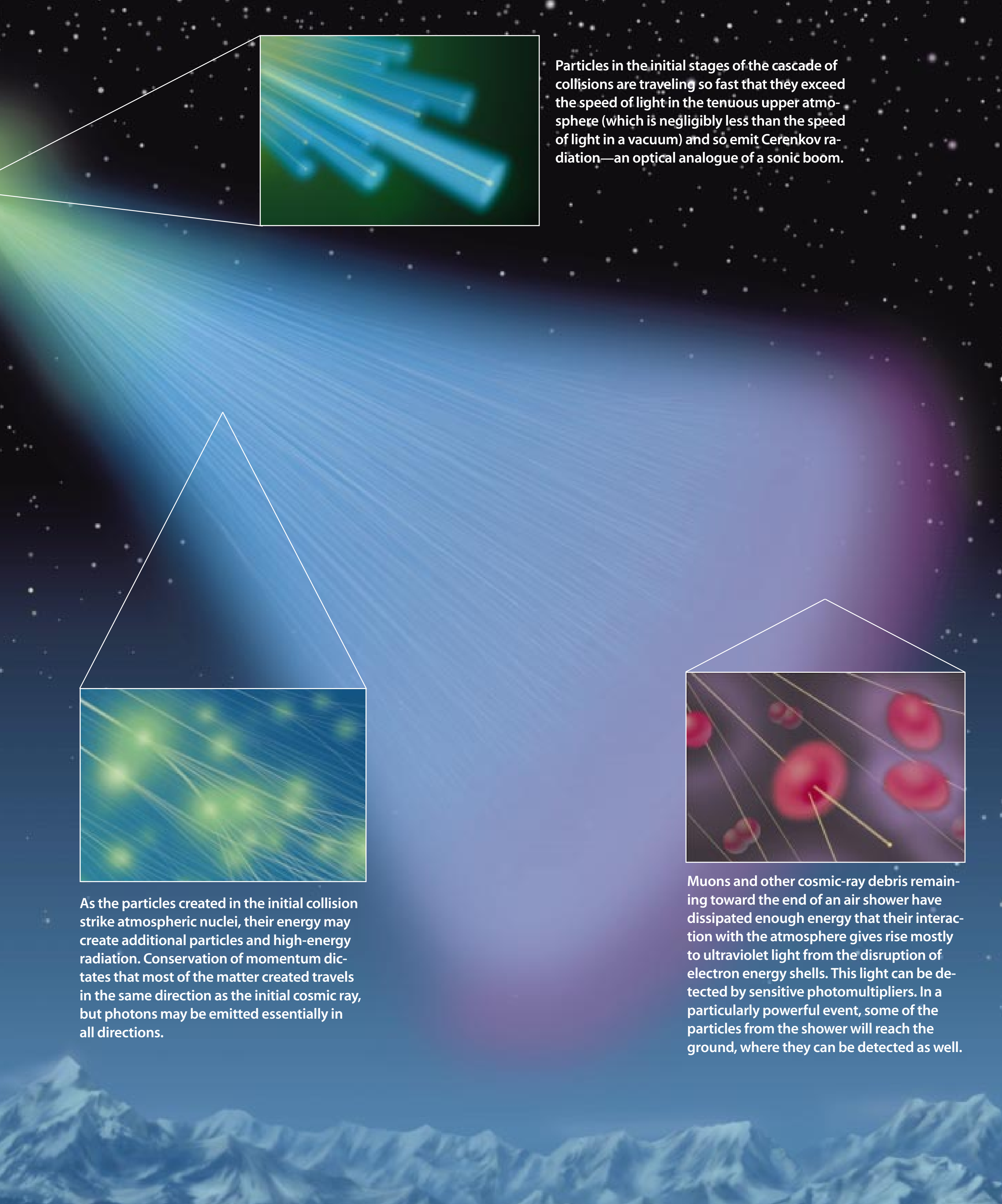
Those questions have occupied physicists since cosmic rays were first discovered in 1912 (although the entities in question are now known to be particles, the name "ray" persists). The interstellar medium contains atomic nuclei of every element in the periodic table, all moving under the influence of electrical and magnetic fields. Without the screening effect of Earth's atmosphere, cosmic rays would pose a significant health threat; indeed, people living in mountainous regions or making frequent airplane trips pick up a measurable extra radiation dose.

Perhaps the most remarkable feature of this radiation is that investigators have not yet found a natural end to the cosmic-ray spectrum. Most well-known sources of charged particles—such as the sun, with its solar wind—have a characteristic energy limit; they simply do not produce particles with energies above this limit. In contrast, cosmic rays appear, albeit in decreasing numbers, at energies as high as astrophysicists can measure. The data run out at levels around 300 billion times the rest-mass energy of a proton because there is no detector large enough to sample the very low number of incoming particles predicted.

Nevertheless, evidence of ultrahigh-energy cosmic rays has been seen at intervals of several years as particles hitting the atmosphere create myriad secondary particles (which are easier to detect). On October 15, 1991, for example, a cosmic-ray observatory in the Utah desert registered a shower of secondary particles from a 50-joule (3×10^{20} electron volts) cosmic ray. Although the cosmic-ray flux decreases with higher energy, this decline levels off somewhat above about 10^{18} eV, suggesting that the mechanisms responsible for ultrahigh-energy cosmic rays are different from those for rays of more moderate energy.

In 1960 Bernard Peters of the Tata Institute in Bombay suggested that lower-energy cosmic rays are produced predominantly inside our own galaxy, whereas those of higher energy come from more distant sources. One reason to think so is that a cosmic-ray proton carrying more than 10^{19} eV, for example, would not be deflected significantly by any of the magnetic fields typically generated by a galaxy, so it would travel more or less straight. If such particles came from inside our galaxy, we might expect to see different numbers coming from various directions because the galaxy is not arranged symmetrically around us. Instead the distribution is essentially isotropic, as is that of the lower-energy rays, whose directions are scattered.

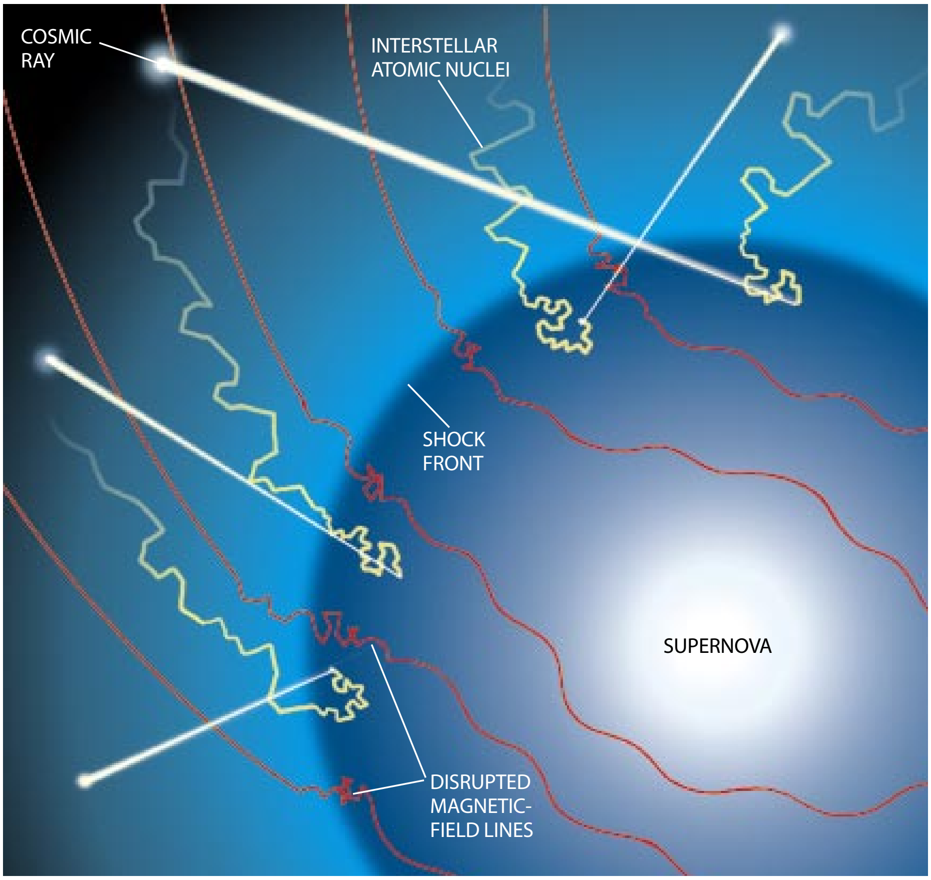
The Life of a Cosmic Ray



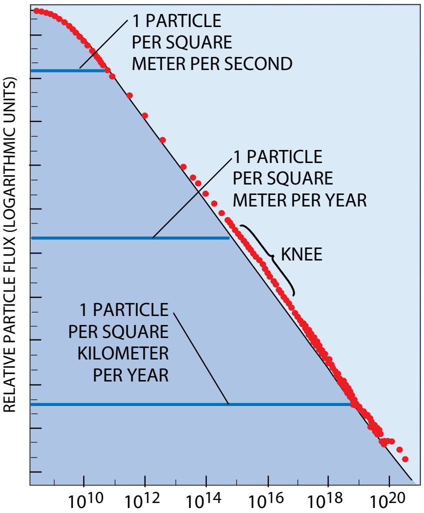
Particles in the initial stages of the cascade of collisions are traveling so fast that they exceed the speed of light in the tenuous upper atmosphere (which is negligibly less than the speed of light in a vacuum) and so emit Cerenkov radiation—an optical analogue of a sonic boom.

As the particles created in the initial collision strike atmospheric nuclei, their energy may create additional particles and high-energy radiation. Conservation of momentum dictates that most of the matter created travels in the same direction as the initial cosmic ray, but photons may be emitted essentially in all directions.

Muons and other cosmic-ray debris remaining toward the end of an air shower have dissipated enough energy that their interaction with the atmosphere gives rise mostly to ultraviolet light from the disruption of electron energy shells. This light can be detected by sensitive photomultipliers. In a particularly powerful event, some of the particles from the shower will reach the ground, where they can be detected as well.



GEORGE KELVIN



JENNIFER C. CHRISTIANSEN

Such tenuous inferences reveal how little is known for certain about the origin of cosmic rays. Astrophysicists have plausible models for how they might be produced but have no definitive answers. This state of affairs may be the result of the almost unimaginable difference between conditions on Earth and in the regions where cosmic rays are born. The space between the stars contains only about one atom per cubic centimeter, a far lower density than the best artificial vacuums we can create. Furthermore, these volumes are filled with vast electrical and magnetic fields, intimately connected to a diffuse population of charged particles even less numerous than the neutral atoms.

Supernova Pumps

This environment is far from the peaceful place one might expect: the low densities allow electrical and magnetic forces to operate over large distances and timescales in a manner that would be quickly damped out in material of terrestrial densities. Galactic space is therefore filled with an energetic and turbulent plasma of partially ionized gas in a state of violent activity. The motion is often hard to observe on human timescales because astronomical distances are so large; nevertheless, those same distances allow even moderate forces to achieve impressive results. A particle might zip through a terrestrial accelerator in a few microseconds, but it could spend years or even millennia in the accelerator's cosmic counterpart. (The timescales are further complicated by the strange, relativity-distorted framework that ultrahigh-energy cosmic rays inhabit. If we could observe such a particle for 10,000 years, that period would correspond to only a single second as far as the particle is concerned.)

Astronomers have long speculated that the bulk of galactic cosmic rays—those with energies below about 10^{16} eV—originate with supernovae. A compelling reason for this theory is that the power required to maintain the observed supply of cosmic-ray nuclei in our Milky Way galaxy is only slightly less than the average kinetic energy delivered to the galactic medium by the three supernova explosions that occur every century. There are few, if any, other sources of this amount of power in our galaxy.

When a massive star collapses, the outer parts of the star explode at speeds of up to 10,000 kilometers (6,000 miles) per second and more. A similar amount of energy is released when a white dwarf star undergoes complete disintegration in a thermonuclear detonation. In both types of supernovae the ejected matter expands at supersonic velocities, driving a strong shock into the surrounding medium. Such shocks are expected to accelerate nuclei from the material they pass through, turning them into cosmic rays. Because cosmic rays are charged, they follow complicated paths through interstellar magnetic fields. As a result, their directions as observed from Earth yield no information about the location of their original source.

By looking at the synchrotron radia-

AIR-SHOWER DETECTOR watches for traces of cosmic rays entering the upper atmosphere. Photodetectors can track flashes of light caused by particles interacting with air molecules and determine the energy and probable identity of the incoming rays. The Fly's Eye detector (*close-up at far right*) is located in Utah.

B. BOLA *U.S. Army*

COSMIC-RAY ACCELERATOR
is believed to arise from a supernova explosion. Astrophysicists hypothesize that atomic nuclei crossing the supernova shock front will pick up energy from the turbulent magnetic fields embedded in the shock. A particle may be deflected in such a way that it crosses the boundary of the shock hundreds or even thousands of times, picking up more energy on each passage, until it escapes as a cosmic ray. Most of the particles travel on paths that result in relatively small accelerations, accounting for the general shape of the cosmic-ray energy spectrum (*far right*), which falls off at higher energies. The "knee," or bend, in the curve suggests that most of the particles are accelerated by a mechanism incapable of imparting more than about 10^{15} electron volts. The relative excess of ultrahigh-energy particles indicates an additional source of acceleration whose nature is as yet unknown.

tion sometimes associated with supernova remnants, researchers have found more direct evidence that supernovae can act as accelerators. Synchrotron radiation is characteristic of high-energy electrons moving in an intense magnetic field of the kind that might act as a cosmic-ray accelerator, and the presence of synchrotron x-rays in some supernova remnants suggests particularly high energies. (In Earthbound devices, synchrotron emission limits a particle's energy because the emission rate increases as a particle goes faster; at some point, the radiation bleeds energy out of an accelerating particle as fast as it can be pumped in.) Recently the Japanese x-ray satellite Asca made images of the shell of Supernova 1006, which exploded 990 years ago. Unlike the radiation from the interior of the remnant, the x-radiation from the shell has the features characteristic of synchrotron radiation. Astrophysicists have deduced that electrons are being accelerated there at up to 10^{14} eV.

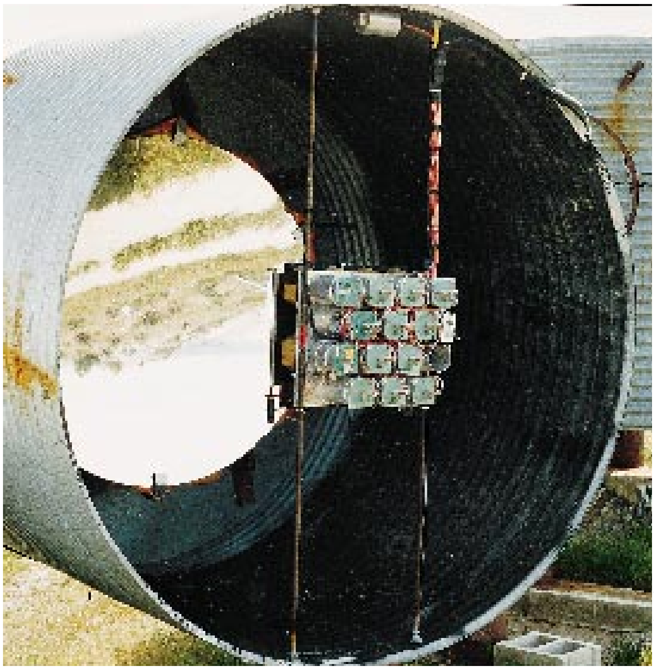
The EGRET detector on the Compton Gamma Ray Observatory has also been used to study point sources of gamma rays identified with supernova remnants. The observed intensities and spectra (up to a billion electron volts) are consistent with an origin from the decay of particles called neutral pions, which could be produced by cosmic rays from the exploding star's remnants colliding with nearby interstellar gas. Interestingly, however, searches made by the ground-based Whipple Observatory for gamma rays of much higher energies from some of the same remnants have not seen signals at the levels that would be expected if the supernovae were accelerating protons to 10^{14} eV or more.

A complementary method for testing the association of high-energy cosmic rays with supernovae involves the elemental composition of cosmic-ray nuclei. The size of the orbit of a charged particle in a magnetic field is proportional to its total momentum per unit charge, so heavier nuclei have greater total energy for a given orbit size. Any process that limits the particle acceleration on the basis of orbit size (such as an accelerating region of limited extent) will thus lead to an excess of heavier nuclei at high energies.

Eventually we would like to be able to go further and look for elemental signatures of acceleration in specific types of supernovae. For example, the supernova of a white dwarf detonation would accelerate whatever nuclei populate the local interstellar medium. A supernova that followed the collapse of a massive star, in contrast, would accelerate the surrounding stellar wind, which is characteristic of the outer layers of the progenitor star at earlier stages of its evolution. In some cases, the wind could include an increased fraction of helium, carbon or even heavier nuclei.

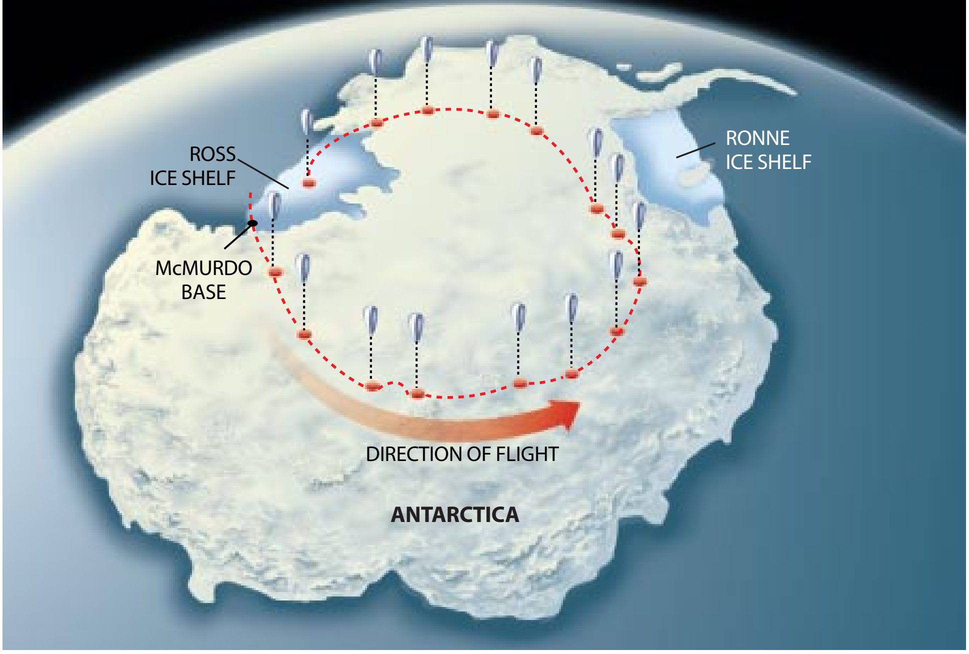
The identity of high-energy cosmic rays is all but lost when they interact with atoms in Earth's atmosphere and form a shower of secondary particles. Hence, to be absolutely sure of the nuclear composition, measurements must be made before the cosmic rays reach dense atmosphere. Unfortunately, to collect 100 cosmic rays of energies near 10^{15} eV, a one-square-meter detector would have to be in orbit for three years. Typical exposures at present are more like the equivalent of one square meter for three days.

Researchers are attacking this problem with some ingenious experiments. For example, the National Aeronautics and Space Administration has developed techniques to loft large payloads (about three metric tons) with high-altitude bal-



UNIVERSITY OF UTAH

HIGH-ALTITUDE BALLOON launched near McMurdo Base in Antarctica carries cosmic-ray detectors above most of the atmosphere. Winds 40 kilometers above the ice cap blow in a circle around the Pole, returning the balloon to the vicinity of its starting point after about 10 days. Balloon detectors are not as sensitive as those placed on board satellites, but they can be made much larger and lofted much more cheaply.



GEORGE KELVIN

loons for many days. These experiments cost a tiny fraction of what an equivalent satellite detector would. The most successful flights of this type have taken place in Antarctica, where the upper atmosphere winds blow in an almost constant circle around the South Pole.

A payload launched at McMurdo Sound on the coast of Antarctica will travel at a nearly constant radius from the Pole and return eventually to near the launch site. Some balloons have circled the continent for 10 days. One of us (Swordy) is collaborating with Dietrich Müller and Peter Meyer of the University of Chicago on a 10-square-meter detector that could measure heavy cosmic rays of up to 10^{15} eV on such a flight. There are efforts to extend the exposure times to roughly 100 days with similar flights nearer the equator.

Across Intergalactic Space

Studying even higher-energy cosmic rays—those produced by sources as yet unknown—requires large ground-based detectors, which overcome the problem of low flux by watching enormous areas for months or years. The information, however, must be extracted from cascades of secondary particles—electrons, muons and gamma rays—initiated high in the atmosphere by an incoming cosmic-ray nucleus. Such indirect methods can only suggest general features of the composition of a cosmic ray on a statistical basis, rather than identifying the atomic number of each incoming nucleus.

At ground level, the millions of secondary particles unleashed by one cosmic ray are spread over a radius of hundreds of meters. Because it is impractical to blanket such a large area with detectors, the detectors typically sample these air showers at a few hundred or so discrete locations.

Technical improvements have enabled such devices to collect increasingly sophisticated data sets, thus refining the conclusions we can draw from each shower. For example, the CASA-MIA-DICE experiment in Utah, in which two of us (Cronin and Swordy) are involved, measures the distributions of electrons and muons at ground level. It also detects Cerenkov light (a type of optical shock wave produced by particles moving faster than the speed of light in their surrounding medium) generated by the shower particles at various levels in the atmosphere. These data enable us to reconstruct the shape of the shower more reliably and thus take a better guess at the energy and identity of the cosmic ray that initiated it.

The third one of us (Gaisser) is working with an array that measures showers reaching the surface at the South Pole. This experiment works in conjunction with AMANDA, which detects energetic muons produced in the same showers by observing Cerenkov radiation produced deep in the ice cap. The primary goal of AMANDA is to catch traces of neutrinos produced in cosmic accelerators, which may generate upward-streaming showers after passing through Earth.

Cosmic rays with energies above 10^{20} eV strike Earth's atmosphere at a rate of only about one per square kilometer a century. As a result, studying them requires an air-shower detector of truly gigantic proportions. In addition to the 1991 event in Utah, particles with energies above 10^{20} eV have been seen by groups elsewhere in the U.S., in Akeno, Japan, in Haverah Park, U.K., and in Yakutsk, Siberia.

Particles of such high energy pose a conundrum. On the one hand, they are likely to come from outside our galaxy because no known acceleration mechanism could produce them and because they approach from all directions even though a galactic magnetic field is insufficient to bend their path. On the other hand, their source cannot be more than about 30 million light-years away, because the particles would otherwise lose energy by interaction with the universal microwave background—radiation left over from the birth of the cosmos in the big bang. In the relativistic universe that the highest-energy cosmic rays inhabit, even a single radio-frequency photon packs enough punch to rob a particle of much of its energy.

If the sources of such high-energy particles were distributed uniformly throughout the cosmos, interaction with the microwave background would cause a sharp cutoff in the number of particles with energy above 5×10^{19} eV, but that is not the case. There are as yet too few events above this nominal threshold for us to know for certain what is going on, but even the few we have seen provide us with a unique opportunity for theorizing. Because these rays are essentially undeflected by the weak intergalactic magnetic fields, measuring the direction of travel of a large enough sample should yield unambiguous clues to the locations of their sources.

It is interesting to speculate what the sources might be. Three recent hypotheses suggest the range of possibilities: galactic black-hole accretion disks, gamma-ray bursts and topological defects in the fabric of the universe.

Astrophysicists have predicted that black holes of a billion solar masses or more, accreting matter in the nuclei of active galaxies, are needed to drive relativistic jets of matter far into intergalactic space at speeds approaching that of light; such

STEVEN PETERZEN *National Scientific Balloon Facility*

jets have been mapped with radio telescopes. Peter L. Biermann of the Max Planck Institute for Radioastronomy in Bonn and his collaborators suggest that the hot spots seen in these radio lobes are shock fronts that accelerate cosmic rays to ultrahigh energy. There are some indications that the directions of the highest-energy cosmic rays to some extent follow the distribution of radio galaxies in the sky.

The speculation about gamma-ray bursts takes off from the theory that the bursts are created by relativistic explosions, perhaps resulting from the coalescence of neutron stars. Mario Vietri of the Astronomical Observatory of Rome and Eli Waxman of Princeton University independently noted a rough match between the energy available in such cataclysms and that needed to supply the observed flux of the highest-energy cosmic rays. They argue that the ultrahigh-speed shocks driven by these explosions act as cosmic accelerators.

Rare Giants

Perhaps most intriguing is the notion that ultrahigh-energy particles owe their existence to the decay of monopoles, strings, domain walls and other topological defects that might have formed in the early universe. These hypothetical objects are believed to harbor remnants of an earlier, more symmetrical phase of the fundamental fields in nature, when gravity, electromagnetism and the weak and strong nuclear forces were merged. They can be thought of, in a sense, as infinitesimal pockets preserving bits of the universe as it existed in the fractional instants after the big bang.

As these pockets collapse, and the symmetry of the forces within them breaks, the energy stored in them is released in the form of supermassive particles that immediately decay into jets of particles with energies up to 100,000 times greater than those of the known ultrahigh-energy cosmic rays. In this scenario the ultrahigh-energy cosmic rays we observe are the comparatively sluggish products of cosmological particle cascades.

Whatever the source of these cosmic rays, the challenge is to collect enough of them to search for detailed correlations with extragalactic objects. The AGASA array in Japan currently has an effective area of 100 square kilometers and can capture only a few ultrahigh-energy events a year. The new Fly's Eye High Resolution experiment in Utah can see out over a much larger area, but only on clear, moonless nights.

For the past few years, Cronin and Alan A. Watson of the University of Leeds have spearheaded an initiative to gather an even larger sample of ultrahigh-energy cosmic rays. This development is named the Auger Project, after Pierre Auger, the French scientist who first investigated the phenomenon of correlated showers of particles from cosmic rays.

The plan is to provide a detection area of 6,000 square kilometers with a 100 percent duty cycle that is capable of measuring hundreds of high-energy events a year. A detector field would consist of many stations on a 1.5-kilometer grid; a single event might trigger dozens of stations. To cover the entire sky, two such detectors are planned, one each for the Northern and Southern hemispheres.

An Auger Project design workshop held at the Fermi National Accelerator Laboratory in 1995 has shown how modern off-the-shelf technology such as solar cells, cellular telephones and Global Positioning System receivers can make such a system far easier to construct. A detector the size of Rhode Island could be built for about $50 million.

Plans exist to cover even larger areas. Detectors in space could view millions of square kilometers of the atmosphere from above, looking for flashes of light signaling the passage of ultrahigh-energy particles. This idea, which goes by the name of OWL (Orbiting Wide-angle Light collectors) in the U.S. and by Airwatch in Europe, was first suggested by John Linsley of the University of New Mexico. To succeed, the project requires developing new technology for large, sensitive, finely segmented optics in space to provide the resolution needed. This development is under way by the U.S. National Aeronautics and Space Administration and in Italy.

As researchers confront the problem of building and operating such gigantic detector networks, the fundamental question remains: Can nature produce even more energetic particles than those we have seen? Could there be still higher-energy cosmic rays, or are we already beginning to detect the highest-energy particles our universe can create? SA

The Authors

JAMES W. CRONIN, THOMAS K. GAISSER and SIMON P. SWORDY work on both the theoretical questions of how cosmic rays are created and the practical problems inherent in detecting and analyzing them. Cronin, a professor of physics at the University of Chicago since 1971, earned his master's degree from the university in 1953 and his doctorate in 1955. In 1980 he shared the Nobel Prize with Val L. Fitch for work on symmetry violations in the decay of mesons. Gaisser, a professor of physics at the University of Delaware, has concentrated on the interpretation of atmospheric cosmic-ray cascades; he earned his doctorate from Brown University in 1967. In 1995 Gaisser spent two months in Antarctica setting up cosmic-ray detectors. Swordy, an associate professor at Chicago, has been active in cosmic-ray measurement since 1976. He earned his Ph.D. from the University of Bristol in 1979. This article updates a version that appeared in *Scientific American* in January 1997.

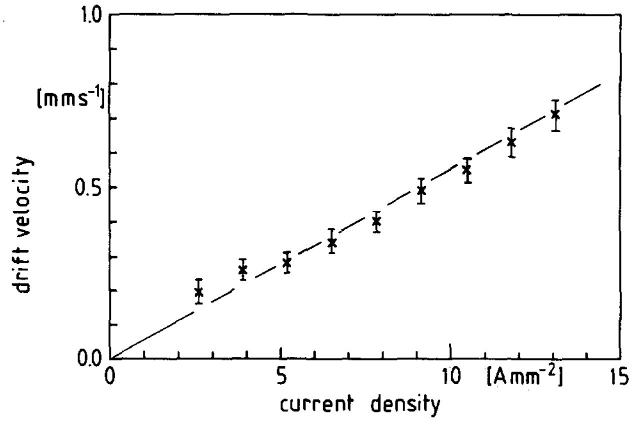


Fig. 2. Drift velocity versus current density in copper.

minals is zero. While turning the magnetic field on, the amplifier must be protected from damage by the large induced emf by closing switch *S*.

With the gear motor moving the Hall specimen in the correct direction, the voltage will decrease and reach zero at about 0.7 mm/s, the drift velocity in copper.

The drift velocity depends upon the strength of the electric field inside the metal. Figure 2 shows the change in drift velocity from 0.3 to 0.6 mm/s^{-1} with a change in current from 5 to 10 A while the carrier mobility remains constant. With a current density of 10 A/mm^{-2} in Sb a drift velocity of -0.5 mm/s^{-1} could be observed. Here the drift motion is opposite to that in copper due to the fact that in Sb the majority carriers are "holes."

DISCUSSION OF RESULTS

The experiment can be explained from two different points of view leading to the same quantitative description according to the principle of relativity in electromagnetic induction.

Let us consider the case where the Hall voltage is exactly compensated, e.g., zero.

In a frame of reference attached to the Hall specimen the magnetic field flux through the circuit formed by specimen, Hall contacts and leads to the voltmeter changes. This gives rise to an induced voltage and hence a Coulomb force which is equal and exactly opposite to the Lorentz force acting on the drifting electrons. The induced voltage thus compensates the Hall voltage.

In a frame of reference attached to the magnet, no Hall voltage will be generated as the electrons are on average at rest with respect to the magnetic field and the mean Lorentz force is zero.

ACKNOWLEDGMENTS

I am greatly indepted to Professor H. Happ who while reading through old notes on demonstration experiments came across this convincing demonstration, to Professor G. Dietz and Dr. T. Pauls for helpful discussions and for reading the manuscript, to Mrs. T. Becker for the reproductions and to Mrs. R. Küpper for the drawings and her help during the measurements.

---

[1] J. Jaumann died on 15 June 1971 at the age of 69.

The early history of cosmic ray research

Qiaozhen Xu[a] and Laurie M. Brown
Department of Physics and Astronomy, Northwestern University, Evanston, Illinois 60201

(Received 22 May 1985; accepted for publication 11 February 1986)

We review the prehistory and early history of cosmic ray studies, concentrating on the period 1900–1927. Following the discoveries of the electron and radioactivity just before the turn of the century, the old problem of leakage of charge from a conductor in air was investigated in terms of the new concept of ionization and ionizing radiation, part of which was found to be highly penetrating and to be of extraterrestrial origin. At first supposed to consist only of ultrahigh energy gamma rays, the cosmic ray primaries are now known to be mainly charged particles. The modern period of cosmic ray research began in 1927, when individual particles were studied by cloud-chamber and coincidence counting techniques.

I. INTRODUCTION

The discovery and exploration of the cosmic rays are heroic chapters in the history of modern physics. The exploration sometimes required adventurous journeys and arduous exertions, involving descents into caves and mines, sea voyages, mountain ascents, balloon flights, etc. The results were no less heroic, yielding profound geophysical and astrophysical insights and initiating such fields as elementary particle physics and high-energy nuclear physics. Begun as a study of the distribution in land, sea, and atmosphere of the sources of the mysterious invisible radiations that seemed inescapable, cosmic ray research gradually revealed the wealth of particles that populate the tiniest portions of space as well as providing a window on the most remote regions of space and time.

The cosmic rays are responsible for a fraction of the small electrical conductivity of the atmosphere observed during fine weather, even at low altitudes. The conductivity of the air, as measured by the rate of discharge of an electroscope, was studied as a part of atmospheric physics (meteorology) from about 1900. At that time, only a few years after the discovery of radioactivity, it was believed that the air's conductivity was due to ions produced by radioactivity in the Earth's crust, and that the "residual ionization" found later at heights above the ground (e.g., on the Eiffel Tower), or above glaciers or on the sea, was due to radioactive emanations mixed with the air. However, balloon flights carried out between 1909 and 1914 showed a large systematic increase in ionization with altitude (on the scale of kilometers), which strongly suggested that an ionizing radiation came down from above and was gradually absorbed by the atmosphere. Beginning about 1926, this radiation was considered well established (even by the most skeptical scientists) and was called by the name "cosmic rays." The present article deals with the history of the cosmic rays up to that date.

II. THE DISCOVERY OF IONS IN THE ATMOSPHERE

In 1785, Charles Coulomb showed that a charged metallic conducting body, placed in the air, gradually loses its charge. That was probably the first reported observation of electrical conduction in the atmosphere in fine weather. After the passage of more than a century, the cause of the air's conductivity was still regarded as an unsolved mystery. We should emphasize that we are concerned with extremely delicate effects, involving only a few ion pairs in a macroscopic volume of air, and that the techniques of observation (which will not be discussed in detail) were being developed at the same time as the observations themselves were being made. Therefore, it is not surprising that there was a considerable measure of disagreement in the pioneering investigations.

Joseph John Thomson, the discoverer of the electron, began to investigate the electrical conductivity of gases at Cambridge University in the early 1880s, applying high voltage from a spark coil to a gas discharge tube. Between 1886 and 1896 he published about ten articles in the *Proceedings of the Royal Society of London*. In his papers, he introduced new terminology: a "polarized molecule" splits into a "positive atom" and a "negative atom."[1] In an 1895 paper entitled "On the Electrolysis of Gases," he began to speak of "ions" in the gas.[2] Beginning in 1896, thus soon after the discovery of x rays, Thomson was using the new radiation to create ions in the air in his apparatus. In 1906, Thomson received the Nobel Prize in Physics "in recognition of the great merits of his theoretical and experimental investigations on the conduction of electricity by gases." That was how the citation described his pathbreaking discovery of the electron![3]

Between 1896 and 1899, Thomson's student, Charles Thomson Rees Wilson published four articles on the effects of x rays on the nucleation of small "clouds" in the laboratory. In his paper "On the Condensation Nuclei Produced in Gases by the Action of Roentgen Rays, Uranium Rays, Ultra-violet Light, and other Agents," Wilson emphasized the difference "between 'ions' and nuclei which carry no charge of electricity."[4] One of his findings was that negative ions have a greater efficiency than positives in nucleating cloud formation.[5] During those years, he studied meteorological effects, and to continue those studies he invented the "cloud chamber," which was later used to observe the tracks of fast charged particles, including those of the cosmic rays.

The German scientists Julius Elster and Hans Geitel formed an interesting research team. Having been friends already in high school, they both became teachers at the Gymnasium in Wolfenbüttel. According to Abraham Pais, "When Elster married and had a house built, Geitel moved in with the young couple and together the two friends built a laboratory in the new home. Here they started their research (often financed from their own pockets) which were to make them internationally renowned. They experimented on photoelectric effects, on spectroscopy, on the conduction of electricity through gases, and especially on atmospheric electricity. These last experiments led to their classic work on the radioactivity of the atmosphere."[6] Pais's article describes their efforts to find an external source for the energy of radioactive substances.

In Thomson's book of 1906 on gas conduction, he wrote:

> In May of 1900, Elster and Geitel...noticed that an electrified body gradually lost its charge...They found that the rate of leak varied,...that it was very much smaller in mist or fog than when the weather was bright and clear, that it was greater at high altitudes than at low ones, and that on the tops of mountains the rate of escape of negative electricity was much greater than that of positive. In plains, they found the rate of leak to be the same for plus and minus charges. They concluded that free ions existed in the atmosphere.[7]

The results of Elster and Geitel were in agreement with some earlier experiments of F. Linss.[8] They did not speculate on a possible origin of the atmospheric ions.

Six months after Elster and Geitel, Wilson reported on experiments done "in a small closed vessel containing dust-free air not exposed to any known ionizing agents," and concluded that the air was a conductor of electricity, the rate of leakage of charge being independent of the sign and proportional to the pressure. A saturation current was produced when the potential was either 120 or 210 V. Its value was taken to measure the ionization, and in this way he estimated that at atmospheric pressure 10 ion pairs were produced per second per cc of air (later revised by him to 14).[9]

In 1903, Canadian physicists John Cunningham McLennan and Eli Franklin Burton tried to relate the conductivity of the air to the thorium emanation that Ernest Rutherford had recently observed. They argued that air is continually ionized, as shown by its weak electrical conductivity, and so "one is forced to conclude there is present in the air an emanation possessing properties similar to that emitted by the thorium compounds."[10]

Wilson also speculated that "some radioactive substance in the atmosphere ...is carried down in the rain."[11] These physicists, then, discovered and named the "ions" in the atmosphere, introduced the term "ionization," and tried to relate it to the radioactivity of the Earth.[12]

III. EXPERIMENTS ON THE CONDUCTIVITY OF THE ATMOSPHERE

After the discovery of ions in the atmosphere, physicists carried out numerous experiments on aspects of this phenomenon, some out of mere curiosity without any apparent aim, but most of them trying to explore the unknown origin of the ions.

A. Experiments with vessels of various materials

In 1902, McLennan and Burton made a series of observations on atmospheric air confined in air-tight vessels of different metals, concluding that, "the effects observed would seem to indicate that all metals in varying degree are the sources of a marked though feeble radioactive emanation."[13] Robert John Strutt also reported, "...there are very marked differences in the rate of the leak, when different materials constitute the walls of the vessel." As a result, he said: "There can therefore, be little doubt that the greater part–if not the whole–of the observed ionization of air is not spontaneous at all, but due to Becquerel rays from the vessel."[14]

B. Experiments with different temperatures and pressures

J. Patterson wrote in 1902 that at constant atmospheric pressure, the conductivity of the air was constant over temperatures ranging from room temperature to 500 °C, and for pressures down to 1/3 atmosphere; for lower pressures it was proportional to the pressure.[15] Wilson found a similar pressure dependence, and wrote that "The falling off from this law at the higher pressures might be taken as indicating that the ionization is due to radiation from the walls of only moderate penetrating power."[16] G. Jaffe, however, found not a simple proportionality of conductivity with pressure, but the linear relation $y = ap + b$, arguing, "This fact seems to indicate that the ionization is (at least partly) due to a radiation from the walls which is not of uniform-type. On this supposition the term b would correspond to a very weak radiation, which is perfectly absorbed by as little as 3 cm of air at 1/5 of an atmosphere pressure."[17] These authors therefore attributed the air's conductivity to local sources.

C. Experiments concerning periodicity

From 1903 to 1904, George C. Simpson made a very complete series of observations on the diurnal and annual variations of the potential gradient, ionization, and dissipation within the Arctic Circle, at Karasjok in Norway at latitude 69° North, on which he wrote a summary in 1905. In his report entitled "Atmospheric electricity in high latitudes," he wrote, "The daily period of the ionization is not so pronounced as that of the dissipation, but the ionization is slightly lower in the evening than in the morning or at midday during the whole year...temperature has a great effect on the ionization while no effect of temperature on the ratio $[I_+/I_-]$ is apparent." And he also pointed out, "The value of the ratio I_+/I_- shows a very distinct yearly period with a maximum in the winter and a minimum during the summer." About the ionization, he wrote "during the summer we have six months' fall from August to February followed by a similar six months' linear rise from Feb. to Aug..."[18]

In 1906 and 1907 Alexander Wood, working with Norman R. Campbell also claimed to have detected a diurnal periodicity in the ionization of gases in closed vessels, having two maxima and two minima per day,[19] a result confirmed by T. Frederick McKeon.[20]

D. Experiments with shielding

Some of the earliest and most significant experiments on shielding were done by Rutherford and H. Lester Cooke, who motivated their experiments with these comments: "Since the excited activity obtained from the atmosphere is very similar in character to the excited radiations from thorium and radium, it was thought possible that some penetrating rays might be given off from the surface of the earth and walls and rooms on which excited activity from the air is distributed. In order to test this point, the amount of ionization was observed in testing vessels of about 1 liter capacity.... The effect of placing metal screens outside the testing vessel was observed." They found little effect on the rate of discharge from a 2 mm thickness of lead placed around their apparatus, but 5 cm of lead cut down the discharge rate by 30%. Beyond 5 cm of lead, they found no effect, although 5 tons of pig lead was placed around the apparatus. On removing the screens, the discharge returned to the original value. They concluded that their results showed that "about 30% of the ionization inside a closed vessel is due to an external radiation of great penetrating power." They claimed that "these effects could not be due to the presence of thorium or radium in the laboratory, for similar results were observed in the library which was free from all possible contamination by radioactive substances."[21]

In order to determine whether the radioactive emanation was given off by the walls of the containing vessel, McLennan and Burton also did a screening experiment. They reported, "The heavy cylinder was filled with air to a pressure of about 400 cm Hg, and allowed to stand until its conductivity had become steady. It was then placed in an insulated galvanized iron tank which was gradually filled with water so as to surround the cylinder with a layer 25 cm in thickness. The initial conductivity before the water was admitted was 21.1. As the water rose, the conductivity decreased and fell to 13.3, when the tank had been filled. The values for the conductivity...show that the loss was almost directly proportional to the rise of the water. The total fall in conductivity was about 37%." "From these results," they arrived at the conclusion that "it is evident that the ordinary air of a room is traversed by an exceedingly penetrating radiation."[22] They made that important observation as early as 1902.

E. Experiments at different localities

Aside from experiments carried out in laboratories, physicists made experiments at different kinds of localities such as in caves, on the sea, on lakes, etc. In 1903, Elster and Geitel, experimenting in caves, observed the interesting phenomenon that the rate of leak in caves and cellars, where the air was stagnant and only renewed slowly, was much greater than in the open air. They found that in a cave, the electric charge leaked off at seven times the rate it did in the outside air, even when it was clear and free of mist. They also found that in a cellar whose windows had been shut for eight days, the rate of leak was considerably greater than it was in the outside air.[23]

Arthur S. Eve had speculated in 1907 that the ionization over the ocean should be less than that over land, for "experimental evidence...indicates that radium is present in seawater to a markedly less degree than in the sedimentary rocks on land. And, since radium emanation decays to half value in four days, the wind is unable to transport the emanation from land to places in mid-ocean before the activity is decreased." Eve did not obtain the result he expected, so he suggested that "The ionization observed is larger than would be anticipated from such a cause (wind carrying radium emanation from land to sea), but it is possible that the rate of recombination of ions over the sea may be less than over the land."[24]

C. S. Wright considered lake water to be a very efficient shield for the Earth's radiation, stating, "that the water of Lake Ontario acts as a perfect screen both for the earth's radiation and, if a sufficient depth be taken, for the gamma rays from radium. On this account and owing to the fact that the water of Lake Ontario contains no active impurity, it has been possible to determine what portion of the ionization in the receivers used in this investigation was due to residual active impurities and to intrinsic activity in the metals of the receivers." By this method, he determined that the ionization due to radioactive impurities in clay soil was about 1 ion per cc per s.[25]

F. Conclusions

During the first years of the present century, physicists primarily working on radioactivity, as well as those engaged in atmospheric science, became aware that the electrical conductivity of the air in fine weather (or alternatively, of gases in closed vessels) was an indicator of the presence of invisible high-energy radiation and they did numerous experiments to study its properties and determine its sources. In summarizing those experiments, Rutherford concluded in 1905, "It is now certain that a large part of the ionization observed in a clean metal vessel results from the emission of ionizing radiations from its walls. A part is due to a very penetrating radiation of the gamma ray type which is everywhere present on the surface of the earth." He noted further that "In most cases the ionization falls off nearly proportionally with the pressure, and is approximately proportional to the density of the gas. Both of these results are to be expected if the ionization observed is due to radiations from the walls or to a penetrating type of radiation passing from the outside through the material of the vessel."[26]

In the same year, Eve wrote that "the natural ionization of the air at the surface of the earth" has obvious causes; namely, "the only ionizing agents under such conditions are (1) radiation due to radioactive matter contained in the air, (2) radiations due to active matter on the surface, or in the material of the sides of the vessel, (3) penetrating radiation through the sides of the vessel, due to radioactive matter in the surrounding bodies."[27]

IV. TENTATIVE SUGGESTIONS OF AN EXTRATERRESTRIAL SOURCE

Robert Andrews Millikan, in a popular book of 1935 surveying particle and cosmic ray physics, stated that "Apart from a passing suggestion of Richardson in 1906,.... I can find no record of the existence anywhere up to 1910 of any ideas even remotely related to those that are now associated with the term 'cosmic rays'."[28] Millikan's statement is not really accurate, for an extraterrestrial source was considered by several scientists, in the early years of the century, as the cause of atmospheric ionization and the Earth's negative electrical potential.

In 1901 Wilson tried "to test whether the continuous production of ions in dust-free air could be explained as being due to radiation from sources outside our atmosphere, possibly radiation like Röntgen rays or like cathode rays, but of enormously greater penetrating power." He carried out experiments in tunnels, but finding no evidence of any decrease in the rate of ionization, no matter how much solid rock was overhead, he concluded: "It is unlikely, therefore, that the ionization is due to radiation which has traversed our atmosphere; it seems to be, as Geitel concludes, a property of the air itself."[29] In 1902, he boiled down freshly collected rainwater and found the dry residue to be radioactive. He also found that "The radioactivity obtained by the evaporation of rain disappears in the course of a few hours, falling to half its initial value in one hour."[30] The same results were obtained from the evaporation of freshly fallen snow.[31] These experiments did not, however, suggest the origin of the radioactivity, and the possibility of an extraterrestrial source remained in his mind.

In 1903, in discussing experiments of Philipp von Lenard, Wilson suggested that sunlight ionizes the air, "especially in the upper atmosphere, while it is still strong in ultraviolet rays."[32] And considering ionized layers of the atmosphere nearer the ground, he wrote, "It is quite conceivable that we may be driven to seek an extraterrestrial source for the negative charge of the earth's surface."[33]

The connection between the source of the air's conductivity and the Earth's electric charge was also considered by Simpson, who stated in 1904: "If we take for granted that the sun continually emits Becquerel rays consisting of positive and negative electrons, one would expect the following to be the consequence. Some of the electrons which reach the earth's atmosphere will be absorbed–probably mainly by the water vapour and dust in the lower atmosphere–but according to Rutherford's experiments more positive than negative; thus we may expect a greater number of negative electrons to reach the surface, a corresponding number of positive electrons being held back by the air. We at once see a cause for the positive charge of the air and the corresponding negative charge on the surface."[34] (The reference here to "positive electrons" is not prescience; it merely refers to the positive unit charge.)

Pursuing the same issue in 1906, Wilson said: "If the existence of a penetrating radiation from cosmical sources were established it would be of the greatest importance in connection with atmospheric electricity. For it would open the question as to whether the negative charge of the earth might not be supplied by these rays. At present I think it is much more likely that precipitation will prove to be a sufficient source."[35] As Millikan acknowledged, Owen W. Richardson also considered the possibility of an external radiation source. He said it could account for a reported diurnal variation in both the Earth's electric field and the conductivity of enclosed air, and he speculated, "In the case of the earth the ionizing rays presumably come from extraterrestrial sources, and will be absorbed to some extent by the earth's atmosphere. They will therefore be more intense further away from the earth's surface,..."[36]

In 1909, a review by Karl Kurz also ruled out radioactive matter in the Earth's crust as an effective ionizing agent at high altitudes because of the absorption of the air, and noting that the conductivity of the air at high altitude is of the same order of magnitude as at the Earth's surface, asked, "What ionizing agent compensates the lack of penetrating radiation from below? Is there perhaps, after all, an ionizing radiation from outside the atmosphere? It could be of such strength and absorbability that it would be without effect in the lower atmosphere... ."[37]

He mentioned preparations at Munich for studies on balloon flights during the forthcoming International Balloon Week in December 1909.

V. FIRST OBSERVATIONS ON IONS IN THE ATMOSPHERE AT DIFFERENT ALTITUDES

Typical experiments on atmospheric ionization studied the rate of discharge of a charged object at different heights above the ground. If the penetrating radiation present at the surface of the Earth were entirely of terrestrial origin, one should detect a diminution in the intensity of this radiation even at moderate distances above the Earth's surface. The altitude dependence was studied, more or less simultaneously, by carrying apparatus to towers and mountain tops or aloft in balloons.

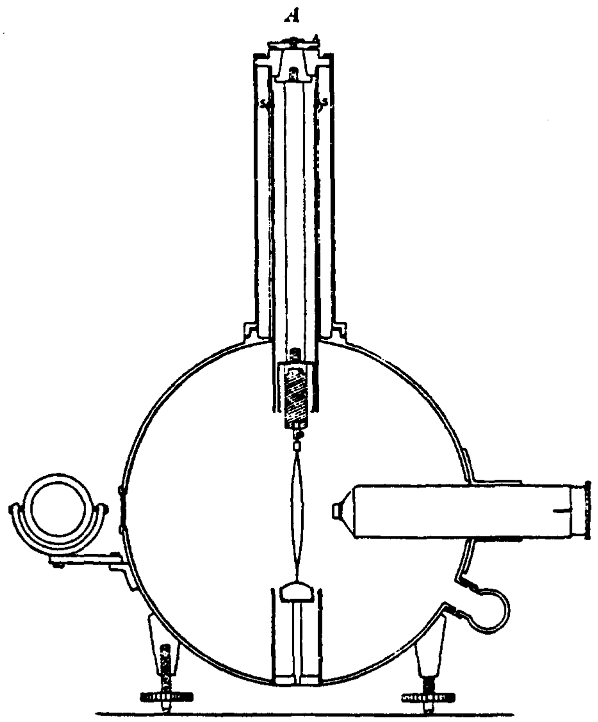


Fig. 1. Electroscope developed by Wulf in 1909.

A. Observations on mountains and towers

In May of 1900, Elster and Geitel found that the rate of leak of electrical charge was greater at high altitudes than at low ones. They compared the effect at sea level and up to 3000 m on the tops of mountains; they reported their observations, but drew no conclusions.[38]

In 1909, Theodor Wulf, of Ignatius College in Valkenburg, Holland, greatly improved the electroscope, replacing the gold leaves with two slender metal wires held, under tension by a light quartz fiber. When charged, the two wires repelled each other and the separation was measured by means of a microscope.[39] This kind of sensitive electroscope was used in later experiments, including those of Hess and Millikan. (See Fig. 1.)

Karl Bergwitz had reported a large decrease of ionization with height, but observations by Albert Gockel contradicted this (see below).[40] On the basis of these balloon results, Wulf decided to make measurements at the top of a tower, since he was skeptical about getting accurate readings in a balloon. The balloon was moving, changed location, and could be influenced by clouds and rain.

In March and April of 1910, Wulf made four measurements with his new sensitive electroscope on the Eiffel Tower in Paris. After substracting the chamber background, he found that six ions were produced per cc per second by the radiation at the surface of the Earth, while at 300 m on top of the tower the number was 3.5. The intensity of ionization was thus reduced to 60% at the top, not what he had expected. According to Wulf's calculation, at 80 m the gamma rays should already reduce to half their value on the ground, and at 300 m it should be only a few percent of the value on the ground. Therefore, Wulf concluded, "either another source of gamma-rays exists in the upper layers of the atmosphere, or the absorption coefficent of gamma-rays in air is apparently smaller than has been assumed."[41]

In 1911, McLennan and E. N. Macallum of Toronto tried to check values calculated by Eve on the effect of altitude in the intensity of the gamma rays from RaC present in the Earth,[42] measuring the ionization on the ground and on buildings at different heights. They found that the ionization at the top of a tower "was only about 48% of the effect produced by similar radiations on the university lawn," in agreement with Eve's calculations.[43]

B. Balloon flight observations

Hermann Ebert made three balloon flights in June and November of 1900, and January 1901 and reached a maximum height of 3770 m. He tried to repeat the Elster–Geitel experiments in "free air," because in mountain-top observations there was always a distortion of the electrical potential levels of the Earth; the effect might also be seasonal. On his first trip up to 2920 m, a negatively charged body discharged faster than a positive one, showing that positive ions were in excess in that layer of the atmosphere. At higher altitude, the discharge rates of both signs of charges became equal. On the second ascent, there was definite evidence of a variation of ionization with height.

From these observations, Ebert concluded: The discharge rate of charged bodies grows appreciably with height, for both positive and negative charges; the unipolarity observed at sea level extends at most to 2000–3000 m; that is, the excess of the leak of charge from a negatively over that from a positively electrified body reaches a maximum at a height between 2000 and 3000 m; at the highest altitudes there is but little difference between the two leaks. However, the leak depends to some extent upon meteorological factors.[44] Those were the first balloon observations on the discharge rate that we have been able to find.

In 1909, Bergwitz made balloon ascents, in which he found a marked decrease of the total ionization, which at 1300 m was only about 25% of the value on the ground.

That was considered to be in better accord with expectation, if the ground were the source of a gamma radiation responsible for the residual ionization. However, there was some question about the validity of Bergwitz's measurements, because his electrometer was damaged during the flight by deformation of its pressure vessel.[45]

Between 1909 and 1911, the Swiss physicist Gockel undertook several balloon ascents, his first flight being carried out to mark a special occasion. *The Society of German Natural Scientists and Physicians*, which met each year in a different German-speaking city, convened in Salzburg, Austria in September 1909. Such famous physicists as Planck, Wien, Sommerfeld, Elster, Stark, Born, Laue, Hahn, Meitner, and Einstein were in attendance.[46] It was at this meeting that Kurz presented his review on the penetrating radiation. He and Wulf had made observations in Zermatt and its surroundings which persuaded them that the source of the radiation was the ground, but they were planning balloon flights to study the variation of the radiation with altitude. In Braunschweig, Germany, balloon flights were also being prepared, and an *International Balloon Week* was to take place.[47]

Gockel played an important part in this activity. As he reported, "For the *International Balloon Week* in December 1909, the Swiss Aeroclub had the kindness to place the Balloon 'Gotthard' at the disposal of Dr. de Quervain and the author. Unfavorable weather had the effect that the ascent had to be put off to the last day of the week, the 11th of the month." The balloon carried three persons (de Quervain for meteorology, Lieutenant Muller to guide the balloon, Gockel to make electric measurements) to a height of 4500 m. Gockel's conclusion was: "The result of the measurement is that in the free atmosphere, there is in fact a lessening of the penetrating radiation, but by far not to the extent that one could expect if the radiation arises mainly from the ground."[48] He made several additional balloon ascents in which he found a slight decrease in gamma radiation with height, but later he realized that his experimental procedure was not free of objection. Gockel's last flight, made on 2 April 1911, was an "Aeroclub Excursion" and carried five persons. Comparing the ionization rate at 2500 m with an earlier result he had obtained at 2800 m, and correcting for barometric pressure, he found a weak *increase* in ionization with increasing height.[49]

VI. DECISIVE EXPERIMENTS BY HESS AND BY KOLHORSTER

From Wulf's Eiffel Tower experiments and Bergwitz's and Gockel's balloon observations, Victor Franz Hess, an Austrian physicist, decided that the accumulated evidence suggested the presence of a previously unknown source of ionization and he initiated an experimental program to check this possibility. He first measured outdoors the absorption in air of gamma rays from an intense radium source, varying the distance between a closed ionization chamber and the source from 10 to 90 m, establishing that gamma rays from the Earth should be almost completely absorbed at a height of 500 m.

Hess was an active amateur balloonist, and he planned a series of manned-balloon ascents. He was aware that during Gockel's ascents, the pressure of the gas in his instrument varied with altitude, which invalidated the measurement. To avoid this problem, he designed an instrument that could survive the rigors of an open balloon gondola–an air-tight ionization chamber with walls sufficiently thick to withstand a pressure differential of one atmosphere, containing a temperature-compensated Wulf fiber electroscope.

A. Flights in 1911 and 1912

In 1911, on the first of a series of ten balloon flights, Hess reached 1070 m. He found that the ionization varied very little and concluded , "Since the radiation of that height was not remarkably different from that at sea level, there must be another source of the penetrating radiation in addition to the gamma-radiation from the radioactive substances in the earth's crust."[50] In the same year he made two other balloon flights, and in 1912 seven more, among which two were of special importance. The first flight in 1912 took place on the occasion of a considerable partial eclipse of the sun in lower Austria on 17 April, from 11 am to 1 pm, up to an altitude of 2750 m. During the eclipse, the balloon descended as a result of the cooling of the gas. Hess found that the radiation at around 2000 m was greater than at ground level, and that the eclipse had no effect on the penetrating radiation. Therefore, he came to the conclusion that "if a part of the radiation is of cosmic origin, it can hardly come from the sun, at least so long as one thinks of a gamma-radiation propagated in straight lines."[51]

B. Decisive flight

The seventh flight of Hess, in a hydrogen-filled balloon was intended to reach a very high altitude. At 6:12 on the morning of 7 August 1912, the balloon ascended from a field near the town of Aussig, in Austria. In its gondola were three persons: a navigator, a meteorologist, and Hess. The flight lasted six hours, reaching the height of 5350 m. At noon, the balloon touched down near the German town

Fig. 2. Photograph of Professor V. F. Hess taken after an important balloon flight in 1912.

Table I. Data taken during Hess's 7th flight in 1912.

Balloon "Böhmen" (1680 cbm hydrogen)
Meteorological observer: E. Wolf.

Leader: Captain W. Hoffory.
Electr. observer: V. F. Hess

		Mean height		Observed radiation					
				Inst. 1	Inst. 2	Inst. 3			
No.	Time	abs. m	rel. m	q_1	q_2	q_3	red. q_3	Temp.	Rel. humidity %
1	$15^{\text{h}}15$–$16^{\text{h}}15$	156	0	17.3	12.9	···	···	1½ days before the ascent (in Vienna)	
2	$16^{\text{h}}15$–$17^{\text{h}}15$	156	0	15.9	11.0	18.4	18.4		
3	$17^{\text{h}}15$–$18^{\text{h}}15$	156	0	15.8	11.2	17.5	17.5		
4	$6^{\text{h}}45$– $7^{\text{h}}45$	1700	1400	15.8	14.4	21.1	25.3	+ 6.4°	60
5	$7^{\text{h}}45$– $8^{\text{h}}45$	2750	2500	17.3	12.3	22.5	31.2	+ 1.4°	41
6	$8^{\text{h}}45$– $9^{\text{h}}45$	3850	3600	19.8	16.5	21.8	35.2	− 6.8°	64
7	$9^{\text{h}}45$–$10^{\text{h}}45$	4800	4700	40.7	31.8	(ended by accident)		− 9.8°	40
		(4400–5350)							
8	$10^{\text{h}}45$–$11^{\text{h}}15$	4400	4200	28.1	22.7			···	···
9	$11^{\text{h}}15$–$11^{\text{h}}45$	1300	1200	(9.7)	11.5			···	···
10	$11^{\text{h}}45$–$12^{\text{h}}10$	250	150	11.9	10.7			+ 16.0°	68
11	$12^{\text{h}}25$–$13^{\text{h}}12$	140	0	15.0	11.6	(After landing at Pieskow, Brandenburg)			

of Pieskow, 50 km east of Berlin. Figure 2 was taken when Hess and his collaborators touched down.

The results of Hess are shown in Table I. Instruments 1 and 2 are thick walled, while instrument 3, with thin walls, was sensitive to both beta and gamma rays. At 1500 to 2500 m, the radiation was about as strong as it was on the ground. There then began a clearly perceptible rise in the radiation with increasing height.

Combining these data with those of his other balloon flights, Hess arrived at these important conclusions: "The results of the present observations seem to be most readily explained by the assumption that a radiation of very high penetrating power enters our atmosphere from above, and still produces in the lowest layers a part of the ionization observed in closed vessels. The intensity of this radiation appears to be subject to transient variations, recognizable in hourly readings. Since I found a reduction in the radiation at the balloon neither by night nor at a solar eclipse, one can hardly consider the Sun as the origin of this hypothetical radiation, at least so long as one thinks only of a direct gamma radiation with rectilinear propagation."[52]

Hess's achievement was recognized in 1936, when Professor H. Pleijel of the Royal Swedish Academy of Sciences addressed these words to Hess, when the latter shared the Nobel Prize for physics with Carl David Anderson:

> By virtue of your purposeful researches into the effects of radioactive radiation carried out with exceptional experimental skill you discovered the surprising presence of radiation coming from the depths of space, i.e., cosmic radiation. As you have proved, this new radiation possesses a penetrating power and an intensity of previously unknown magnitude; it has become a powerful tool of research in physics, and has already given us important new results with respect to matter and its composition. The presence of this cosmic radiation has offered us new, important problems on the formation and destruction of matter, problems which open up new fields of research."[53]

C. Confirmative flight

Werner Kolhörster, a German physicist, made five dangerously high balloon flights with more refined techniques, in 1913 and 1914. He checked the effect of low temperature

Table II. Kolhörster measurements in 1914.

Kolhörster

Flight 1		Flight 2		Flight 3		Flight 4	
Alt.	Ions	Alt.	Ions	Alt.	Ions	Alt.	Ions
310	− 1.2	500	− 2.0	1090	− 1.2	1000	− 1.5
760	− 1.3	600	− 1.4	2130	+ 2.1	2000	+ 1.2
1650	+ 0.8	1000	− 2.1	3550	+ 7.0	3000	+ 4.3
2110	+ 1.3	1400	− 1.7	4700	+ 14.5	4000	+ 9.3
2400	+ 3.1	1500	− 0.8	5600	+ 27.5	5000	+ 17.2
2600	+ 4.3	2400	+ 3.1	6200	+ 29.3	6000	+ 28.7
3000	+ 7.5	3300	+ 4.5			7000	+ 44.2
3400	+ 8.9	4000	+ 6.7			8000	+ 61.3
3500	+ 11.1					9000	+ 80.4

on the Wulf electroscopes before his balloon ascents, something to which Hess had not paid much attention. Kolhörster reached a maximum altitude of 9300 m, enabling him to greatly extend Hess's observations.

Table II shows the average difference between the ionization observed at various heights and that at sea level, in ions per cc per second.

These provided clear confirmation of Hess and should have made his conclusions incontrovertible. The ionization increased in Kolhörster's observations until it attained a value about 50 times that at sea level. The radiation from space thus had an attenuation coefficient of 1 × 10–5/cm, much more penetrating than any known gamma rays. Extrapolating Kolhörster's measurements back to ground level suggested an ionization of 2.5 ions per cc per second.[54]

Although radiation of cosmic origin seemed to be well established at this point, it still had no commonly accepted name. Egon Von Schweidler introduced the term "Hess rays" for it, but Hess himself used the term "ultragamma radiation" and Kolhörster called it "Höhenstrahlung."[55]

Probably because of the First World War and its aftermath, there were no reports of significance for this history from 1914 to 1922. Gockel, Hess, and Martin Kofler continued their researches by mountain ascents and balloon flights.[56] While in the army, Kolhörster continued cosmic ray observations at his meteorological observation station.[57] However, no big progress was made.

VII. FURTHER EXPERIMENTS AND MILLIKAN'S DOUBTS

The conclusion of Hess and Kolhörster that there was an extraterrestrial source for the penetrating radiation was accepted gradually by most physicists, but Millikan claimed that it was not convincing to himself nor to some other specialists, including William Francis Gray Swann, Frederick A. Lindemann, and G. Hoffmann.[58] Millikan was the most skeptical of them. When he moved to the California Institute of Technology after the First World War, he set out energetically to determine whether the experimental data reported by Hess and other workers were correct, and whether there were compelling reasons to believe in the existence of a radiation from outer space. With collaborators, he carried out three series of experiments.

A. First series; ascent into the stratosphere

In March and April of 1922, from Kelly Field, San Antonio, Texas, Millikan and Ira Sprague Bowen sent four unmanned balloons into the stratosphere, carrying self-registering light-weight electroscopes and thermometers. With all its recording and driving mechanisms, each apparatus weighed only about 200 g.

Three of the four balloons were recovered and two of these had made satisfactory records of their flights, during which they reached altitudes of 11.2 and 15.5 km, respectively. A comparison of the electroscope reading recorded at the 5-km level during ascent with the reading at the same level during descent showed that the average discharge rate of the electroscope while above the 5-km level was about three times its discharge rate at the surface of the Earth. Millikan found that that was only 25% of the value expected from the Hess–Kolhörster curve. Thus, Millikan and Bowen concluded, "The results then of the whole Kelly Field work constitute definite proof that there exists no radiation of cosmic origin having such characteristics as we had assumed." Their paper, published in 1926, continued, "They show that the ionization increased much less rapidly with altitude than would be the case if it were due to rays from outside the earth having an absorption coefficient of .57 per meter of water."[59]

B. Millikan's mountain peak and airplane observations

In the summers of 1922–23, Millikan and Russel M. Otis carried out his second series of experiments. At Ross Field near Pasadena, Otis made measurements in captive balloons. These yielded results in agreement with those of other observers, in that up to an altitude of 2000 m the number of ions per cc per second was 1 to 3 less than that on the ground. Again, Otis sent his equipment on several airplane flights in 1922 at Marsh Field near Riverside and in 1923 at Rockwell Field near San Diego. These flights reached heights of more than 5000 m. The results obtained were in agreement with those we have quoted of Millikan and Bowen, in that they showed a markedly lower rate of leak at the highest altitudes than those reported by Hess and Kolhörster.

During the summers of 1922 and 1923, Millikan's team made a long series of observations on Mt. Whitney (4130 m) and Pike's Peak (4300 m). They found variations with altitude, but no dependence of penetrating radiation upon daylight or darkness, or upon the position of any of the heavenly bodies. In September 1923, on Pike's Peak, they made experiments outdoors and indoors, with all sides of the vessel shielded with lead, or with one or two sides unshielded and found that the penetrating radiation came equally from all sides.[60] Because of the screening effect, Millikan concluded that "there exists no such penetrating radiation as we have assumed (of cosmic origin)." Continuing, they "found as a result of a snowstorm on the mountain as large a percentage change (about 10%) in the ionization inside our 5-cm lead shield as outside it. We interpret this result also as meaning that the whole of the penetrating radiation is of local origin. How such quantities of radioactive material get into the upper air is as yet unknown."[61] The experiments of series (a) and (b) confirmed Millikan's doubts about the existence of radiation of cosmic origin.

C. Measurements in snow-fed lakes: removing Millikan's doubts

In September 1925, Millikan and G. Harvey Cameron carried out a series of experiments in snow-fed lakes at high altitudes. They went first to Muir Lake at 3590 m above sea level, just under the brow of Mount Whitney, the highest peak in the United States (except for Alaska). The lake is very deep and some 700 m in diameter. They worked there for the last ten days of September, sinking two electroscopes to various depths down to about 20 m. The electroscope readings decreased steadily down to a depth of 15 m below the surface. The atmosphere above the lake was equivalent in absorbing power to 7 m of water, so that they were observing rays so penetrating that if they came from outside the atmosphere, they had the power of passing through 22 m of water before being completely absorbed.

In order to obtain definite evidence as to whether these very hard rays were of cosmic origin, coming wholly from above, the atmosphere acting merely as an absorbing air

Table III. Snow-fed lakes measurements by Millikan.

Readings in Lakes Muir and Arrowhead Electroscope No. 3 Muir Lake									
Depth below surface (m)	0	0.45	1.0	2.8	3.0	5.0	10.0	15.0	20.0
Ionization (ions/cc/s)	13.3 13.2	9.7 ···	7.7 7.8	6.0 5.8	5.45 ···	4.9 4.6	4.0 4.0	3.6 ···	3.6 3.7
Means	13.25	9.7	7.75	5.9	5.45	4.75	4.0	3.6	3.65
				↙	↙	↙			
Arrowhead									
Depth below surface (m)	0	0.7	1.0	1.1	3.0	5.0	···	15.0	···
Ionization (ions/cc/s)	7.0 7.2 7.5 6.9 7.2	5.8 ··· ··· ··· ···	5.5 ··· ··· ··· ···	5.15 ··· ··· ··· ···	4.85 4.9 ··· ··· ···	4.4 ··· ··· ··· ···	··· ··· ··· ··· ···	3.7 ··· ··· ··· ···	··· ··· ··· ··· ···
		↗	↗		↗				
Means	7.0	5.8	5.5	5.15	4.9	4.4	···	3.7	···

blanket, they next went to another deep snow-fed lake, Lake Arrowhead in the San Bernardino mountains, 480 km farther south and 2060 m lower in altitude. The atmosphere between the altitudes of the two lakes has an absorbing power equivalent to about 2 m of water. The data in Table III were from the more sensitive of the two electroscopes, No. 3. The arrows in the table show the equivalent mean depths in the two lakes, when the difference in altitude is taken into account. Using these data, Millikan plotted curves, with the ionization readings in the two lakes as ordinates, and as abscissas, the depths in meters beneath the top surface of the atmosphere, reduced to the equivalent depths in water.

Millikan came to these conclusions: "Within the limits of observational error, every reading in Arrowhead Lake corresponded to a reading 6 feet farther down in Muir Lake, thus showing that the rays do come in definitely from above, and that their origin is entirely outside the layer of atmosphere between the levels of the two lakes." He stated further: "No single absorption coefficient is found to fit the absorption curve, the lower end of which requires a coefficient of 0.18 per meter of water; the upper end, a coefficient of 0.30 per meter of water. These coefficients correspond by Compton's equation to wave-length $\lambda_1 = 0.00038 A$ and $\lambda_2 = 0.00063 A$. These are fifty times the frequencies of ordinary gamma rays, $\lambda = 0.025 A$, and the former corresponds to an energy of 32,000,000 volts."[62]

It was this high penetrating power of the radiation observed in these experiments that convinced Millikan and the other doubters (such as Swann and Hoffmann) about the correctness of Hess's claim. And it was Millikan who gave this radiation the name *cosmic rays.*

D. "Millikan rays"?–a misunderstanding

The work of Millikan's team was important, not only because of the precise scientific results obtained, but also because of the novel and ingenious techniques employed. One great innovation was the application of sounding balloons in cosmic ray research. Hess and Kolhörster had to accompany their electroscopes in order to observe them. The use of unmanned balloons eliminated the danger and the high cost of manned balloon flights, but also the adventure. Millikan's electroscopes, masterpieces of ruggedness and sensitivity for their time, were borne aloft by two balloons; at a certain altitude one of the balloons would burst and the other would then bring the equipment gently back to Earth. During the flight, a simple device continuously recorded the electroscope readings on photographic film, to be developed and examined after recovery.

After Millikan won the Nobel Prize in 1923 for his work on the elementary charge of electricity and on the photoelectric effect, he was the most popular and respected physicist in America. He had a flair for publicity, and he obtained high praise for his experiments in snow-fed lakes. Exulting in Millikan's success, The New York Times published an editorial entitled "Millikan Rays," which said, "Dr. R. A. Millikan has gone out beyond our highest atmosphere in search for the cause of a radiation mysteriously disturbing the electroscopes of the physicists... His patient adventuring observations through twenty years have at last been rewarded... He found wild rays more powerful and penetrating than any that have been domesticated or terrestrialized, traveling toward the earth... The mere discovery of these rays is a triumph of the human mind that should be acclaimed among the capital events of these days. The proposal that they should bear the name of their discoverer is one upon which his brother-scientists should insist... 'Millikan rays' ought to find a place in our planetary scientific directory all the more because they would be associated with a man of such fine and modest personality."[63]

TIME reported, "Dr. R. A. Millikan...told the Academy about a new ray which he had discovered–a ray which begins in eternity... The Millikan Ray stabs earthward,...the Millikan Rays, wherever they are present in any quantity, have a sterilizing effect fatal to life,... The Millikan Ray will pierce six feet of lead..."[64] A picture of Millikan peering through a microscope was printed on the cover of TIME, while its caption breathlessly exclaimed. "DR. ROBERT ANDRES MILLIKAN...detected the cosmic pulse."[65]

The Scientific Monthly said, "Discovery of ultra-x-rays a hundred times more penetrating than ordinary x-rays were announced at the Madison meeting of the National Academy of Sciences on November 9 by Dr. R. A. Millikan,...some of his colleagues have suggested calling them 'Millikan rays' in his honor."[66] And SCIENCE, under the title "Millikan Rays," quoted the whole above-mentioned editorial from the New York Times.[67] All this did little to reduce Millikan's self-esteem.

In fact, however, because of Millikan's scientific dogmatism, he was twice on the losing side in controversies on cosmic rays. The first controversy concerned the existence of radiation of cosmic origin. The second time was in the 1930s and concerned the question of whether the primary cosmic rays were charged particles or, as Millikan believed, gamma rays. In the first case, after his own experiments, carried out over a long period of time, led to results that were contrary to his expectation, Millikan's skepticism evaporated.

The term "Millikan Rays" was used quite often, and Millikan enjoyed being the "discoverer" of cosmic rays. But after Hess and others expressed their chagrin about this, Millikan wrote Hess a letter, in which he said, "I made no claims of any sort about the discovery of penetrating radiations... If anybody has suffered from misrepresentation so far, it seems to me that I am the sufferer... The really important thing is that between all of us we have been able to make pretty certain the existence of a radiation which comes to the earth from outside... The evidence seems to me now to be unambiguous,... That such cosmic rays, if they exist, must be of nuclear origin is altogether obvious. It has been suggested literally scores of times."[68]

In 1936, when Hess won the Nobel Prize for his discovery of cosmic rays, Millikan expressed his warm congratulations and wrote:

> Every informed physicist will acclaim the award of the Nobel Prize in Physics to Victor F. Hess; for after a decision had been made that the first significant work in the field of cosmic rays was to be honored by a Nobel Prize there was certainly no living person who could for a moment be considered for that award except Dr. Hess. The Swiss, Gockel, the Austrian, Hess, and the German, Kolhörster, were undoubtedly the three persons who opened up this field. Their early work was done from 1910 to 1914, and no other particularly important work of this sort appeared until about a decade later, when the modern era of cosmic ray research was entered in. Gockel died about a decade ago, and Kolhörster's important work definitely followed that of Hess... . Hess...was the earliest living experimenter in the initiation of a new field of physical knowledge.[69]

VIII. THE END OF ONE ERA AND A NEW BEGINNING

In 1927, individual charged particles of cosmic rays were observed by Dmitrii Vladimirovich Skobelzyn with a Wilson cloud chamber. In 1928, Walter Bothe and Kolhörster applied the coincidence-counting method to the study of viewing fast charged particles in the cosmic rays. With these two new kinds of observation, the study of cosmic rays entered a new era: the positron and pair production were discovered; the mesotron (muon) and pion were discovered; the cosmic rays gave rise to many important discoveries. Until the early 1950s, when high energy particle accelerators took over this role, the cosmic rays were the only source of very high energy particles used for the study of elementary particle physics. Even now, the cosmic rays still provide much higher energy particles than can be produced on Earth. Studies of cosmic rays with rockets and Earth satellites have contributed greatly to the field of modern astrophysics.

NOTE: The reader who wishes to learn more about cosmic rays before 1927 will find few sources other than the original scientific papers we have quoted, especially in the English language. However, several modern books with some historical material are D. J. X. Montgomery, *Cosmic Ray Physics* (Princeton U. P., NJ, 1949); Bruno Rossi, *Cosmic Rays* (McGraw-Hill, New York, 1964); Martin A. Pomerantz, *Cosmic Rays* (Van Nostrand Reinhold, New York, 1971); Satio Hayakawa, *Cosmic Ray Physics* (Wiley, New York, 1969); A. M. Hillas, *Cosmic Rays* (Pergamon, Oxford, 1972). The forthcoming Vol. 5 of Jagdish Mehra and Helmut Rechenberg's, *The Historical Development of Quantum Theory* (Springer, New York) will contain an account especially of Austrian contributions before World War One, including Hess, and a discussion of Erwin Schrödinger's cosmic ray analysis of 1912. (We thank Dr. Rechenberg for showing us this material, and for useful advice.) For late (nearly) contemporary accounts, see Victor F. Hess, *The Electrical Conductivity of the Atmosphere and its Causes* (translation, Van Nostrand, New York, 1928); Karl K. Darrow, "Data and Nature of Cosmic Rays," Bell Syst. Tech. J. **11**, 148 (1932); R. A. Millikan, *Electrons (+ and −), Protons, Photons, Neutrons, and Cosmic Rays* (University of Chicago Press, Chicago, IL, 1935). Of some interest also is W. F. G. Swann, "The History of Cosmic Rays," Am. J. Phys. **29**, 811 (1961) and Stephen Rosen, Ed., *Selected Papers on Cosmic Ray Origin Theories* (Dover, New York, 1969).

---

[a] Permanent address: Physics Department, Xiamen University, Xiamen, Fujian Province, China.

[1] J. J. Thomson, Proc. R. Soc. London **49**, 93 (1891).

[2] J. J. Thomson, Proc. R. Soc. London **58**, 244 (1895).

[3] Nobel Lectures, *Physics* (Elsevier, Amsterdam, 1967), 1901–1921, p. 142.

[4] C. T. R. Wilson, Proc. R. Soc. London **64**, 127 (1898).

[5] C. T. R. Wilson, Proc. R. Soc. London **65**, 289–290 (1899).

[6] A. Pais, Rev. Mod. Phys. **49**, 930 (1977).

[7] J. Elster and H. Geitel, Ann. Phys. **2**, 425–446 (1900); J. J. Thomson, *Conduction of Electricity through Gases* (Cambridge U. P., Cambridge, 1906), 2nd ed., p. 3.

[8] F. Linss, Meteorol. Z. **4**, 345 (1887).

[9] C. T. R. Wilson, Proc. Cambridge Philos. Soc. **11**, 32 (1900); C. T. R. Wilson, Proc. R. Soc. London **68**, 159 (1901).

[10] J. C. McLennan and E. F. Burton, Phys. Rev. **16**, 184 (1903).

[11] C. T. R. Wilson, Proc. Cambridge Philos. Soc. **11**, 428 (1902).

[12] C. T. R. Wilson, Proc. R. Soc. London **69**, 277 (1901).

[13] J. C. McLennan and E. F. Burton, Phys. Rev. **16**, 185 (1903).

[14] R. J. Strutt, Nature **67**, 369 (1903).

[15] J. Patterson, Proc. Cambridge Philos. Soc. **12**, 44 (1903).

[16] C. T. R. Wilson, Proc. R. Soc. London **69**, 281 (1901).

[17] G. Jaffe, Philos. Mag. **8**, 564–565 (1904).

[18] G. Simpson, Trans. R. Soc. **A205**, 61–97 (1906).

[19] A. Wood, Nature **73**, 583 (1906).

[20] T. F. McKeon, Phys. Rev. **25**, 399–421 (1907).

[21] E. Rutherford and H. L. Cooke, Phys. Rev. **16**, 183 (1903).

[22] J. C. McLennan and E. F. Burton, Phys. Rev. **16**, 191–192 (1903).
[23] J. Elster and H. Geitel, Phys. Z. **2**, 560 (1901); J. J. Thomson, *Conduction of Electricity Through Gases* (Cambridge U. P., Cambridge, 1906), 2nd ed., p. 7.
[24] A. S. Eve, Philos. Mag. **13**, 248, 258 (1907).
[25] C. S. Wright, Philos. Mag. **17**, 311, 317–318 (1909).
[26] E. Rutherford, *Radioactive Transformation* (Yale U. P., New Haven, 1905), p. 197.
[27] A. S. Eve, Philos. Mag. **10**, 108 (1905).
[28] R. A. Millikan, *Electrons (+ and −), Protons, Photons, Neutrons, and Cosmic Rays* (Chicago U. P., Chicago, 1935), p. 303.
[29] C. T. R. Wilson, Proc. R. Soc. London **68**, 159, 161 (1901).
[30] C. T. R. Wilson, Proc. Cambridge Philos. Soc. **11**, 428–430 (1902).
[31] C. T. R. Wilson, Proc. Cambridge Philos. Soc. **12**, 85 (1903).
[32] C. T. R. Wilson, Nature, **68**, 103 (1903).
[33] C. T. R. Wilson, Nature, **68**, 104 (1903).
[34] G. Simpson, Nature **69**, 270 (1904).
[35] C. T. R. Wilson, Proc. Cambridge Philos. Soc. **13**, 378, 382 (1906).
[36] O. W. Richardson, Nature **73**, 607 (1906).
[37] Karl Kurz, Phys. Z. **10**, 834 (1909).
[38] J. Elster and H. Geitel, Ann. Phys. **2**, 441 (1900).
[39] Th. Wulf, Phys. Z. **10**, 153 (1909).
[40] Refer to Refs. 45 and 48 below.
[41] Th. Wulf, Phys. Z. **11**, 811 (1910).
[42] A. S. Eve, Philos. Mag. **21**, 27–40 (1911).
[43] J. C. McLennan and E. N. Macallum, Philos. Mag. **22**, 640–642 (1911).
[44] H. Ebert, Ann. Phys. **5**, 718–724 (1901).
[45] K. Bergwitz, Habilitationsschrift, Braunschweig, 1910, s37; M. A. Pomerantz, *Cosmic Rays* (Van Nostrand Reinhold, New York, 1971), p. 4.
[46] Armin Hermann, *The New Physics* (Heinz Moos, Munich, 1979), pp. 11–13.
[47] Karl Kurz, Phys. Z. **10**, 834–845 (1909).
[48] A. Gockel, Phys. Z. **11**, 280–283 (1910).
[49] A. Gockel, Phys. Z. **12**, 595–597 (1911).
[50] V. F. Hess, Phys. Z. **12**, 998–1001 (1911).
[51] V. F. Hess, Phys. Z. **13**, 1085–1086 (1912).
[52] V. F. Hess, Phys. Z. **13**, 1088–1091 (1912).
[53] Nobel Lectures, *Physics* (Elsevier, Amsterdam, 1965), 1922–1941, p. 356.
[54] W. Kolhörster, Phys. Z. **14**, 1153 (1913); Dtsch. Phys. Ges. **16**, 719 (1914).
[55] V. F. Hess, *Electrical Conductivity of the Atmosphere and its Causes* (Van Nostrand, New York, 1928), p. 117.
[56] A. Gockel, Dtsch. Schweiz. Natur. Ges. **54**, 1 (1917); Hess and Kofler, Phys. Z. **18**, 585 (1917).
[57] W. Kolhörster, Z. Phys. **11**, 379 (1922).
[58] Millikan to Hess, July 24, 1926, Millikan Collection Caltech.
[59] Millikan and Bowen, Phys. Rev. **27**, 360–361 (1926).
[60] Millikan and Otis, Phys. Rev. **27**, 650–657 (1926).
[61] Otis and Millikan, Phys. Rev. **23**, 778–779 (1924).
[62] Millikan and Cameron, Phys. Rev. **28**, 851, 856 (1926).
[63] The New York Times, November 12, 1925, p. 24.
[64] TIME **6**, 26, November 23, 1925.
[65] TIME **9**, Cover, and p. 16, April 25, 1927.
[66] The Scientific Monthly **21**, 661 (1925).
[67] SCIENCE **62**, 461–462 (1925).
[68] Millikan to Hess, July 24, 1926, Millikan Collection Caltech.
[69] Millikan, Award of Nobel Prize in Physics to V. F. Hess and Carl D. Anderson, November 23, 1936, Millikan Collection Caltech.

The vibrating string controversy

Gerald F. Wheeler
Department of Physics, Montana State University, Bozeman, Montana 59717
William P. Crummett
Department of Physics and Geophysical Engineering, Montana College of Mineral Science and Technology, Butte, Montana 59701

(Received 6 December 1985; accepted for publication 10 February 1986)

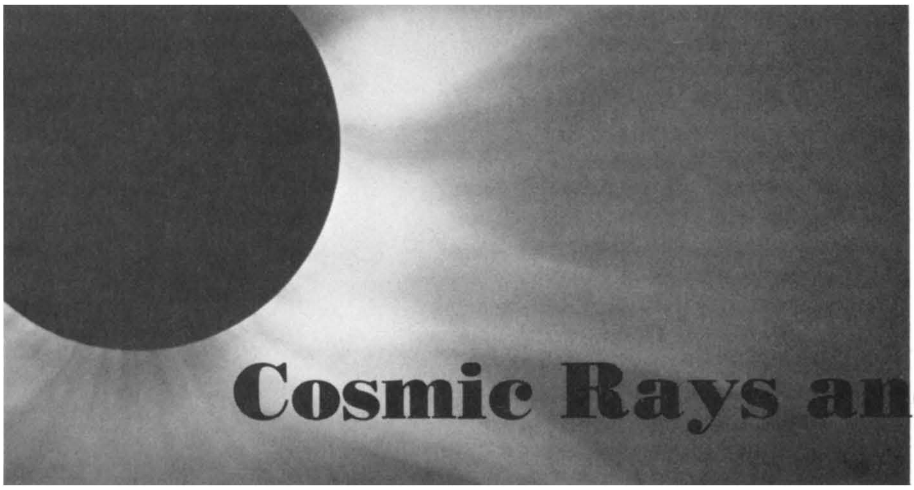
In the mid-1700s a debate raged between Jean d'Alembert, Leonhard Euler, and Daniel Bernoulli concerning the proper solution to the classical wave equation. This controversy was partially solved by Lagrange and, more conclusively, by Fourier (50 years later) and it provides an interesting case study for the role of mathematics in the modeling of physical phenomena. Of particular note in this debate, was the meaning of boundary conditions. The controversy is summarized from the point of view of this mathematical physics perspective.

INTRODUCTION

Mathematical descriptions of wave phenomena are fundamental to many areas of physics. A clear understanding of the relations which describe the vibrating string is required to comprehend more complex wave motions. Few physicists, however, are aware of the intense controversy that existed over the original descriptions of the vibrating string proposed during the eighteenth century. At the height of the controversy one of the most fundamental and powerful theorems of mathematical physics emerged, was overlooked, and had to wait 50 years for its rediscovery.

While the debate has long held interest for mathematical historians, there has been little discussion of the way this debate signaled the emergence of a new kind of physicist. There are excellent reviews of the controversy,[1,2] each presented as a topic from the history of mathematics. In presenting our view of the debate, we have drawn extensively from these sources, as well as the original papers.

Physicists will find the controversy enlightening. Many



Cosmic Rays and Eclipses

Serge A. Korff

One of the many interesting facets of nature is the way that apparently diverse fields of research turn out to have common connections. This is especially true of the field of cosmic rays, which connects with a wide variety of other branches of physics, geophysics and astrophysics. In examining one of these connections we must necessarily limit ourselves and omit much that is interesting and could be said about each field by itself, not to mention all the peripheral ones.

Cosmic rays are composed mostly of protons, with some alpha particles, and a few heavier nuclei, plus a small number of electrons and photons, arriving at the earth from outside the solar system. A semantic confusion occurs, because some high energy particles originating in the sun also arrive at earth. In this article I shall use the adjective "cosmic" to refer to entities from outside the solar system, the "cosmos." The energies of these entities are individually very high, especially when compared with those produced by man-built accelerators. A few of our largest machines generate beams in the range of 100 Gev (10^{11} ev), yet this is the bottom of the cosmic ray spectrum, which goes up here by some ten orders of magnitude. The lower energies, up to 10^{11} ev or so, are measured by geomagnetic effects, since charged particles in a magnetic field follow orbits calculable from electrodynamics. At the higher energies the curvatures of such particles in a field as weak as the earth's, of the order of 0.3 G, are unmeasurably small, even given the radius of the earth as the measuring-scale. The higher energies are known from a study of secondary-production effects. For example, cascade theory permits the calculation of the energy of the particle causing a "giant shower" of tens of thousands of secondary particles. Such giant showers may cover tens or more of square kilometers of ground at the earth's surface. They can be studied by distributing trays of Geiger counters and other detecting equipment in appropriate arrangements. Indeed we do not know today where the top of the energy spectrum is, but only that each extension or improvement in technique shows us that we have not yet reached it.

Further, we do know from astronomical evidence that space, for many thousands of light-years in all directions from the solar system, is filled with a very diffuse and low density gas, perhaps averaging one proton per cubic centimeter or an order of magnitude either way. These protons have been exuded from stars, or may indeed have survived from much earlier epochs. They move with average speeds in the range of tens to hundreds of km/sec, and therefore generate a magnetic field in space. The order of magnitude of this field is perhaps a microgauss, although this is again not well known and the number cited is uncertain

Serge A. Korff *received his Ph.D. at Princeton in 1931. He has been on the New York University faculty for 35 years and is now Professor Emeritus. He is also Past President and current Chairman of the Council, American Geographical Society, Past President and current member of the Board of Governors, New York Academy of Sciences, and Past President of the Explorers Club. (New York University, New York, New York 10003)*

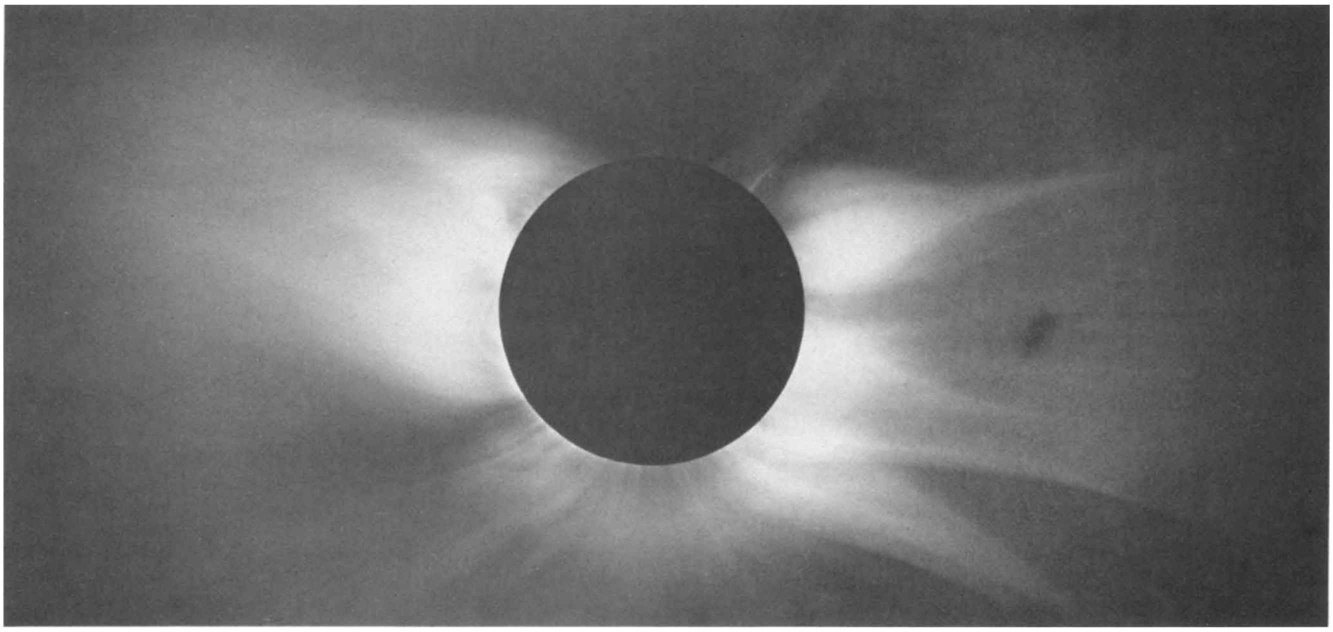


Fig. 1. Photograph of the corona of the sun, taken with a radial density gradient filter. Eclipse of 30 June 1973, taken at Loiengalani, Kenya, by High Altitude Observatory, a division of National Center for Atmospheric Research, sponsored by the National Science Foundation. Typical solar minimum, long streamers. Note that at one of the poles is a large area where no corona seems attached, and also at several other smaller places are areas that can be called "holes" in the corona. The major structures are quiescent, and show little variation with time over a period of some hours. The solar wind may originate mainly in the holes.

by an order of magnitude, but is given to define the "ball-park" we are discussing. The direction of the vector defining this field is not known, and presumably varies from place to place. The net result of this tenuous diffuse and random galactic field is that the directions from which the primary cosmic ray particles have come are scrambled and beyond experimental reach with present technology. The earth is embedded in a gas of high energy cosmic ray particles traversing a gas of lower-energy matter.

We speak of protons in space. There must of course be an equal number of electrons, for space cannot have a net charge. But the electrons do not have to share the high energies of the protons, and of course if they share the average velocities they have much lower average energies. Electrons have much more efficient energy-loss mechanisms, such as x-ray production (Bremsstrahlung) than protons.

We do not know the origin nor the accelerating mechanism of the cosmic radiation. Many theories have been proposed and each is ardently espoused by a group of enthusiasts. For example, some believe that cosmic rays originate in supernova explosions. In this model, they are accelerated in complicated shock-wave mechanisms of magnetohydrodynamic types. A supernova, spectacular though it is in terms of the huge acceleration imparted to an enormous mass of material, is a very low-energy event in terms of velocities of individual protons moving outward in the giant blast. But the fascinating problem of the origin of the cosmic radiation is another story which we must defer in this case.

In this article we look at a more mundane problem. What happens when the cosmic rays impinge on the earth's atmosphere? This study is aided by controllable experiments with accelerators, for in nature a broad spectrum of energies and particle types is incident from random directions on a complex target. In the laboratory we can isolate the primary types, the target composition, the beam energy and impact momentum. At energies well above the nuclear binding levels of a few MeV's the production of secondaries by collision of primary particles with nuclei in the upper atmosphere becomes quite complex. Secondary nucleons are formed, as well as all the intermediate entities, such as mesons, hyperons, leptons and neutrinos presently known and perhaps some not yet discovered. Among the secondary nucleons are neutrons, which are present in the target atoms in the atmosphere. Because of their short radioactive half-life, however, they would decay in flight, and therefore cannot be a part of the primary radiation. Once again, we cannot here examine all these processes in detail, but shall confine our remarks to the neutrons.

Neutrons have become the principal indicator of cosmic ray intensities for two reasons. The first is that good techniques such as neutron detectors (often consisting of large proportional counters filled with BF_3 gas in which the incoming neutron is captured and an alpha particle is emitted) permit easy detection, and the second is that because of a multiplying effect, variations in neutron intensities are extremely sensitive indicators of the variations in cosmic ray intensities. A primary proton can produce many neutrons. The number depends in a complex way on the energy of the proton and the target type. The details again we do not have the space to present here, but the net result is that an effect in space which produces a five per cent change in the intensity of the ionizing component of the cosmic radiation may show as a twenty per cent change in the neutron intensity.

Quite incidentally, the neutrons generate a multitude of interesting phenomena in altering isotopic types. They form both radiocarbon (C^{14}) and tritium (H^3) in the atmosphere when captured by nitrogen nuclei, according to:

$$_0n^1 + {}_7N^{14} \rightarrow {}_7N^{15*} \rightarrow {}_6C^{14} + {}_1H^1 \qquad (1a)$$

$$\text{or} \rightarrow {}_6C^{12} + {}_1H^3 \qquad (1b)$$

The first reaction has led to Libby's great contribution to archaeology, carbon-14 dating, and the second to useful advances in oceanography. Many other isotopes also are formed, which again we do not have space to discuss here. It will suffice for this article to say that both the neutron intensity and the intensity of the ionizing component of the cosmic radiation are today monitored with considerable accuracy at various stations, at different latitudes, longitudes and altitudes on earth, and by instruments carried in space vehicles.

Let us temporarily set aside consideration of the details of particle physics and turn to a brief description of the sun, the corona, and solar physics. Eclipses of the sun by the moon occur with some regularity, more often than newspaper accounts sometimes imply. The path of totality, of the moon's shadow as it sweeps across the earth, is seldom a hundred miles wide and often much less. The travel speed of this shadow is of the order of 1000 miles an hour, and the duration of totality is between zero or a few seconds and a few minutes, seldom exceeding seven minutes. Since 0.71 of the earth's surface is water, and much of the land area accessible only with great difficulty, total eclipses are seldom seen in convenient places. Ground stations are necessary for stable platforms, for long exposure photographs or heavy equipment.

It is only during totality that the sun's corona can be seen (Fig. 1). The problem revolves around relative intensities, for the brightness of the corona is of the order of a millionth of that of the bright solar disk or photosphere. There is an instrument called a coronagraph which attempts to overcome this difficulty, and enables some idea of the corona to be obtained between eclipses. Yet even the best of these instruments do not today give the ability to discern the fine detail that can be seen during an eclipse. The scattering of sunlight by the earth's atmosphere is the principal villain of the piece, since scattered skylight near the sun is many times brighter than the corona. Scattered light within the instrument is the source of the most difficult intensity problem that the coronagraph itself faces. The astronauts have obtained excellent photographs of the corona from space stations by using a coronagraph with an occulting disk appreciably larger than the solar disk image. Still we cannot do all our experiments in space stations, so eclipses are of interest and importance. The total time that the corona has been seen since the development of good instruments is at most a few hours. That we have learned as much about it as we have is largely due to careful preparation of experiments and equipment before the eclipse takes place.

We have learned that the corona is extremely tenuous and extremely hot. The temperature estimates come from a study of the spectral lines emitted by the coronal material. The individual energies needed to excite these lines, for example that of Fe XIV, (not a French king, but iron that has lost thirteen electrons) indicates temperatures in the corona between 1.5 and 2 million degrees K. The pressures depend on how far out radially from the disk we made the measurement, but are so low that they correspond to a good laboratory vacuum. We only see the material at all because of an optical path length through the corona of many hundred thousand kilometers. The composition of the corona is roughly that of the sun itself, largely hydrogen, next helium, then some heavier atoms, most of which lack one or many electrons, plus the free electrons that have been subtracted from these ions. Since in the scattering of light by small particles, including scattering by free electrons and ions (treated in detail by Rayleigh, Mie,

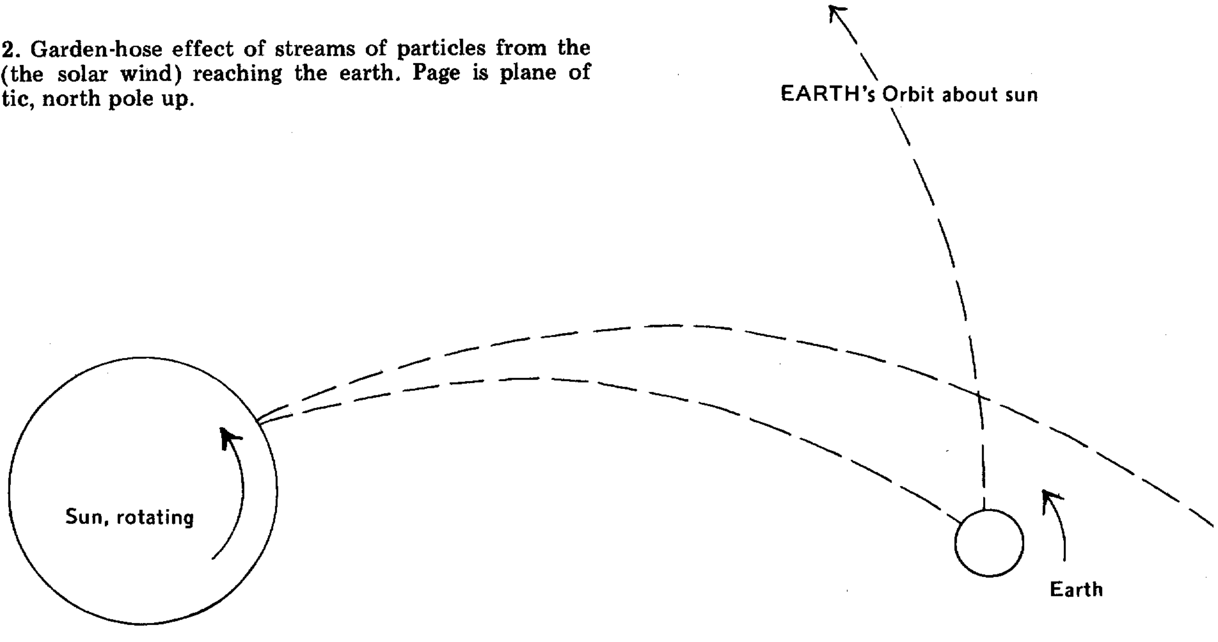


Fig. 2. Garden-hose effect of streams of particles from the sun (the solar wind) reaching the earth. Page is plane of ecliptic, north pole up.

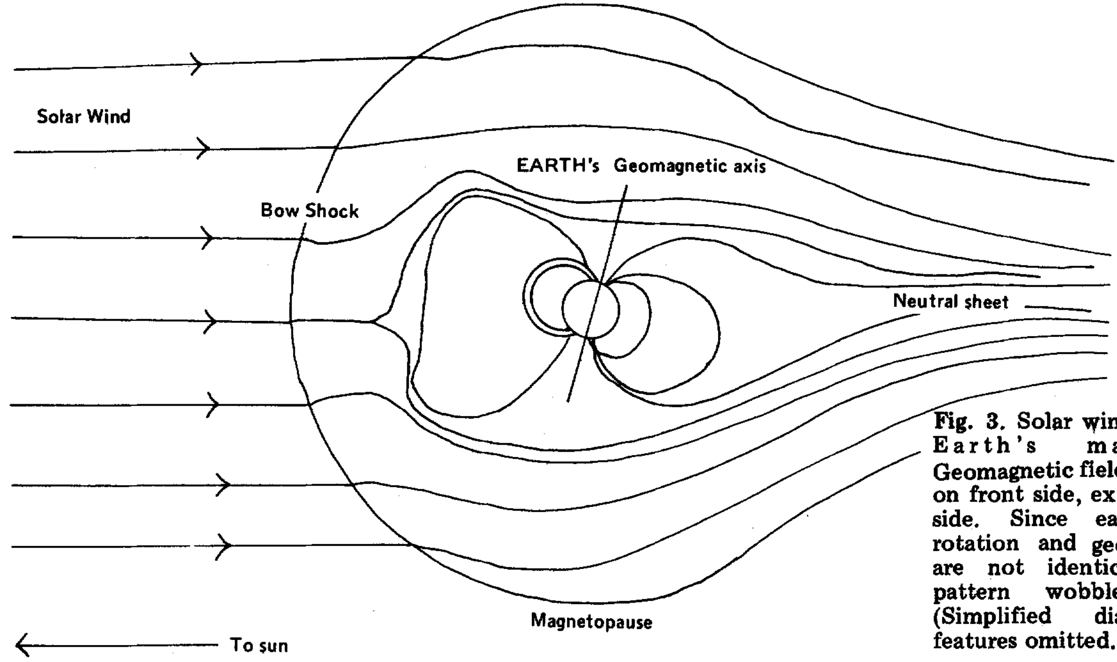
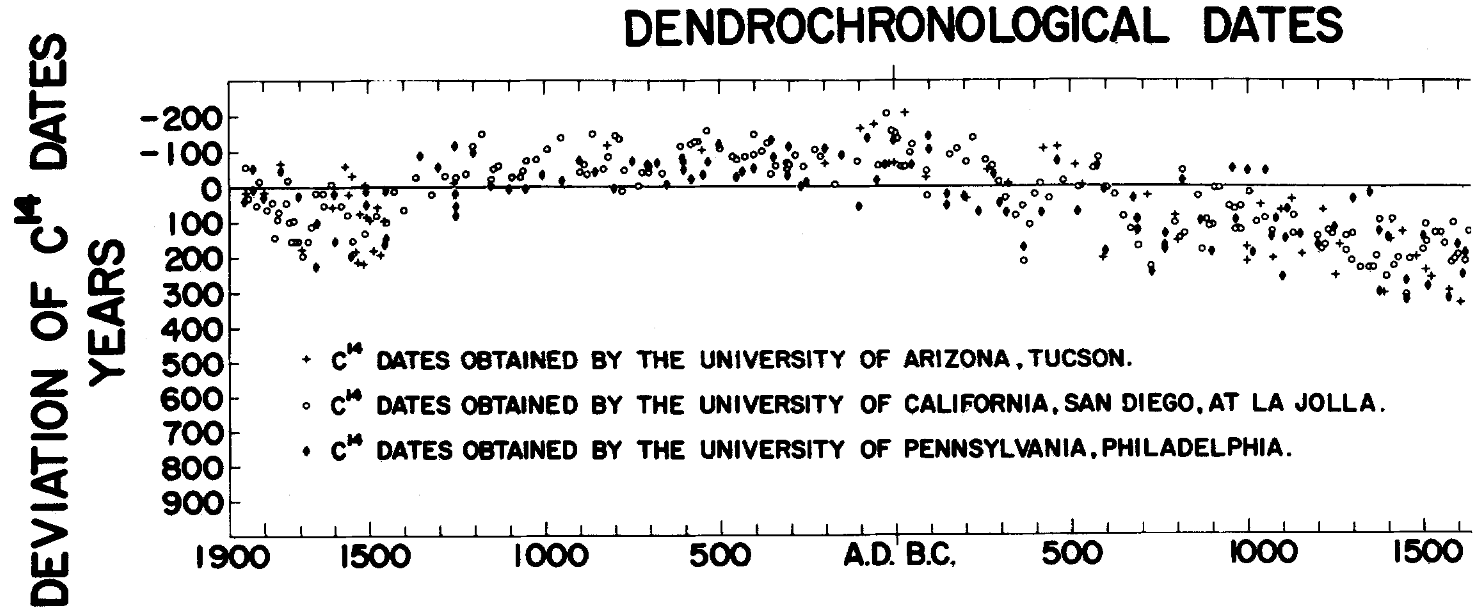


Fig. 3. Solar wind impinging on Earth's magnetosphere. Geomagnetic field is compressed on front side, extended on back side. Since earth's axis of rotation and geomagnetic axis are not identical, the entire pattern wobbles in space. (Simplified diagram. Many features omitted.)

and others) the energy scattered depends inversely on the mass of the scattering particle, the free electrons do most of the scattering. Because of the high velocities which these electrons have, corresponding to the high coronal temperatures, Doppler broadening smears out scattered Fraunhofer absorption lines beyond recognition. The Fraunhofer spectrum is the normal dark-line spectrum of the sun, in which the continuous spectrum of the hot photosphere is crossed by absorption lines generated in the layer (called the reversing layer) immediately above the disk.

The forms of the corona vary greatly from one eclipse to the next. Sometimes it is fairly uniform, but at other times long streamers can be traced for four or more solar radii out into space. Some of the shapes of the streamers, especially in the inner parts of the corona and near the poles of the sun, remind one immediately of magnetic field shapes, occasionally simple dipole fields but more usually rather complex structures. Turbulent processes at and below the solar disk or photosphere often generate strong magnetic fields which impart inductive accelerations to ions and electrons. Such phenomena as solar flares and sunspots have associated magnetic fields. The net result is a very complex motion of the material. Because of the high solar gravity and high escape velocity, over 800 km/sec, most of the material is held by the sun, but occasionally large structures have been seen to blow away and completely leave the sun.

Today space probes carry plasma detectors, ionization-measuring devices and magnetometers. These have permitted the development of extensive data on the region between the earth and the sun, and indeed going out now to beyond the orbit of Jupiter. We have learned that the sun continually exudes a plasma (Fig. 2), which shows long-term fluctuations following the sunspot cycle, an eleven-year period, and short-term fluctuations of a few



days or a few hours which take place when active regions develop on the sun. These may be correlated with solar flares and are often called "solar flare effects," although not all flares generate observable disturbances in the solar plasma or normal solar wind. Because of the very long free paths in the material between us and the sun, a burst of high energy charged particle radiation can go right through the solar wind with very few collisional effects. However, a substantial burst of particles will also generate a magnetic field by virtue of its motion, and this will interact at appreciable distances with other charged particles. The normal solar wind has particle densities of a few particles per cubic centimeter moving with speeds of 300 to 600 km/sec, and generating or carrying with it from the sun interplanetary magnetic fields of several gammas (a gamma is 10^{-5} Gauss) when measured at the earth's distance from the sun.

The solar wind impinges on the earth's magnetic field and creates a shock called a "bow shock," which moves with the earth in the direction of the earth's motion relative to the plasma, and a long tail in the opposite direction (Fig. 3). When a blast wave of additional particles comes from the sun the net magnetic field at the earth's surface which is the vector sum of the earth's own field and the field carried by the particles, often shows substantial variations. This effect is called a "magnetic storm" and is also manifest by other activities. The particles may actually be seen forming aurorae, the ionization levels in the earth's atmosphere may change, and the field which controls the entrance of cosmic rays to the earth's environs changes. Therefore cosmic ray intensity variations are observed, both on earth and in space vehicles. Thus the solar wind by itself modulates, and its changes change the cosmic ray intensity. The changes in number of neutrons observed here on the earth's surface is thus another means for studying the solar wind and solar physics. The development of this topic and the experimental study of the interactions is one of today's hottest research areas, both thermally and in its inherent interest, as a function for generating new knowledge.

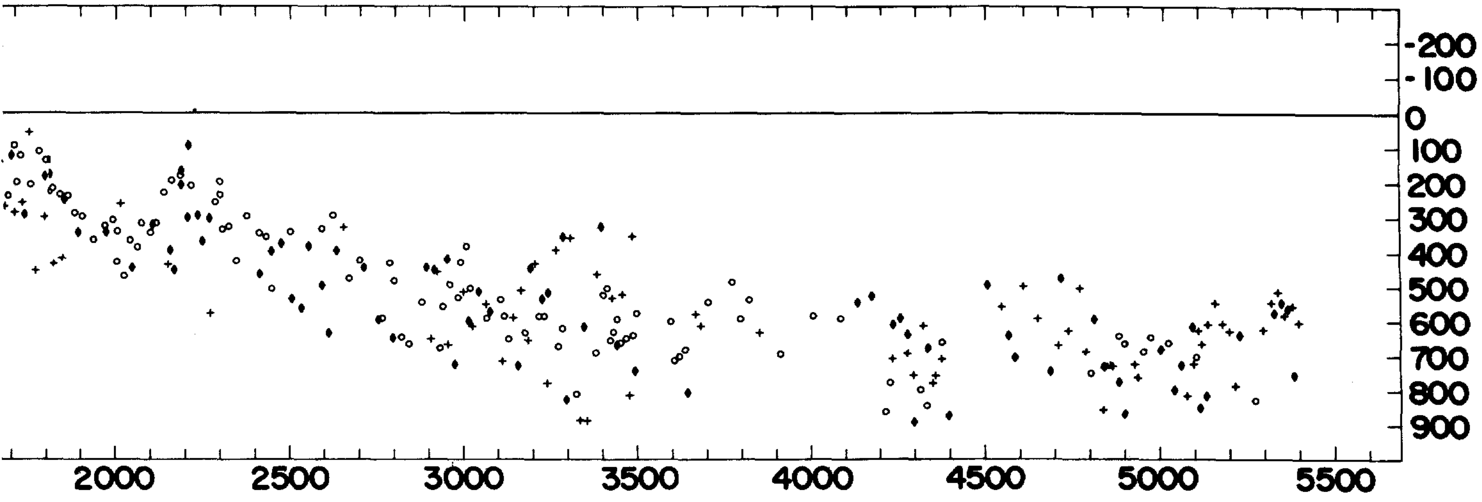
Somehow the solar wind manages to blow through the corona. We do not know the mechanisms that originate the solar wind, nor those that accelerate it. We do not even know where on the sun the wind has its source. Many suggestions have been made. Through them all runs the thread that about the only way to get charged particles up to high energies is by having them in a varying magnetic field, such as those found in betatrons or other accelerators, where inductive acceleration processes transfer energy from the field to the particle.

For sometime our group at NYU has been interested in attempting to see if we could learn more about this process by looking closely at the corona. We wondered whether the corona might show motions of ionized material which in turn might give some clue to an acceleration mechanism. To this end we decided to take photographs of the corona at various time-intervals. We had tried this on earlier eclipses, both on the time scale of 100 sec during an eclipse, and 1000 sec by cameras along an eclipse path. Now a unique opportunity developed, for identical cameras located for the eclipse of June 30, 1973 on the west coast of Africa in Mauritania, and near the east coast, in Kenya, would be separated by almost 10 000 sec of shadow travel time. We took such photographs, and projected them overlapping one another. By shutting off one or the other projector, it was possible to look for differences in the corona. No differences were found. The experiment would have revealed motions of as little as 10 km/sec in any streamer or intensity feature. Yet the plasma observers reported that the solar wind blew constantly throughout this time interval.

This means that the main optical features of the solar corona, the long streamers and other shapes and structures, are very stable, at least on the time scale of 10^4 sec. This clearly leaves two possibilities, first that the solar wind is generated in areas where there is no appreciable optically visible corona, or second that the acceleration processes, whatever they may be, take place out beyond the corona or wherever the optical corona is missing. We use the word "optical" repeatedly for the observations we made were in the visible portion of the spectrum, and we did not in this experiment observe the situation in other spectral regions. The further conclusion is that the optical corona has its

Fig. 4. Radiocarbon dates minus dendrodates. Example: 4000 dendro is 4700 radiocarbon. Sun may have changed average emission, at 700 BC and at 3800 BC. Data from H. Seuss et al. Ref. MASCA Newsletter V 9 #1, Aug. 1973.

(As explained on page 338, "dendro"-dates are dates obtained from analyzing tree rings.)



shape controlled by magnetic fields anchored in the solar surface. Further, there do appear in the photographs to be "holes" in the corona, or regions where there are no long streamers attached. Such holes are hard to see in pictures, for any photograph is an integration along the line of sight.

Just recently we have seen some beautiful pictures taken from Skylab in which a nice stable corona was in place when a disturbance on the surface of the sun sent a great pulse out through it. After the pulse the corona was still stable but had changed in shape. Thus the magnetic fields which control its shape had undergone a rather abrupt change from one configuration to another.

At the present time there are also excellent pictures taken in the x-ray region of the sun. These pictures show activity in the form of generation of x radiation from active areas on the sun. These areas cover only a small fraction of the solar surface. We tend to think of them as the "hot" areas, the word being used to refer to excitation conditions rather than to thermal heating on any massive scale.

Many persons believed that the solar wind originated in the hot areas, of high excitation, by some process not well understood. However, the fact that the solar and coronal plumes, helmets and streamers are so stable over such long time-scales, and show little evidence of outward motion, throws doubt on this interpretation. On the other hand if the wind originates in the "cold" areas, and blows out through holes in the corona, the major optically visible features could retain stability and yet the wind could blow. The interpretation therefore which seems better to fit the observations is that the solar wind is accelerated by some electrodynamic mechanism operating in the "cold" regions, where the strong fields that maintain the coronal forms are absent. Thus our opinion today is the opposite of what many believed earlier.

The details of the mechanism that cause a solar wind to blow have not yet been worked out to the satisfaction of more than a few enthusiasts. However, the general principles are evident, since to accelerate charged particles what is needed is a magnetic field that varies as a function of time. A charged particle in such a field will pick up energy from the field, or lose it to the field, depending on circumstances. This is essentially the betatron mechanism, and is presumably in general terms the operative mechanism, although the exact details remain to be worked out.

If one looks at the sun with radio telescopes, one sees that the sun's apparent diameter increases as the frequency decreases, in other words radiofrequency radiation is generated far out radially from the optical disk. This observation also is material to our search for understanding, for radiofrequency radiation is a consequence of motion of the charged particles at that place. Here again the evidence shows that charged particles are in motion far out beyond the solar limb. It may be out here that the solar wind is given the velocity that permits it to escape from solar gravity and drives it out past the earth into outer space.

Another piece of evidence about long term solar variations comes from a comparison of the radiocarbon dates with the tree-ring or "dendro-" dates. Since the dendrodates now extend back for some 8500 years, a comparison of dendro- and radiocarbon dates is very revealing (Fig. 4). What it shows is that as one goes backwards, the two systems are in good step until about 500 BC, where the curve starts to change. Then at about 3500 BC, there occurs another change in the slope of the radiocarbon curve. This suggests that twice the sun has changed the average amount of solar wind that it exudes. It also means that radiocarbon dates must have a correction applied. The older interpretation was made on the assumption that the rate of production of radiocarbon was constant. This is evidently true only to a first approximation. Actually, the production is very complex, varying both with latitude, altitude and state of the sunspot cycle. However, because of mixing in the oceans and atmosphere, the large variations are much smoothed out. The other smoothing factor is the long (5700 year) half life of radiocarbon.

Thus we see several important facets of the situation. First, cosmic rays, solar physics and archaeological dates are closely tied to one another, a general commentary on the unity of nature. Second, we also see how advances in one field, for example cosmic rays, may quite unexpectedly provide an important tool, in this case radiocarbon dating, for an apparently quite different one, archaeology. And finally, just about all the frontier fields of our knowledge are in a state of flux. What we thought we knew, and what seemed safe assumptions, need to be considered anew from time to time, and often require modification as we go along. Further, these modifications, irritating as they are when they occur to the workers who just thought they had the situation figured out, actually lead to better understanding of the many complex factors at work, to more accurate measurements, and to more glimpses of the complex interrelationships between various apparently disparate and unconnected natural phenomena. In science we seek to understand how this world of ours operates. It is always more complex and interrelated than we had at first thought. Yet by following the interrelationships we at the same time improve our overall understanding.

CHANGES AT AAPT

Executive Officer Arnold Strassenburg has begun a two year leave of absence from AAPT and the State University of New York to go to Washington D.C. with the National Science Foundation. Melba Phillips will take his place as Acting Executive Officer of AAPT. Also joining the AAPT staff is Dean Zollman who will be Staff Physicist during his two year leave of absence from Kansas State University in Manhattan, Kansas.

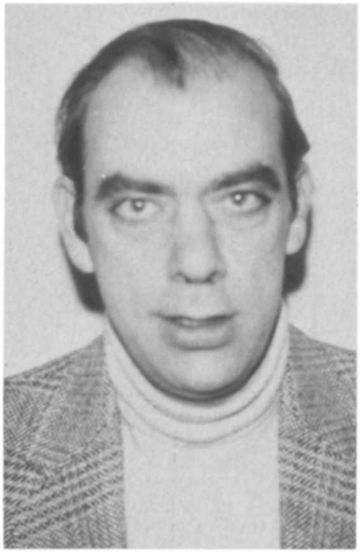
During the summer the AAPT offices moved a half mile from their old quarters into the new physics building at Stony Brook University. All correspondence with AAPT, including subscriptions for *The Physics Teacher*, should be addressed:

AAPT
Graduate Physics Building
SUNY at Stony Brook
Stony Brook, NY 11794

Matter from outside our solar system - new insights

Part I: The astrophysical framework

JOHN P. WEFEL



John Paul Wefel *received his B.S. from Valparaiso University and his M.A. and Ph.D. degrees in Physics/Astrophysics from Washington University, St. Louis. He worked at the Naval Research Laboratory before coming to the University of Chicago as a McCormick Fellow in 1975, becoming a Senior Research Associate in the Fermi Institute in 1977. In 1982 he was appointed Assistant Professor of Physics at Louisiana State University. Dr. Wefel holds memberships in APS, AAPT, AAAS, and AIAA. (Laboratory for Astrophysics and Space Research, Enrico Fermi Institute, Chicago, IL)*

One of the most fundamental problems in the fields of astronomy and astrophysics is the question of the origin and distribution of the elements in nature. Is a rock from the earth of the same composition as one from the moon or Mars? Is a bucket of matter from the sun identical to one obtained from other stars, from the dark clouds in the direction of the galactic center, or from the rarefied gas of the interstellar medium? If there are differences, are these due to spatial or temporal changes within our solar system or in other parts of the galaxy? What causes such changes, and can they be predicted within models of chemical evolution for the galaxy or for specific regions within it? Researchers in many fields, astronomy, space science, geophysics, astrophysics, to name a few, have been trying to answer these questions for many years.

With the advent of the space program we have seen the return of samples from the moon, on-site observations of Mars, remote sensing of the sun and the planets in the inner solar system, and fly-bys of the gas giants Jupiter and Saturn. These new sources of information have been combined with the more traditional data from solar spectroscopy and from the analysis of terrestrial and meteoritic samples to give, today, a detailed picture of the distribution of elements within our solar system and the composition of the matter at the birth of the solar system about 4.5 billion years ago. It was at this time that the solar nebula withdrew from any further mixing with the gas and dust in the galaxy and began the inexorable process of gravitational collapse to form a new star – our sun – and its planetary system.

The galaxy, however, has not remained static since the birth of the solar system but has continued to evolve and change, and astronomical observations have revealed some of this change. In the last several decades both ground-based and satellite-borne observations have opened new electro-magnetic "windows" to the universe such that now the abundances of certain elements can be measured in many different types of stars, in molecular gas clouds, in some supernova remnants, and in the diffuse interstellar medium. This astronomical revolution has revealed a surprising diversity to the element patterns in different objects and in separate parts of the galaxy. These patterns constitute the fundamental clue to the current conditions within the objects and to their history.

Into this arena come the cosmic rays. These high-energy particles, composed of nuclei of all of the atoms in the periodic table, represent the only sample of matter from beyond our solar system available for direct study. The cosmic rays are themselves relatively young, having traveled in the galaxy for approximately 20 million years, and although their exact sources are still unknown, the information contained in the distribution of elements and isotopes is already bringing new insights into the processes of nucleosynthesis and chemical evolution in the galaxy.

Following a "capsule" review of the history of cosmic-ray astrophysics and a discussion of the accumulated data, pointing to the need for isotope measurements, the theory of heavy-element nucleosynthesis is reviewed to provide a basis for the interpretation of the cosmic-ray data. In the second installment of this article, the experimental techniques, which represent a revolution in their own right, are described before concentrating on the isotopic ratios and their interpretation. The element neon provides the most significant result, showing 3 to 4 times as much of the neutron-rich isotope ^{22}Ne (relative to the more common isotope ^{20}Ne) in the cosmic-ray source as is observed in most solar-system materials. The explanation of this "neon anomaly" has already generated a large amount of theoretical work, some of which will be described.

Cosmic-ray physics: an historical overview

The cosmic radiation was first discovered in the early years of this century as the agent responsible for the discharging of electroscopes in the laboratory. In the years 1912-1914 Hess and Kolhorster, in a remarkable series of manned balloon ascents, showed that this ionizing radiation came from "the sky" and not from the ground. These experiments established the extraterrestrial character of the cosmic rays and were followed by several decades of investigation designed to elucidate the nature of the cosmic-ray phenomenon. Experiments were performed at various altitudes, from sea level to the tops of mountains, and on ships during voyages around the world. These latter experiments showed that the radiation was affected by the earth's magnetic field. With the discovery of an east-west asymmetry, the nature of the cosmic rays – positively charged particles – was established. The names of Millikan, Wilson, Compton, Clay, Stormer, and Rossi, to mention only a few, are intimately connected with the discoveries made in this initial period in the history of cosmic-ray research.[1-3]

In 1937 the nature of cosmic-ray research underwent a change with the discovery of the muon in the cosmic radiation. This discovery ushered in an era, encompassing almost a decade and a half, during which the cosmic radiation was viewed as a natural source of high-energy particles – the accelerator in the sky – for the study of elementary-particle phenomena. Many new tools, such as cloud chambers, nuclear emulsions, and Geiger-tube arrays, were developed for research into the masses, decay schemes, and lifetimes of particles such as mesons and hyperons. Nuclear interactions of cosmic-ray protons and the nature of cosmic-ray-induced showers became the prominent areas of research with only a modest effort devoted to the study of the nature of the radiation itself. With the advent of the synchrotron and later particle accelerators in the 1950s, research in elementary particle (or high-energy) physics moved from the cosmic-ray laboratory into the new accelerator centers.[1,4,5]

In the late 1940s it was discovered that the arriving cosmic rays, those observed in balloon experiments near the top of the atmosphere, contained more than just protons. The nuclei of helium and still heavier elements were found.[6] This discovery led to a profound change in the way in which the cosmic radiation was viewed, since it became clear that the cosmic rays were actually a sample of galactic matter. In the early 1960s primary electrons and gamma rays were found, followed in the mid-1960s by the discovery of ultra-heavy cosmic rays, the nuclei of atoms heavier than nickel (encompassing the upper two thirds of the periodic table). In the early 1970s the low-energy anomalous component was identified (a class of particles whose origin may be quite different from that of most cosmic rays), and recently evidence has been obtained for antiprotons among the cosmic rays. These discoveries ushered in the modern era of cosmic-ray astrophysics[7] in which the cosmic rays are used as "probes" of the astrophysical processes occurring in other parts of our galaxy. In particular, this paper deals with the heavy elements in the cosmic rays as probes of nucleosynthesis and chemical evolution in the galaxy.

The heavy elements are relatively rare – only about 14% of the cosmic rays are helium and less than 1% of the particles have a nuclear charge $Z \geqslant 3$. The cosmic rays cover an enormous energy range, 15 orders of magnitude, from $\sim 10^6$ eV (electron volts) to the highest recorded energy of 10^{21} eV. (In the remainder of this paper the notation MeV = 10^6 eV and GeV = 10^9 eV will be employed.) The energy spectrum is a power law with the intensity decreasing as the energy increases. The bulk of the particles have energies below several GeV/nucleon, and most of the experimental data on heavy elements has been obtained between 100 MeV/nucleon and 1 GeV/nucleon. The absolute intensity of the cosmic radiation can be illustrated by considering a one-square-foot detector located above the earth's atmosphere and magnetic field. If this hypothetical detector records only particles with energies in the range 100-300 MeV/nucleon, then it will "see" about 200 protons and 35 helium nuclei every second. However, there will be only three medium nuclei (carbon, nitrogen, or oxygen) every two seconds, one neon every six seconds, and four iron nuclei per minute. Unfortunately, most cosmic-ray detectors on satellites are considerably smaller than a square foot (typically square inches) which extends the time required to collect the events. For the ultra-heavy cosmic rays the situation is much worse. Our hypothetical detector will observe one krypton ($Z = 36$) nucleus every two weeks and six lead nuclei ($Z = 82$) in a full year of operation. Clearly, much larger detectors are needed to study the ultra-heavy component.

Figure 1 from the work of Garcia-Munoz and Simpson[8] gives a comparison, for the elements hydrogen through nickel, between the element abundances measured at earth in the cosmic rays at low energy and the abundances of the elements in our solar system.[9] There is a general similarity between the two distributions, but the cosmic rays show a large overabundance of the light elements, lithium, beryllium, and boron, the elements just below iron, and the odd-Z nuclei. This overabundance of rare species indicates that the cosmic rays have undergone nuclear spallation reactions in which nucleons have been removed from the nucleus before reaching the earth. Such reactions break up the abundant elements, such as carbon, oxygen, and iron, forming fragments (secondary particles) concentrated just below the parent element. For example, the spallation of the numerous carbon and oxygen nuclei forms the lithium, beryllium, and boron observed in the cosmic rays at a level thousands of times higher than these same elements are found in the solar system. Thus the cosmic rays observed at earth are a mixture between particles accelerated at the cosmic-ray source and nuclei produced by fragmentation reactions in interstellar space. In addition, the intensity and composition of the low-energy particles is observed to change on an 11-year cycle correlated with sunspot activity, thereby indicating that the

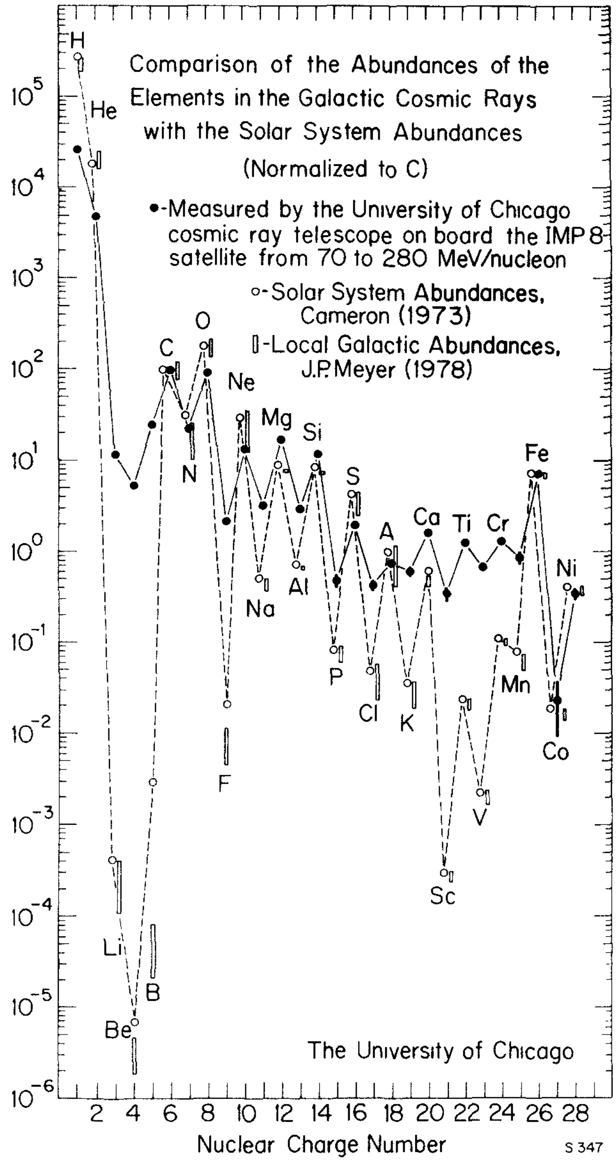


Fig. 1. Comparison of the abundances (relative to carbon) measured in the low-energy cosmic rays at earth with matter in or near the solar system.

cosmic rays at earth are affected by processes within the solar system.

A schematic representation of the complete cosmic-ray propagation problem is given in Fig. 2, and may be divided roughly into two parts: galactic propagation and solar modulation. After leaving their source regions as high-energy particles, the cosmic rays travel within the galaxy for about 20 million years[10] before arriving in the vicinity of the solar system. In that time some of the particles can escape from the confinement region (never to be seen at earth) while others encounter various amounts of interstellar matter. It is these encounters with the interstellar medium that produce spallation and also reduce the energy of the particles through ionization energy loss. Fortunately, the details of this galactic propagation can be deduced by studying the ratios of secondary (spallation produced) to primary elements, for example, B/C and V/Fe (see Fig. 1), using cross sections for the nuclear spallation reactions measured in accelerator laboratories. Such investigations determine not only the average amount of matter encountered by the cosmic rays but also the distribution of pathlengths followed by the particles. Armed with this information, the secondary (spallation) contribution can be unfolded, element by element, from the measured charge spectrum.

The difficulties of a cosmic ray are not over when it arrives in the vicinity of our solar system. The region around our sun, the heliosphere [indicated as extending out to 50 Astronomical Units (A.U.) on Fig. 2] is filled with plasma and magnetic fields which are carried out from the sun by the expanding solar wind. A cosmic ray must "swim upstream" a long way in order to reach the earth located at 1 A.U. In this solar modulation process the particles are decelerated, losing energy, and undergo diffusion among the magnetic field irregularities. Since these processes depend on the charge, mass, and energy of the particles, the composition measured at earth is affected. Solar modulation has been studied by measuring the cosmic-ray intensity at different locations within the heliosphere and by determining the energy spectrum of different components (e.g., protons and electrons).

The ultimate goal of the study of cosmic-ray

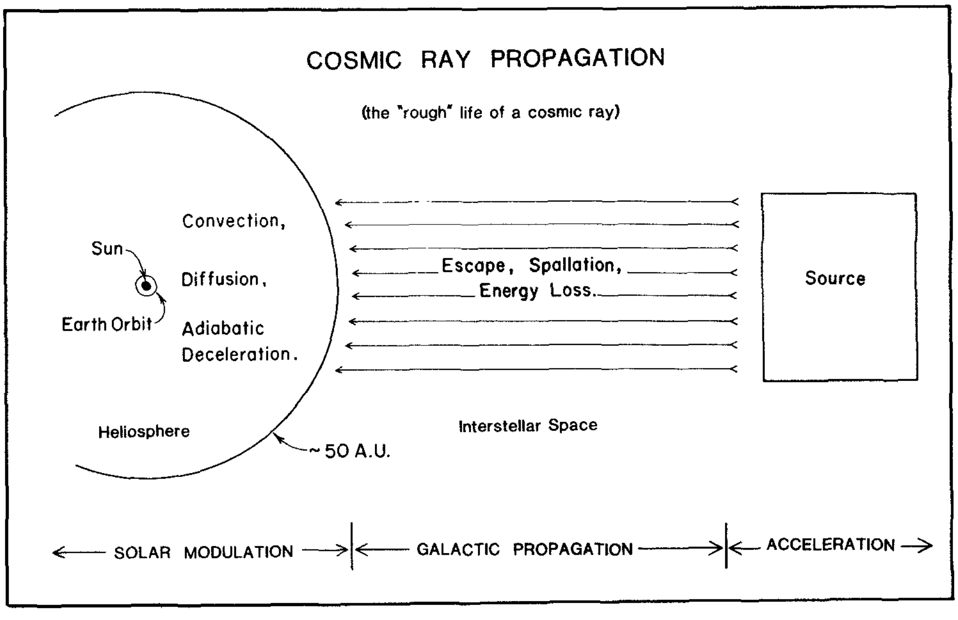


Fig. 2. A schematic representation of the history of a cosmic ray particle from acceleration to observation at earth.

composition is to determine the relative abundances of the matter at the source of the cosmic rays. This necessitates unfolding the effects of galactic propagation and solar modulation from measured abundances, such as on Fig. 1. Both solar modulation and galactic propagation are now understood sufficiently well so that their effects can be calculated, using large computer codes, thereby permitting the cosmic-ray source abundances to be extracted.[11]

What do these source abundances reveal? To answer this question, we take as a point of reference the distribution of elements in our solar system.[9] Each set of relative abundances (cosmic-ray source or solar system) must be normalized to the same element (usually carbon or silicon) before they can be compared. Once normalized, however, we can look at the ratio CRS/SS = Cosmic Ray Source Abundance divided by the Solar System Abundance. If the material at the cosmic-ray source is the same as matter in the solar system, then this ratio should be unity within the experimental uncertainties. For example, the oxygen to carbon ratio in solar system material is 1.7[9] while the same ratio for the cosmic-ray source matter is found (after unfolding propagation and modulation effects) to be 1.16 ± 0.02, where the uncertainty is derived from the cosmic-ray measurement errors plus uncertainties in unfolding the secondary contributions. For relative abundances normalized to carbon = 100, the oxygen abundance would be 170 in the solar system and 116 ± 2 in the cosmic-ray source. Thus, the ratio CRS/SS is 0.68 ± 0.01. In this example the element oxygen is clearly less abundant in the matter at the cosmic-ray source than it is in the solar system. Looking at other elements in a similar fashion reveals that the ratio CRS/SS varies considerably, with the cosmic-ray source matter (for a carbon normalization) overabundant in heavy elements such as silicon, calcium, or iron and underabundant in the lightest elements, hydrogen and helium. Thus at first look, the material at the cosmic-ray source appears to be very different from matter in our solar system!

It must be remembered that the cosmic-ray source abundances determined by unfolding galactic propagation and solar modulation effects represent the composition of the cosmic rays after their acceleration to high energy, and it is possible that the acceleration process can affect the relative composition. This possibility is addressed in Fig. 3 where the ratio CRS/SS (here the normalization is to silicon) is plotted as a function of the first ionization potential of the element. A distinctive correlation is evident. The elements with the lowest ionization potentials have the highest CRS/SS ratios while high ionization potential elements show low ratios. The simplest interpretation of such a correlation is that the cosmic rays originate from material very similar to solar system matter except that during the acceleration process the ease with which a given element is ionized determines the amount of that element which eventually is accelerated to become high-energy cosmic-ray particles. Note that the correlation is not all-encompassing. The elements hydrogen, neon, and sulphur seem to be anomalous, suggesting that this simple approach is not the complete answer.

Alternatively, the matter from which cosmic rays originate may be completely unlike solar system material, being instead rich in the heavy elements relative to the lighter species, and the correlation on Fig. 3 may have nothing to do with the acceleration process. This situation presents the cosmic-ray physicist with a dilemma since it is impossible to decide between these alternatives based solely on the element abundances. Of course, the actual physical situation in the galaxy might involve some combination of the two alternatives.

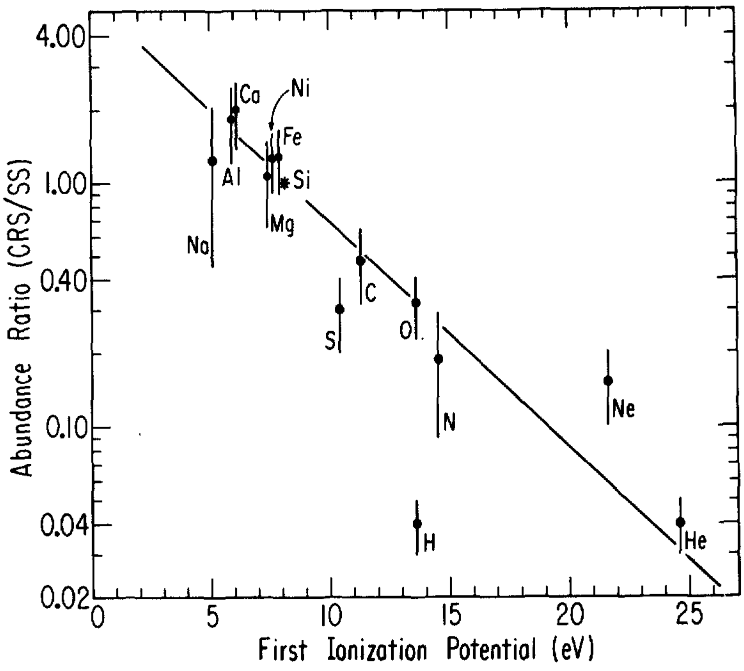


Fig. 3. Correlation between the ratio, cosmic-ray-source abundance/solar-system abundance, and the first ionization potential of the element.

A solution to this dilemma is offered by studying the isotopes of the primary cosmic-ray elements. Since the acceleration process, involving, for example, first ionization potential, appears to select on the basis of an atomic property of the element, it will treat two different isotopes of a given element identically and will not alter the relative distribution of the isotopes. Solar-system material is characterized not only by the relative abundances of the different elements but also by a distinctive pattern to the isotopes of each element, and often this isotopic signature is more precisely determined than the relative element abundances. Thus, if the isotopic composition of different elements in the cosmic-ray source matter can be determined, then comparison with the isotopic pattern of solar-system material should reveal whether or not these two samples of matter are the same. If the cosmic rays come from matter with a totally different nucleosynthetic history, the isotopic patterns will, in all probability, be quite different, and it is precisely these patterns that form the principal clues to the nature of the processes that formed this matter.

For isotope studies the effects of galactic propagation become extremely important. We learn nothing about the matter at the cosmic-ray source by studying the isotopic composition of an element formed principally as a secondary by spallation reactions in interstellar space. The isotopic pattern of such a secondary element is determined mainly by the spallation cross sections and not by the source matter. Thus, the isotopes of the elements lithium, beryllium, boron, and fluorine, for example, are essentially all secondary and provide information on galactic propagation but not on the composition of cosmic-ray source matter. Hydrogen, helium, carbon, and nitrogen each have two stable isotopes, but in each case one of them (deuterium, helium-3, carbon-13, nitrogen-15) is completely dominated by spallation production. A similar situation exists for oxygen with three stable isotopes of mass numbers 16, 17, and 18. Oxygen-16 is the principal isotope, and the two heavier isotopes are again dominated by secondary production. The next three even-charge elements, neon, magnesium, and silicon each have three stable isotopes (^{20,21,22}Ne, ^{24,25,26}Mg, ^{28,29,30}Si), and only one of these, ^{21}Ne, is mainly a spallation-produced isotope. Thus, it is not surprising that these three elements are the first to yield information on the isotopic composition of the matter at the cosmic-ray source.

This brings us to the newest era in cosmic-ray physics, beginning in the 1970s with the development of the experimental techniques necessary to measure the individual isotopes of the elements in the cosmic rays. This new phase is just beginning and will probably extend for the next several decades, based on the enormity of the job ahead, but it has already produced some fascinating experimental results which will be described in detail in the second part of this article. The interpretation of the cosmic-ray isotope measurements involves the processes of nucleosynthesis that form these elements, and the next section reviews the theory of heavy element formation in stars, with special emphasis on the elements neon, magnesium, and silicon.

Nucleosynthesis: The origin of heavy elements

Matter in the universe consists mainly of hydrogen and helium, with only a small component of "metals" (in astronomical parlance "metals" refers to the sum of all elements with $Z > 2$). The hydrogen and (most of) the helium are primordial elements, formed in the initial "big bang" at the beginning of the universe. The metals, however, are second-generation elements owing their existence largely to nucleosynthesis within stars.[12] This situation is illustrated by the existence in our galaxy of two distinct classes of stars; the stars of the galactic disk, population I, with a metals content of ~ 2% by mass and the older halo stars, population II, in which the metals content is at least 20 to 30 times smaller. This implies that the population I objects probably formed later, after the heavy element content of the galactic gas had increased due to nucleosynthesis in previous generations of stars, some of which still exist in the halo population. Indeed, the very "shining" of the stars in the sky requires a tremendous amount of energy which is provided by the nuclear fusion reactions whose products are heavy elements.

The evolution of stars and the formation of heavy elements is, qualitatively, quite simple.[13] Starting with an initial gravitationally bound "clump" of interstellar gas, the force of gravity contracts the gas and heats the interior until the temperature exceeds about a million degrees and hydrogen fusion begins. During this "hydrogen burning" stage, hydrogen is converted into helium either by the proton-proton cycle or, if the central temperature is high enough (~ 20 million degrees), by the carbon-nitrogen-oxygen (CNO) cycle.[14,15] In the proton-proton chain the principal reactions are:

$$^{1}\mathrm{H} + {}^{1}\mathrm{H} \rightarrow {}^{2}\mathrm{H} + e^{+} \qquad 1(a)$$

$$^{2}\mathrm{H} + {}^{1}\mathrm{H} \rightarrow {}^{3}\mathrm{He} \qquad 1(b)$$

$$^{3}\mathrm{He} + {}^{3}\mathrm{He} \rightarrow {}^{4}\mathrm{He} + 2\,{}^{1}\mathrm{H} \qquad 1(c)$$

for a net reaction in which four hydrogen atoms (^{1}H) are converted, through deuterium (^{2}H) and helium-3 (^{3}He), into one helium-4 nucleus (^{4}He). The quantity e^{+} denotes a positron from the beta decay accompanying deuterium production. In addition to this basic chain, there is an alternate reaction channel in which ^{3}He interacts with ^{4}He to form ^{7}Be which eventually is transformed (again, there are two alternate reaction channels) into two ^{4}He. Approximately 25 MeV of energy is released for each four hydrogen atoms converted into a helium atom. Our sun, with a central temperature of approximately 13 million degrees burns hydrogen by proton-proton reactions.

In stars more massive than the sun, the gravitational contraction produces higher internal temperatures ($\gtrsim$ 20 million degrees) which permit hydrogen burning by the CNO cycle, so named because atoms of carbon, nitrogen and oxygen catalyze the transformation of hydrogen into helium. The principal reactions in the CNO cycle are:

$$^{12}\mathrm{C} + {}^{1}\mathrm{H} \rightarrow {}^{13}\mathrm{N} \qquad 2(a)$$

$$^{13}\mathrm{N} \rightarrow {}^{13}\mathrm{C} + e^{+} \qquad 2(b)$$

$$^{13}\mathrm{C} + {}^{1}\mathrm{H} \rightarrow {}^{14}\mathrm{N} \qquad 2(c)$$

$$^{14}\mathrm{N} + {}^{1}\mathrm{H} \rightarrow {}^{15}\mathrm{O} \qquad 2(d)$$

$$^{15}\mathrm{O} \rightarrow {}^{15}\mathrm{N} + e^{+} \qquad 2(e)$$

$$^{15}\mathrm{N} + {}^{1}\mathrm{H} \rightarrow {}^{12}\mathrm{C} + {}^{4}\mathrm{He} \qquad 2(f)$$

where again the net change is the conversion of four hydrogen atoms into one helium with the release of about 25 MeV of energy. There is an alternative cycle, which will be important later in the discussion, in which reaction 2(f) is replaced by:

$$^{15}\mathrm{N} + {}^{1}\mathrm{H} \rightarrow {}^{16}\mathrm{O} \qquad 2(f)'$$

This is then followed by the reactions:

$$^{16}\mathrm{O} + {}^{1}\mathrm{H} \rightarrow {}^{17}\mathrm{F} \qquad 2(g)$$

$$^{17}\mathrm{F} \rightarrow {}^{17}\mathrm{O} + e^{+} \qquad 2(h)$$

$$^{17}\mathrm{O} + {}^{1}\mathrm{H} \rightarrow {}^{14}\mathrm{N} + {}^{4}\mathrm{He} \qquad 2(i)$$

The net effect of this alternate chain is to convert four hydrogen atoms into one helium *and* to convert an atom of ^{12}C into ^{14}N. This latter transformation is the basis for production of neutron-rich isotopes during later stages in the evolution of the star.

The conversion of hydrogen into helium, by either the proton-proton chain or the CNO cycle, provides the energy to both support the star against the force of gravity and to fuel the luminosity of the object – to make the stars shine! Stars spend most of their lifetime converting hydrogen to helium, about 10^{10} years for our sun. For larger stars, however, the hydrogen is burned much more rapidly, requiring only about 10^{6} years for a star 20 times the mass of the sun, and these large stars can complete their life cycle very rapidly compared to, for example, the age of our solar system.

When the hydrogen in the core is finally exhausted, the star contracts gravitationally until the central temperature is high enough (100 to 200 million degrees) to ignite helium fusion. This process converts helium into carbon,

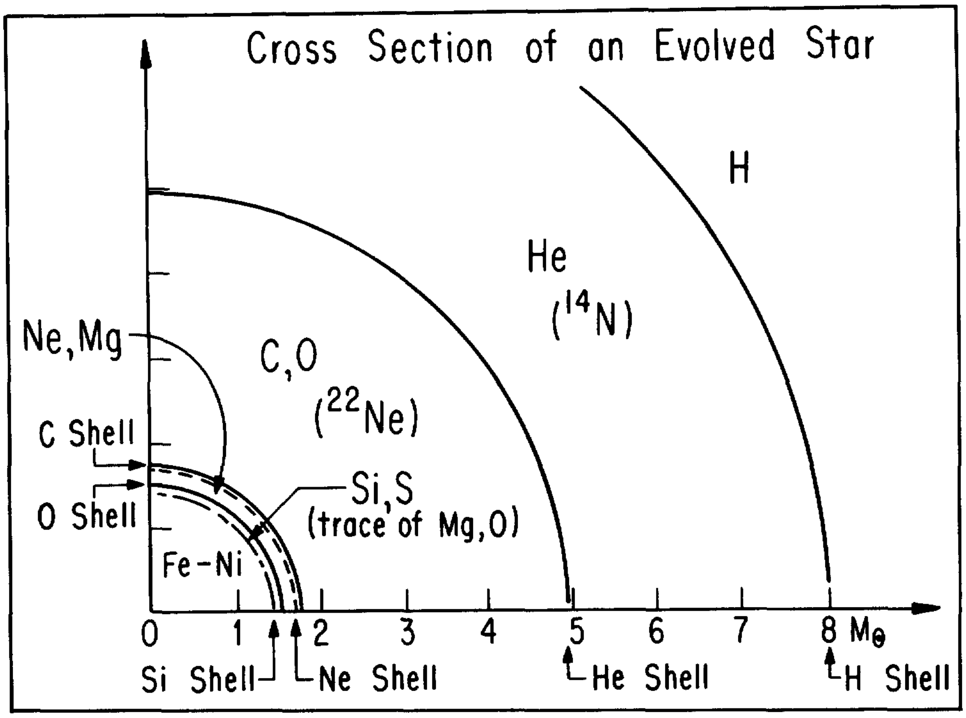


Fig. 4. The internal structure of a 22-solar-mass star at the end of its evolution prior to a supernova explosion (see text for details).

some of which also reacts with a helium nucleus to form oxygen. The energy generation from helium burning supports the star for a much smaller fraction of its lifetime (the red-giant phase), and when the helium is exhausted the star again begins to contract. At this stage there is a breakpoint in the evolution. In stars smaller than about six solar masses, the degenerate carbon-oxygen core can support itself against further gravitational collapse. These stars probably shed their excess mass via stellar winds or pulsations and eventually evolve to a white dwarf.

For the larger stars the gravitational contraction following the red-giant phase increases the core temperature to about 800 million degrees where carbon itself begins to burn forming neon and magnesium. Then follow successive contractions, temperature increases, and thermonuclear reactions which in turn burn neon to magnesium and some silicon, oxygen to silicon and sulphur, and finally silicon to iron and nickel. The nuclear burning at each stage commences in the core before the previous stage has completed burning out through the entire star, leaving shells that are undergoing each of the different nuclear-burning processes. These stars, therefore, have a layered structure, as illustrated in Fig. 4 for a star of 22 solar masses. The outermost parts of the star consist of unburned material, separated by a hydrogen-burning shell from the next region, the helium zone. Proceeding inward there are regions containing carbon and oxygen, neon, silicon, and finally the iron-nickel core. At this stage the nuclear evolution of the star is complete because the iron peak elements are the most tightly bound nuclei, and no further energy can be extracted by nuclear fusion.[16]

What happens next to this star? As expected, gravity again takes hold and the star contracts, squeezing the core to higher and higher density. Eventually an instability develops, and the core region explodes giving one of the most spectacular astronomical displays, a supernova. The explosion throws off the outer layers of the star, and the central core collapses to a compact remnant, a white dwarf, a neutron star, or possibly a black hole.[17] In the process of ejecting the outer layers, the material is heated, and some regions may attain high enough temperatures for additional nuclear reactions to occur, called "explosive nucleosynthesis," before the expansion cools the matter. Explosive nucleosynthesis augments the previous production of heavy elements and may be the principal source of many of the less abundant isotopes. With the exception of the core (~ 1.5 solar masses) which contracts to a compact remnant, all of the remaining material of the star is ejected into the surrounding medium where it is absorbed and mixed. Subsequent generations of stars forming from this interstellar medium will have an increased "metal" abundance. The largest stars return the most processed matter to the interstellar medium, but there are fewer of them. Taking a mass average over the stellar population gives a star of 20 to 25 solar masses (c.f. Fig. 4) as the "average" star for heavy element production.[18]

The basic nucleosynthesis in stars, following hydrogen burning, is one of alpha-particle addition. Helium burning forms ^{12}C and ^{16}O; carbon burning forms ^{20}Ne and ^{24}Mg; oxygen burning forms ^{28}Si and ^{32}S. Each of these products is a combination of alpha particles. For the cosmic-ray problem, however, it is necessary to understand the formation of both the major alpha-particle isotopes and the neutron-rich species (in particular, ^{22}Ne, ^{25,26}Mg, ^{29,30}Si). To accomplish this, it is necessary to look in more detail at the nuclear reactions occurring both before and after the supernova outburst.

As a by-product of hydrogen burning in massive stars via the CNO cycle, all of the carbon and oxygen atoms with which the star was formed originally are converted to

^{14}N, as given by equations 2(f)′ – 2(i). For a star with the composition of our sun this amounts to ~ 2% by mass. When helium burning commences, this ^{14}N also undergoes a transformation via the reactions,

$$^{14}N + He \rightarrow {}^{18}F \qquad (3a)$$

$$^{18}F \rightarrow {}^{18}O + e^{+} \qquad (3b)$$

$$^{18}O + He \rightarrow {}^{22}Ne \qquad (3c)$$

forming a substantial amount of ^{22}Ne, as indicated in parenthesis in Fig. 4. If the star is massive enough so that helium burning proceeds at high temperatures, some of the ^{22}Ne may be consumed during the helium burning phase by,

$$^{22}Ne + He \rightarrow {}^{25}Mg + n \qquad (4)$$

producing ^{25}Mg and, most important, a neutron. The important reactions (3) and (4) form the neutron source both for some production of neutron-rich isotopes of the elements below iron, such as $^{25}Mg + n \rightarrow {}^{26}Mg$ or $^{28}Si + n \rightarrow {}^{29}Si$, and for neutron captures by the iron-peak elements to form most of the elements in the periodic table heavier than nickel.

Some of the ^{22}Ne survives the evolution to be burned after the supernova outburst during explosive nucleosynthesis providing, again, a source of neutrons for neutron-capture reactions. It is these reactions, during explosive carbon or explosive neon-burning nucleosynthesis, that are commonly believed to be the main source of the neutron rich isotopes $^{25,26}Mg$ and $^{29,30}Si$, since the amount of these species formed during helium burning is too small to explain their natural abundance in the solar system. However, the real situation may not be this simple. Recent studies of convective shell burning (c.f. Fig. 4) at high temperatures, during pre-supernova evolution, have produced abundance patterns that "mimic" the results from explosive nucleosynthesis and thereby offer an alternate source for the heavy magnesium and silicon isotopes. In either case the basic reactions involve neutron capture from neutrons produced by burning ^{22}Ne by Eq. (4). Thus, the exact mix between nucleosynthesis before and after the supernova explosion in a massive star remains an open question whose resolution requires detailed calculations for realistic stellar models, a subject of much current research activity.

It is important to stress that the formation of ^{22}Ne, $^{25,26}Mg$, and $^{29,30}Si$ are coupled, but are independent, to first order, of the major isotopes ^{20}Ne, ^{24}Mg and ^{28}Si. For example, if the "metals" content of a massive star were substantially higher than the 2% mentioned above and all else being equal, following the evolution each of the ratios $^{22}Ne/^{20}Ne$, $^{25}Mg/^{24}Mg$, $^{26}Mg/^{24}Mg$, $^{29}Si/^{28}Si$ and $^{30}Si/^{28}Si$ would be increased by approximately the same multiplicative factor.[19]

Most of the evolution described in the preceeding paragraphs is relevant to stars larger than ~ 15 solar masses, those that return the most processed matter to the interstellar medium. However, the smaller stars, whose evolution takes much longer, can also be quite important. In stars around 8 solar masses a complicated series of thermal instabilities (flashes) in helium-burning shells permits ^{22}Ne synthesis by Eq. (3), but the temperatures remain low enough so that ^{22}Ne is not consumed via Eq. (4). Stars of this type may be a source of essentially pure ^{22}Ne! However, for slightly larger stars, the shell temperatures during thermal flashes can be high enough to burn some of the ^{22}Ne, and neutron-capture products are formed. These stars are thought to produce much of the solar system abundance of the elements heavier than nickel and would be expected to have some enhancement of the heavy magnesium and silicon isotopes along with a large amount of ^{22}Ne. Such stars, enriched in elements heavier than nickel, have been observed astronomically, and their study forms one of the best sources of information on the nuclear evolution of stars in this intermediate mass range. It is now apparent that the details of the evolution and the types of nucleosynthesis occurring in stars of different initial mass can be quite complicated, and the extent to which stars in one mass range may be important for synthesis of particular isotopes in the galaxy (or in the cosmic rays) remains, currently, an important area of investigation.

For the discussion of the isotopic composition of the cosmic rays in the second part of this article, the elements neon, magnesium, and silicon are the most important. The principle isotopes, ^{20}Ne, ^{24}Mg and ^{28}Si are formed during the normal course of evolution of most massive stars, as indicated in Fig. 4. The neutron-rich isotopes of these elements, however, require more specialized circumstances. The heavy magnesium and silicon isotopes, in particular, are formed principally in massive stars probably just before or during a supernova explosion. The one exception is ^{22}Ne, formed during helium burning in both high and intermediate mass stars, whose burning to release neutrons is the pivotal reaction on which the production of the other species depends.

Summary

The cosmic rays are indeed an interesting sample of galactic matter! The element abundances contain information on the source matter, the acceleration process, and the history of particle propagation in the galaxy. Unfortunately, this information cannot be unlocked completely due to the unknown nature of the acceleration process. However, the cosmic-ray isotopes provide the key for studying the matter at the cosmic-ray source. Is the nucleosynthetic history of this matter the same as matter in our solar system? This long-standing question is beginning to be answered due, at least partly, to the new research results in cosmic-ray physics.

The theory of nucleosynthesis for heavy elements during stellar evolution has been reviewed to provide a basis for the interpretation of the differences or similarities between the isotopic composition of cosmic-ray source matter and material from the solar system. Particular attention was devoted to the elements neon, magnesium, and silicon since these are the elements whose isotopes have been studied recently in the cosmic rays and provide the best current information on the composition of the cosmic ray source matter. However, this is the second part of the story which will be described in detail in the next part of this article, to appear in these pages next month.

Acknowledgments

This work was supported by the National Aeronautics and Space Administration under NASA Grant NGL 14-001-006 and Contract NAS 5-25731. The author would like particularly to thank J.A. Simpson, D. N.Schramm, W. D. Arnett and M. G. Munoz for their help and encouragement and T. G. Guzik for a critical review of the manuscript and also to acknowledge the aid of many others, too numerous to mention, who contributed data, ideas, and useful discussions.

References

1. A. M. Hilas, *Cosmic Rays* (Pergamon, Oxford, 1972), pp. 3-70.
2. J. R. Winckler and D. J. Hofmann, Am. J. Phys. **35**, 1 (1967).
3. B. Rossi, *Cosmic Rays* (McGraw-Hill, New York, 1964), pp. 1-268.
4. B. Rossi, *High Energy Particles* (Prentice-Hall, Englewood Cliffs, 1952), pp. 388-527.
5. C. F. Powell, P. H. Fowler, and D. H. Perkins, *The Study of Elementary Particles by the Photographic Method* (Pergamon, New York, 1959), pp. 1-669.
6. E. P. Ney, Sci. Am. **184**, 26 (1951).
7. F. B. McDonald and C. E. Fichtel, Eds., *High Energy Particles and Quanta in Astrophysics* (MIT Press, Cambridge, MA, 1974), pp. 1-170.
8. M. Garcia-Munoz and J. A. Simpson, Proc. 16th International Conference on Cosmic Rays (Kyoto, Japan) **1**, 270 (1979).
9. A. G. W. Cameron, Sp. Sci. Rev. **15**, 121 (1973); also J. P. Meyer, *Les Elements et leurs Isotopes dans l'Univers* (U. of Liege Press, Liege, France, 1979) pp. 153-489. The sources of the data used in these two compilations of solar system abundances are nearly, but not completely, the same. The differences in adopted values demonstrate that some uncertainty remains in assessing relative abundances within our "well-studied" solar system! For the remainder of this paper the Cameron compilation will be adopted.
10. M. Garcia-Munoz, G. M. Mason and J. A. Simpson, Ap. J. **217**, 859 (1977).
11. M. M. Shapiro and R. Silberberg, Ann. Rev. Nucl. Sci. **20**, 323 (1970).
12. W. A. Fowler and W. E. Stephens, Am. J. Phys. **36**, 1 (1968). This excellent bibliography contains references to all of the major work in stellar evolution, abundance determinations, nuclear reaction rates and nucleosynthesis up to about the end of 1967.
13. D. D. Clayton, *Principles of Stellar Evolution and Nucleosynthesis* (McGraw-Hill, New York, 1968), pp. 362-606. For a more general treatment see also D. D. Clayton, Phys. Today, **22**, 28 (1969).
14. G. Wallerstein, Science **162**, 625 (1968).
15. A. Unsold, *The New Cosmos* (Springer-Verlag, New York, 1969), pp. 250-259.
16. D. Halliday, *Introductory Nuclear Physics* (Wiley, New York, 1955), pp. 261-263.
17. R. Ruffini and J. A. Wheeler, Phys. Today **24**, 30 (January, 1971).
18. D. N. Schramm and W. D. Arnett, *Mercury*, p. 16 (May/June, 1975).
19. S. E. Woosley and T. A. Weaver, Ap. J. **243**, 651 (1981).

PRACTICAL BIOPHYSICS

With great conscientiousness Maxwell repeated Cavendish's experiments. Something of the background of electrical science as Cavendish knew it will be recovered if we remember that in his day there was no known effect of an electric current by which measurements could be made. Cavendish was driven to the dire expedient of passing the current through his own body and estimating its magnitude by the intensity of the resulting shock! According to Sir Arthur Schuster the necessary apparatus was set up in the laboratory and all visitors were required to submit themselves to the ordeal of impersonating a galvanometer. On one occasion a young American astronomer expressed his severe disappointment that after traveling to Cambridge on purpose to meet Maxwell and consult him on some astronomical topic he was almost compelled to take his coat off, plunge his hands into basins of water and submit himself to a series of electrical shocks!

Alexander Wood
The Cavendish Laboratory

REVIEWS

The dawn of the particle astronomy era in ultra-high-energy cosmic rays

Pablo M. Bauleo[1]* & Julio Rodríguez Martino[2,3]*

Cosmic rays are charged particles arriving at the Earth from space. Those at the highest energies are particularly interesting because the physical processes that could create or accelerate them are at the limit of our present knowledge. They also open the window to particle astronomy, as the magnetic fields along their paths are not strong enough to deflect their trajectories much from a straight line. The Pierre Auger Observatory is the largest cosmic-ray detector on Earth, and as such is beginning to resolve past observational disagreements regarding the origin and propagation of these particles.

In 1912, after a series of balloon flights, Hess discovered a penetrating radiation that originated in outer space. Years later, in 1926, Millikan called this radiation 'cosmic rays'. The name has survived since then, generally referring to charged particles impinging on the Earth's atmosphere. In the late 1930s, Auger and his group measured coincident signals generated by detectors separated by distances of more than a few hundred metres[1,2]: they concluded that these signals were caused by an 'extensive air-shower' (EAS) of charged particles. Auger and his co-workers assumed that the air-shower was originated by a single photon, high in the atmosphere, and used the recently developed quantum electrodynamics theory to estimate its energy, which they found to be in excess of 10^{15} electron volts (eV). Figure 1 shows a schematic representation of an EAS.

Cosmic rays of energies larger than about 10^{13} eV are small in number, and so can only be detected through the secondary particles produced when they enter Earth's atmosphere. The EAS starts with the interaction of a cosmic ray with a nucleus in the upper atmosphere. All the available energy is distributed among the secondary particles—of which there can be billions if the primary energy is above 5×10^{17} eV—that can spread over several tens of square kilometres at ground level. Two methods are mainly used to register these particle cascades. The particle density can be sampled at the ground using an array of detectors; alternatively, the shower path can be tracked through the atmosphere, collecting the fluorescence light induced by electrons in the atmospheric nitrogen molecules.

Here we review the developments in ultra-high-energy cosmic ray (UHECR) physics over the past 15 years: we cover the controversy about the existence of the theoretically predicted suppression of the cosmic-ray energy spectrum, and its later confirmation. The most relevant topic is the discovery that the arrival direction of the most energetic cosmic rays follows the distribution of nearby extragalactic objects. This implies that their origin is not cosmological, but instead they are accelerated inside extragalactic objects, by some still unclear physical process. Three large experimental facilities—AGASA (Akeno Giant Air Shower Array), HiRes (High Resolution Fly's Eye) and the Pierre Auger Observatory—have already started what will eventually become a new era in astronomy. In the near future, further observations and more accurate instruments will identify the cosmic-ray acceleration sites and will lead to the study of the energy spectrum of individual sources. This, combined with the study of the attenuation of cosmic rays through space, could give valuable information on the cosmic microwave background. The deflection produced on the cosmic-ray path by Galactic and extragalactic magnetic fields will be an indirect tool to measure their strength. In addition, accurate measurements of the interaction of cosmic rays with the Earth's atmosphere will hint at the particle physics interaction models, at an energy range beyond what can be achieved in

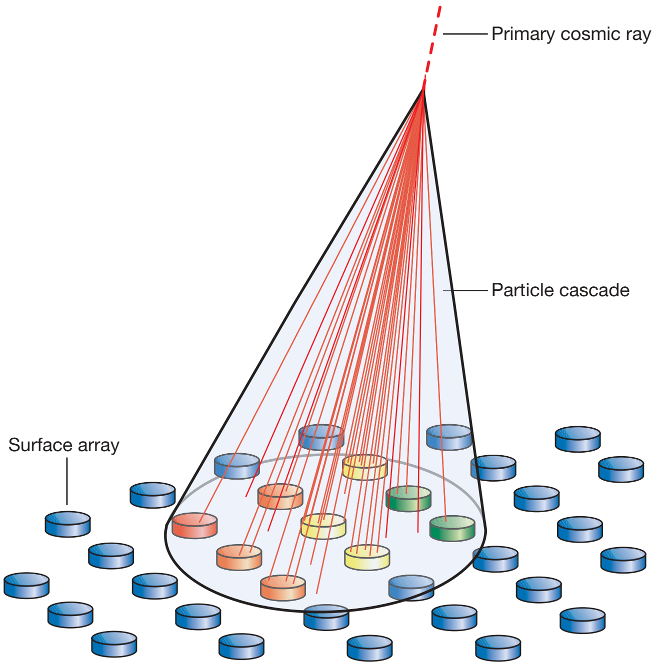


Figure 1 | Scheme of an extensive air-shower. The primary cosmic ray (dashed line) undergoes a nuclear interaction in the upper atmosphere (typically 20 km above sea level), producing a cascade of elementary particles (represented as solid red lines within a conical shape). These particles propagate across the atmosphere and could reach ground level. The cascade footprint at the ground could be of tens of square kilometres. A network of particle detectors at ground level (surface array) can detect the arrival of the particles, allowing reconstruction of the whole cascade. Different colours in the scheme represent different arrival times of the particles.

---

[1]Physics Department, Colorado State University, Fort Collins, Colorado 80523, USA. [2]Dipartimento di Fisica e Astronomia, Università di Catania, [3]INFN Sezione di Catania, I-95123 Catania, Italy.
*These authors contributed equally to this work.

human-made accelerators. These observations are within reach of the current and next-generation observatories, and will herald the dawn of the era of charged-particle astronomy.

Properties of cosmic rays

The observed cosmic-ray energy spectrum spans from 10^8 eV to more than 10^{20} eV. Particles with energies lower than 10^{10} eV mainly come from the Sun, as the solar wind prevents particles in that energy range from reaching the Earth from outside the Solar System. For energies higher than 10^{18} eV, a convincing explanation of the acceleration processes and sources is still unknown. Some theories suggest that these cosmic rays originate in stellar winds within our Galaxy and later accelerate in supernova shocks or similar high-energy environments[3]. Active galactic nuclei (AGN), galaxies with very intense emission in a broad wavelength range, are possible source candidates of UHECRs above 10^{19} eV (ref. 4), but so far there are only experimental hints suggesting this.

The cosmic-ray flux follows a power law ($E^{-\gamma}$) as a function of energy E, with an approximate index $\gamma = 3$. This index value remains remarkably constant, showing only small variations across the whole measured cosmic-ray energy spectrum.

At the highest energies, above 10^{20} eV, the estimated number of particles is only a few per km^2 per millennium. This extremely low flux calls for the construction of huge observatories, covering a very large area with detectors. For instance, the Pierre Auger Southern Observatory in Argentina covers 3,000 km^2, which is about 30 times the size of the district of Paris.

Cosmic rays with energies above 4×10^{19} eV cannot travel through space without being attenuated[5,6]. Propagation is mainly affected by the presence of the cosmic microwave background radiation, consisting of photons with a black-body radiation distribution corresponding to an equivalent temperature of 2.7 K. In the rest frame of an extremely energetic proton, these low-energy photons are seen as very high energy photons (γ-rays), of about 10^8 eV. If the photon energy in the rest frame is above 150 MeV, pion-production reactions become possible. The proton loses energy in each reaction, reducing the mean distance it can travel undisturbed to about 50 Mpc. This effect produces a dip in the spectrum, known as the 'GZK suppression' (named after Greisen, Zatsepin and Kuzmin, who predicted its existence), and it reduces the number of high energy particles able to arrive at Earth, if originated at larger distances.

The HiRes observatory data suggested the presence of suppression in the flux of cosmic rays in the GZK energy region[7], whereas the AGASA collaboration announced that the cosmic-ray spectrum continued, with a power law dependence, above GZK energies[8]. This last result was revisited a few years later, without being able to arrive at a definite conclusion owing to the limited number of events in the GZK energy region, even though the existing, limited data collected by AGASA is still being re-analysed[9,10].

The Pierre Auger Southern Observatory data seem to agree with the HiRes result in the GZK energy region[10], resolving the controversy between the two previous experimental results. The observation of the GZK suppression[11] is another interesting result. The larger data set of the Auger Observatory made it also possible to establish a correlation between some high energy events and AGN (or any other astronomical objects that follow the same spatial distribution) closer than 75 Mpc to the Earth[12]. For protons with energy larger than 6×10^{19} eV, the magnetic deflection of the trajectory of the cosmic rays is only a few degrees[12], hence enabling the possibility of particle astronomy. This small deflection would imply that the particles 'point back' to their sources, making it possible to identify the origin of cosmic rays and even study the spectra of individual sources. By studying the distribution of cosmic-ray arrival directions (such as clustering, thread-like structures, and so on), it would be possible to analyse the properties of Galactic and inter-galactic magnetic fields.

The AGASA observatory

AGASA[13] was located at Akeno, Japan. It ran in full operation mode from 1993 to 2004, being able to take data continuously, independently of weather conditions. Each ground station, composed of plastic scintillators, sampled the secondary particles of an EAS. The trigger time of each individual station was used in the reconstruction of the EAS arrival direction, while the energy measurement was based on the number of particles at each station.

The energy of a shower detected using a ground array is not measured directly. The particle density at a given distance from the EAS axis is correlated with the energy of the primary cosmic ray through computer simulations. The models used in the simulations are based on the knowledge about interactions acquired in particle accelerator experiments. This means that the models extrapolate the physical processes to several orders of magnitude in energy beyond what has been measured until now. One of the Large Hadron Collider experiments (LHCf) will be dedicated to reducing the uncertainty in hadron interaction models of cosmic-ray showers, by measuring the forward particle production in proton interactions[14]. Apart from this, the computational effort of producing and tracking about 10^{11} particles is too large to be practical. Hence, only a statistically representative sample of the EAS secondary particles is propagated to the ground in the simulation. All these facts lead to an energy measurement that is strongly model dependent, and to large uncertainties in its value.

The HiRes observatory

The HiRes Fly's Eye[15] was located in Dugway, Utah, USA. HiRes commissioned its first location in 1997 and its second location in 1999. Both locations were decommissioned in 2006. This observatory collected fluorescence light induced in the atmosphere by the passing EAS. The total brightness of an EAS, in fluorescence light, averages a few watts. The amount of light collected is so faint that these detectors can only operate on clear, moonless nights. Typical observation duty factors of fluorescence detectors lie between 10% and 15%.

Each of the HiRes locations had mirrors that focused the fluorescence photons into a light sensor array, or 'camera'. It is conceptually similar to a CCD camera, with each light sensor playing the role of one pixel. By considering the relative trigger times and geometric pattern of the pixels in the camera, it is possible to reconstruct the arrival direction of the shower. The energy is calculated by integrating the total amount of light measured at the detector location. The total number of photons induced in the atmosphere by the EAS is proportional to the total available energy, that is, the energy of the primary cosmic ray. Some particles in the cascade do not induce fluorescence light and the total energy of the EAS must be corrected to account for this fact. The atmospheric conditions are other factors to include when estimating the primary cosmic-ray energy. An atmospheric attenuation correction, based on the distance from the EAS to the detector, needs to be applied.

Discrepant results

The limited sample of cosmic rays in the GZK energy region, together with intrinsic differences in the way each experiment measured the cosmic-ray energy, set the stage for a controversial difference between the measured spectra.

A comparison of both measured spectra is shown in Fig. 2, where the discrepancy is clear. The AGASA data seemed to favour the absence of a suppression, while HiRes spectrum followed the expected curve. Both results should be interpreted carefully, as the calculations involved are not straightforward and, again, the number of detected events was not enough to firmly establish either claim[16].

When computing a cosmic-ray spectrum, it is critical to calculate the instrument exposure, or time-integrated collection area. In the case of AGASA, the exposure is reduced to the convolution of the detector array geometrical area, the acceptance solid angle and the effective running time. The acceptance of a surface array, like AGASA, becomes

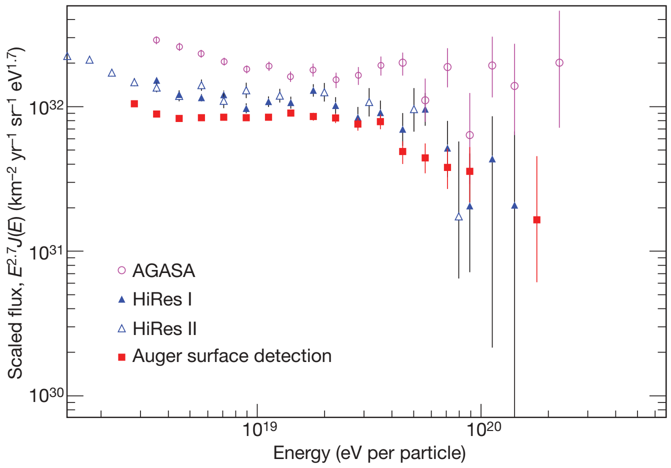


Figure 2 | UHECR data from different experiments. Comparison between AGASA (circles), HiRes monocular spectra (open and filled triangles correspond to each HiRes location) and Pierre Auger Southern Observatory. Error bars are 1σ.

constant with energy, once its trigger efficiency reaches 100%. The cumulative exposure for this detector is about 1,600 km^2 sr yr (ref. 17).

A fluorescence detector requires a more complicated exposure calculation. The collection volume is a hemisphere (centred at the detector location), the radius of which indicates the maximum observation distance for a given EAS. This distance changes with the atmospheric conditions (atmospheric aerosols, cloud coverage, position of the clouds) and depends on the EAS energy. The acceptance of a fluorescence detector, like HiRes, is a function of the EAS energy (the brighter the EAS, the further away it can be detected). This implies that, in order to calculate the energy spectrum of an EAS, it is necessary to accurately know how the instrument acceptance depends on the EAS energy and the atmospheric conditions at the time of measurement. It is very difficult to deduce the exposure of HiRes from the published results, but it is quoted as "more than twice that of AGASA above the GZK-threshold"[7].

In any case, both these experiments have statistically limited data samples, given the extremely low cosmic-ray flux at those energies[16]. In response to the AGASA results, numerous speculations about how cosmic rays could avoid energy loss on their way to Earth were proposed. Either new particles[18,19] or interactions with magnetic fields[20] were invoked to avoid the problem. These articles are just a small sample of a long list showing different (and sometimes quite ingenious) arguments.

Besides measuring the energy spectrum, both experiments analysed the arrival direction distributions of cosmic rays. An '*n*-plet' is defined as a set of *n* independent events whose arrival directions are the same, within experimental uncertainties. The AGASA collaboration found 5 doublets and 1 triplet[21] in their data sample, where only 2 doublets were expected statistically. This result was not confirmed by HiRes[22]. On the other hand, correlations were found in the HiRes sample with the locations of BL Lacertae objects (AGN with their jets pointing towards Earth)[23], although they have not been confirmed by an independent data sample. It should be remarked that anisotropy in the arrival direction of cosmic rays is not expected at lower energies. However, at higher energies—combining data from different observatories—an excess of events coming from the supergalactic plane (a plane defined by the locations of the galaxies in the local cluster) was found for events with energies above 4×10^{19} eV, giving a hint that the origin of UHECRs is most likely to be extragalactic[24]. This result was independently suggested later by analysing the shape of the cosmic-ray spectrum[25,26].

The relatively low exposures of these experiments could only provide hints about the arrival direction of the cosmic rays, making it possible to search for clustering and sources, but not to confirm them. Still, these results were extremely important in that they showed anisotropy studies (and potentially the identification of cosmic-ray sources) to be within reach.

The Pierre Auger Observatory

The Pierre Auger Southern Observatory[27], schematically shown in Fig. 3, is located in the province of Mendoza, Argentina. It covers an approximate area of 3,000 km^2, which makes it the largest cosmic-ray observatory to date. Its northern counterpart will be built in the vicinity of Lamar, Colorado, USA. When finished, the joint instruments will have full sky coverage as observed from both hemispheres. The Southern Observatory has been collecting cosmic-ray data since 2004, while increasing its size up to the installation of the last surface detector on June 2008. As of 31 August 2007, the accumulated exposure of the Southern Observatory is 9,000 km^2 sr yr (ref. 28). The yearly accumulated exposure is about 6,000 km^2 sr yr and the observatory is expected to operate for a total of 20 years.

This observatory combines the techniques used in previous experiments, by means of a 'hybrid detector', that is, having a fluorescence detector and an array of surface detectors working together. The fluorescence detector follows the shower cascade across the atmosphere and the surface detector array—in this case water Cherenkov detectors—detects the particles on arrival at ground level. Hybrid measurements can set an absolute energy scale, improve the determination of the primary particle type and give better energy and angular resolution[27]. This approach provides a model-independent energy calculation, using the fluorescence detector data together with the simple surface array aperture calculation.

In hybrid mode, for any given EAS measured simultaneously by both instruments, the energy deposited in the atmosphere—as recorded by the fluorescence detector—is then related to a surface detector energy parameter. Then, this model-independent correlation can be used as energy calibration for events measured only with the surface detector array[29].

The Pierre Auger Collaboration is taking advantage of the unique characteristics of the observatory. Although the limit is arbitrary, EAS detected by water Cherenkov arrays are typically reconstructed only up to 60°. Auger Collaboration members have developed analysis techniques to extend the acceptance up to 75°, which increases the

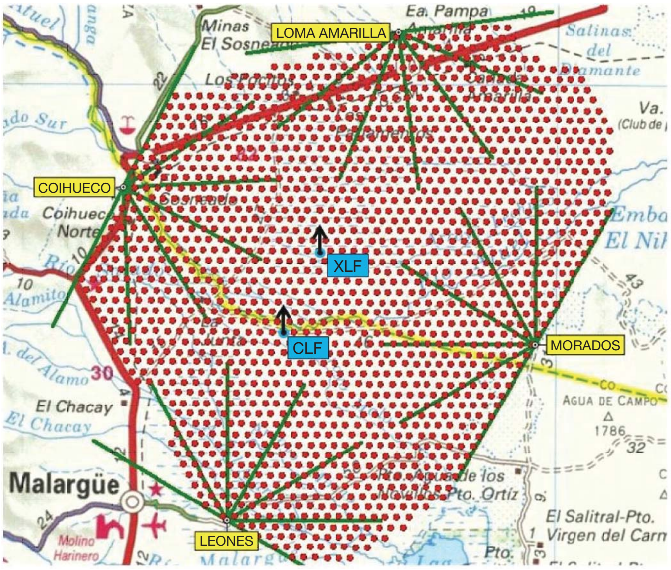


Figure 3 | The Pierre Auger Southern Observatory. It consists of an array of 1,600 surface detectors (red filled circles), complemented by 24 fluorescence detectors, grouped in four buildings (yellow labels; Leones, Morados, Loma Amarilla and Coihueco). Green lines represent the field of view of each detector. Two laser facilities (blue labels; CLF and XLF) are available for energy calibration and atmospheric monitoring. Observatory information is superimposed on a map of the area.

instrument exposure even further[30]. As a comparison, plastic scintillator arrays (like AGASA, or the Telescope Array in Utah, see below) typically limit their reconstructed events up to 45°.

Evidence for the GZK suppression

Results recently published by the Auger Collaboration[11] report the existence of a deficit of cosmic rays at the highest energies. Still, this result alone is not enough as to prove that the GZK suppression has been observed. It could be that the energy spectrum is limited by the maximum energy available at the cosmic-ray acceleration sites.

When the evidence on the deficit in the flux of cosmic rays is put together with the energy at which the correlation with nearby extragalactic objects[28] sets in, one could then argue that the GZK suppression has been observed. If objects beyond an approximate distance of 75 Mpc were to be included in the analysis, the correlation would very rapidly diminish.

Although both HiRes and the Pierre Auger Southern Observatory have observed a suppression in the cosmic-ray flux above an energy of approximately 4×10^{19} eV, differences still exist in the measured spectrum index and the overall energy normalization. The energy scales of these two observatories differ by about 17% (ref. 31).

The sources

One of the main questions to be answered regarding UHECRs is how these particles can be accelerated to such energies. The size of the acceleration region and the magnetic field present in it must follow a relation, usually represented in a Hillas plot like that shown in Fig. 4. Only a few astrophysical objects could then be potential sources.

Arguably, the most relevant recent observation has been the discovery of a correlation between cosmic-ray arrival directions and nearby extragalactic objects[12,28]. The correlation found in the Pierre Auger Southern Observatory data becomes significant for cosmic rays above 5.7×10^{19} eV and AGN within 75 Mpc. With those parameters, 20 events (out of a total of 27) lie within 3.1° from an object listed in the Veron-Cetty-Veron catalogue[32].

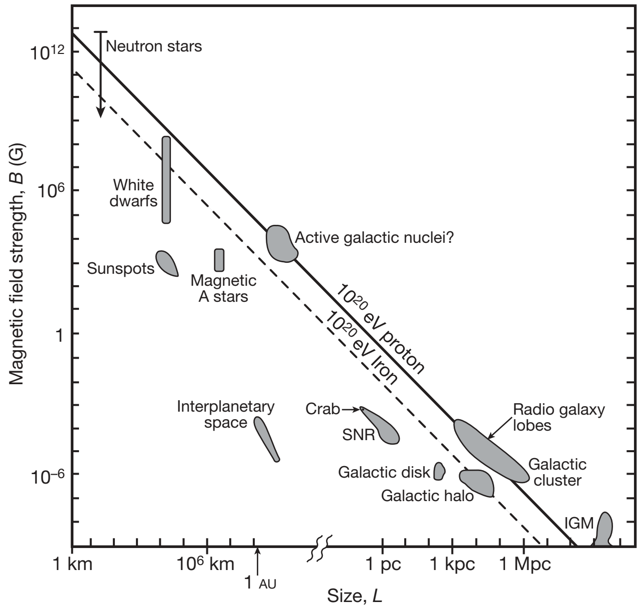


Figure 4 | Hillas diagram. Non-exotic acceleration processes require a particle to be confined within a region (of size L) where magnetic field shocks are present (with a field intensity value of B). Once the particle reaches its maximum energy, then the magnetic field is not able to keep the particle confined within the acceleration region and the particle escapes. This gives an approximate value for the maximum achievable energy of $E_{\max} = BL$, shown as a solid/dashed line for a 10^{20} eV proton/iron nuclei, respectively. We show data for a variety of astrophysical objects; only those above the line can accelerate particles to energies into the GZK region. Crab indicates the Crab nebula; SNR, supernova remnant, IGM, intergalactic magnetic field.

AGN have traditionally been considered as possible candidates for cosmic-ray acceleration sites. However, any other astrophysical object close enough to Earth to avoid the GZK suppression, with a spatial distribution similar enough to that of AGN, could be the source.

The AGN hypothesis seems to be supported by the correlation found between the arrival direction of cosmic rays reported by the Auger Collaboration[12] and the positions of the Swift hard X-ray catalogue of AGN, when weighted by the X-ray flux and constrained to distances less 100 Mpc (ref. 33). At the same time, using the same events measured by the Pierre Auger Southern Observatory, a correlation was also found with the HIPASS catalogue of H I spiral galaxies (when weighted by their H I flux)[34]. The latter results do not contradict the correlation found with AGN, as all these objects trace the distribution of matter. The hypothesis of H I galaxies as cosmic-ray sources is interesting, as it would explain the lack of events from the Virgo cluster (which is not rich in H I galaxies).

HiRes members have searched their data for correlations[35] based on the Pierre Auger Southern Observatory parameters, and their analysis does not support the result published by the Auger Collaboration. Reference 31 shows that if corrected by the energy mismatch between both experiments, HiRes would have only 5 events in their stereo data sample, which might not be enough as to establish or reject any correlation.

Open questions

Despite having measured a suppression in the spectrum compatible with the GZK suppression and arrival direction anisotropies (or perhaps because of those facts), some exciting and intriguing questions still remain to be solved.

Sources and acceleration models. Nearby extragalactic objects have been found to correlate with the arrival direction of cosmic rays, but it is not yet possible to study the energy spectrum of individual sources. Such a spectrum would lead to a better understanding of acceleration processes at the sources. At the same time, the search for other potential sources should continue. Cosmic rays could be generated by different astrophysical objects.

Energy spectrum. The GZK suppression is produced by the interaction of nucleons with photons, at energies higher than 4×10^{19} eV. At energies higher than 3×10^{20} eV, the interactions become much less probable. Hence, cosmic rays with those energies could propagate almost undisturbed through space, allowing the study of the Universe at extreme energies. This feature, predicted by quantum physics, is known as the 'GZK recovery'. Observing it would prove quantum physics at an energy range that has not been explored before. The lack of a GZK recovery could imply new physics.

Mass composition and particle physics. A very important point to be studied is the mass composition of cosmic rays. This will either prove or reject different acceleration and propagation models, which favour either light or heavy primary particles. Moreover, at these high energies, cosmic-ray interactions with atoms in the upper atmosphere are in the range of a few hundred TeV (in the centre of mass frame). Studies of shower development in the atmosphere (known as elongation rate) will give an opportunity to unveil features of hadronic interactions at these energies, which are more than one order of magnitude higher than those achievable by the Large Hadron Collider, the most powerful human-made particle accelerator[36].

Magnetic fields. Magnetic fields could be studied by looking at the arrival direction pattern of cosmic rays as a function of energy. If 'strings' of events were identified, their relative deviation at different energies would allow us to set limits (or possibly even measure) the strength of Galactic and extragalactic magnetic fields.

A larger set of events, measured with good resolution, will answer several questions. As it is true for so many scientific disciplines, the main problem to be solved regarding the study of UHECRs is obtaining a significantly larger number of events.

The future

The Pierre Auger Southern Observatory is the largest-aperture observatory currently taking data and its exposure is larger than that of any previous detector. There is a proposal to increase the size of the projected Northern Observatory to cover an area 7 times larger than that of the Southern Observatory.

The Telescope Array in Utah, a hybrid instrument combining a surface scintillator array and fluorescence detectors, is the only observatory in the Northern Hemisphere currently taking data in this energy regime. Its yearly cumulative exposure will depend on the final operation conditions, but it could be estimated to be about 1,400 km^2 sr yr (ref. 37).

New techniques and observation methods are being considered. The collection of fluorescence light with space-based instruments, looking down into the Earth's atmosphere, has been proposed. JEM-EUSO[38] and OWL[39] are examples of this technique. Radio-wave detection of EAS is also currently being developed[40].

The past decade has proven fruitful and exciting in cosmic-ray physics. We have witnessed revisions and improvements in the instrumental techniques, which in turn have paid off by establishing the existence of the GZK suppression and by the discovery of anisotropies in the cosmic-ray arrival directions. In cosmic-ray physics, discoveries have been achieved by seeking the largest exposure possible. History has shown us that in this field, exposure matters.

In the near future, within 4 years or so, the Pierre Auger Southern Observatory should have observed about 100 events above $\sim 5 \times 10^{19}$ eV. In contrast, the proposed Pierre Auger Northern Observatory could be collecting the above-mentioned statistics every 9 months. In 20 years of combined operation, about 2,000 events (above $\sim 5 \times 10^{19}$ eV) could have been observed. Such data from the Northern and Southern Observatories could be used to accurately search for point sources, to study the energy spectra of different sources, and to understand Galactic and extragalactic magnetic fields, as well as to investigate and perhaps uncover particle physics beyond accelerator energies. A new window to the Universe has been opened; we are witnessing the dawn of the particle astronomy era.

1. Auger, P. & Maze, R. Extensive cosmic showers in the atmosphere. *C.R. Acad. Sci. II* **207**, 228–229 (1938).
2. Auger, P., Ehrenfest, P., Maze, R., Daudin, J. & Robley, A. F. Extensive cosmic-ray showers. *Rev. Mod. Phys.* **11**, 288–291 (1939).
3. Biermann, P. & Sigl, G. Introduction to cosmic rays. *Lect. Notes Phys.* **576**, 1–26 (2001).
4. Biermann, P., Gaisser, T. & Stanev, T. The origin of galactic cosmic rays. *Phys. Rev. D* **51**, 3450–3454 (1995).
5. Greisen, K. End to the cosmic ray spectrum? *Phys. Rev. Lett.* **16**, 748–750 (1966).
6. Zatsepin, G. T. & Kuzmin, V. A. Upper limit of the spectrum of cosmic rays. *Zh. Eksp. Teor. Fiz. Pisma Red.* **4**, 114–117 (1966).
7. Abbasi, R. *et al.* First observation of the Greisen-Zatsepin-Kuzmin suppression. *Phys. Rev. Lett.* **100**, 101101 (2008).
   First observation of the cosmic-ray flux suppression at the highest energies.
8. Sakaki, N. *et al.* in *Proc. 27th Int. Cosmic Ray Conf. (ICRC 2001)* (eds Kampert, K. H., Hainzelmann, G. & Spierling G.) 333–336 (Copernicus Gesellschaft, 2001).
9. Sokolsky, P. & Thomson, G. B. Highest energy cosmic rays and results from the HiRes experiment. Preprint at ⟨http://arXiv.org/abs/0706.1248v1⟩ (2007).
10. Watson, A. Recent results from the Pierre Auger Observatory – Including comparisons with data from AGASA and HiRes. *Nucl. Instrum. Meth. A* **588**, 221–226 (2008).
11. Abraham, J. *et al.* Observation of the suppression of the flux of cosmic rays above 4×10^{19}eV. *Phys. Rev. Lett.* **101**, 061101 (2008).
    Confirmation of the cosmic-ray flux suppression at the highest energies.
12. Abraham, J. *et al.* Correlation of the highest energy cosmic rays with nearby extragalactic objects. *Science* **318**, 938–943 (2007).
    Discovery of cosmic-ray arrival direction anisotropy.
13. Chiba, N. *et al.* Akeno giant air shower array (AGASA) covering 100-km^2 area. *Nucl. Instrum. Meth. A* **311**, 338–349 (1992).
14. Pilkington, A, *et al.* in *Proc. XVI Int. Workshop on Deep-Inelastic Scattering and Related Topics (DIS 2008)* (eds Devenish, R. & Ferrando, J.) doi:10.3360/dis.2008.50 (Science Wise Publishing, 2008).
15. Abu-Zayyad, T. *et al.* The prototype high-resolution Fly's Eye cosmic ray detector. *Nucl. Instrum. Meth. A* **450**, 253–269 (2000).
16. De Marco, D., Blasi, P. & Olinto, A. V. On the statistical significance of the GZK feature in the spectrum of ultra high energy cosmic rays. *Astropart. Phys.* **20**, 53–65 (2003).
17. Takeda, M. *et al.* Energy determination in the Akeno giant air shower array experiment. *Astropart. Phys.* **19**, 447–462 (2003).
18. Kachelriess, M., Semikoz, D. V. & Tortola, M. A. New hadrons as ultra-high energy cosmic rays. *Phys. Rev. D* **68**, 043005 (2003).
19. Gibilisco, M. Monopolonium decay as a source of ultrahigh energy cosmic rays. *Nucl. Phys. Proc.* **100** (Suppl.), 357–359 (2001).
20. Deligny, O., Letessier-Selvon, A. & Parizot, E. Magnetic horizons of UHECR sources and the GZK feature. *Astropart. Phys.* **21**, 609–615 (2004).
21. Teshima, M. *et al.* in *Proc. 28th Int. Cosmic Ray Conf. (ICRC 2003)* (eds Kajita, T., Asaoka, Y., Kawachi, A., Matsubara, Y. & Sasaki, M.) 437–440 (Universal Academy, 2003).
22. Bellido, J. *et al.* in *Proc. 28th Int. Cosmic Ray Conf. (ICRC 2003)* (eds Kajita, T., Asaoka, Y., Kawachi, A., Matsubara, Y. & Sasaki, M.) 425–428 (Universal Academy, 2003).
23. Abbasi, R. U. *et al.* Search for cross-correlations of ultra-high-energy cosmic rays with BL Lacertae objects. *Astrophys. J.* **636**, 680–684 (2006).
24. Uchihori, Y. *et al.* Cluster analysis of extremely high energy cosmic rays in the northern sky. *Astropart. Phys.* **13**, 151–160 (2000).
25. Berezinsky, V. On transition from galactic to extragalactic cosmic rays. *J. Phys. Conf. Ser.* **47**, 142–153 (2006).
26. Stanev, T. The transition from galactic to extragalactic cosmic rays. *Nucl. Phys. Proc.* **168** (Suppl.), 252–257 (2007).
27. Abraham, J. *et al.* Properties and performance of the prototype instrument for the Pierre Auger Observatory. *Nucl. Instrum. Meth. A* **523**, 50–95 (2004).
28. Abraham, J. *et al.* Correlation of the highest-energy cosmic rays with the positions of nearby active galactic nuclei. *Astropart. Phys.* **29**, 188–204 (2008).
29. Sommers, P. in *Proc. 29th Int. Cosmic Ray Conf. (ICRC 2005)* Vol. 7 (eds Sripathi Acharya, B. *et al.*) 387–390 (Tata Institute of Fundamental Research, 2005).
30. Newton, D. in *Proc. 30th Int. Cosmic Ray Conf. (ICRC 2007)* (eds Caballero, R. *et al.*) 323–325 (Universidad Nacional Autonoma de Mexico, 2007).
31. Kampert, K. H. Ultra high-energy cosmic ray observations. *J. Phys. Conf. Ser.* **120**, 062002 (2008).
32. Veron-Cetty, M. *et al.* A catalogue of quasars and active nuclei: 12th edition. *Astron. Astrophys.* **455**, 773–776 (2006).
33. George, M. *et al.* On active galactic nuclei as sources of ultra-high energy cosmic rays. *Mon. Not. R. Astron. Soc.* **388**, L59–L63 (2008).
34. Ghisellini, G. *et al.* Ultra-high energy cosmic rays, spiral galaxies and magnetars. Preprint at ⟨http://arxiv.org/abs/0806.2393⟩ (2008).
35. Abbassi, R. U. *et al.* Search for correlations between HiRes stereo events and active galactic nuclei. Preprint at ⟨http://arXiv.org/abs/0804.0382⟩ (2008).
36. LHC. 2008. ⟨http://lhc2008.web.cern.ch/lhc2008⟩ (2008).
37. Kawai, H. *et al.* Telescope array experiment. *Nucl. Phys. Proc.* **175–176** (Suppl.), 221–226 (2008).
38. Ebisuzaki, T. *et al.* The JEM-EUSO project: Observing extremely high energy cosmic rays and neutrinos from the International Space Station. *Nucl. Phys. Proc.* **175–176** (Suppl.), 237–240 (2008).
39. Orbiting wide-angle light-collectors (OWL). ⟨http://owl.gsfc.nasa.gov⟩ (2004).
40. Falcke, H. *et al.* Detection and imaging of atmospheric radio flashes from cosmic ray air showers. *Nature* **435**, 313–316 (2005).

Acknowledgements We express our gratitude to many people, too many to mention individually, in the AMANDA and Auger Collaborations, who guided us through the cosmic-ray world over the years. We especially thank our mentor A. Filevich, whom introduced us to cosmic-ray physics and the scientific world.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence should be addressed to P.M.B. (bauleo@lamar.colostate.edu).

the amounts of a nutritionally responsive transcriptional activator Gcn4, and demonstrated that this is required for full lifespan extension from dietary restriction[9]. Similarly, autophagy must be induced for lifespan to be extended by dietary restriction in *C. elegans*[10].

On the basis of these studies, it is tempting to speculate that rapamycin may be functioning as a dietary-restriction mimetic — a small molecule that provides the benefits of dietary restriction without requiring a reduction in food intake. Like dietary restriction, TOR inhibition not only increases lifespan, but also confers protection in invertebrate and rodent models against age-associated disorders, including cardiovascular dysfunction, diet-induced obesity and cancer[7]. Cancer inhibition in particular is a hallmark of dietary restriction in rodents, and rapamycin analogues are already used clinically as a treatment for certain forms of cancer.

Despite these links, Harrison *et al.*[1] do not strongly favour the idea that rapamycin is mimicking dietary restriction in mice. This is based on their data that rapamycin extends lifespan without reducing body weight, and when treatment is initiated during middle age (late-life onset of dietary restriction has shown inconsistent effects on longevity in previous studies). It is worth pointing out, however, that a true dietary-restriction mimetic may not reduce body weight if it mimics the signalling events (and downstream responses) associated with dietary restriction without changing food consumption. Also, dietary restriction has not yet been extensively characterized in mice of the genetically diverse background used by Harrison *et al.*, so it is difficult to predict whether dietary restriction in these animals would have effects similar to rapamycin. Thus, although it is premature to say for certain that rapamycin is functioning as a dietary-restriction mimetic in mice, the known role of TOR in the nutrient response, and the genetic relationship between TOR signalling and dietary restriction in invertebrates, make this a reasonable possibility.

Is this the first step towards an anti-ageing drug for people? Certainly, healthy individuals should not consider taking rapamycin to slow ageing — the potential immunosuppressive effects of this compound alone are sufficient to caution against this. On the basis of animal models, however, it is interesting to consider that rapamycin — or more sophisticated strategies to inhibit TOR signalling — might prove useful in combating many age-associated disorders. Also, as relevant downstream targets of TOR are better characterized, it may be possible to develop pharmacological strategies that provide the health and longevity benefits without unwanted side effects. So, although extending human lifespan with a pill remains the purview of science-fiction writers for now, the results of Harrison *et al.*[1] provide a reason for optimism that, even during middle age, there's still time to change the road you're on. ■

Matt Kaeberlein and Brian K. Kennedy are in the Departments of Pathology and Biochemistry, University of Washington, Seattle, Washington 98195, USA.
e-mail: kaeber@u.washington.edu

1. Harrison, D. E. *et al. Nature* **460,** 392–395 (2009).
2. Jia, K., Chen, D. & Riddle, D. L. *Development* **131,** 3897–3906 (2004).
3. Kaeberlein, M. *et al. Science* **310,** 1193–1196 (2005).
4. Kapahi, P. *et al. Curr. Biol.* **14,** 885–890 (2004).
5. Powers, R. W. 3rd, Kaeberlein, M., Caldwell, S. D., Kennedy, B. K. & Fields, S. *Genes Dev.* **20,** 174–184 (2006).
6. Vellai, T. *et al. Nature* **426,** 620 (2003).
7. Stanfel, M. N., Shamieh, L. S., Kaeberlein, M. & Kennedy, B. K. *Biochim. Biophys. Acta* advance online publication doi:10.1016/j.bbagen.2009.06.007 (2009).
8. Smith, E. D. *et al. Genome Res.* **18,** 564–570 (2008).
9. Steffen, K. K. *et al. Cell* **133,** 292–302 (2008).
10. Hansen, M. *et al. PLoS Genet.* **4,** e24 (2008).

Competing financial interests: declared (see online article for details).

ATMOSPHERIC PHYSICS

Cosmic rays, clouds and climate

Ken Carslaw

Galactic cosmic rays could influence Earth's cloudiness by creating aerosol particles that prompt cloud formation. That possible effect looks to be smaller than thought, but the story won't end there.

Striking correlations have been observed between Earth's cloud cover and the flux of galactic cosmic rays entering our atmosphere. The decrease in galactic cosmic ray (GCR) flux by about 15% over much of the twentieth century has led to the hypothesis that GCRs could influence climate through their effect on cloudiness. This controversial possibility is revisited in a paper in *Geophysical Research Letters* by Pierce and Adams[1].

There are several plausible mechanisms that could link GCR flux and cloud properties[2]. A leading candidate is the 'ion–aerosol clear-air mechanism', in which atmospheric ions created by GCRs act as nuclei for the formation of atmospheric particles. The nucleation of new nanometre-sized aerosol particles is observed frequently, and in many parts of the atmosphere, and is thought to be a major source of cloud-condensation nuclei (CCN) — particles large enough for cloud droplets to form around them. The link between GCRs and climate is therefore plausible because any change in GCR-ionization rate might be expected to drive changes in cloud-droplet concentrations, and hence the amount of solar radiation that clouds reflect back to space.

Atmospheric ions can indeed seed new particles[3], but two outstanding questions have hampered progress. What fraction of nuclei is created this way? And what fraction of these particles grows large enough to influence CCN? To be relevant to recent climate change, it would be necessary to show that the decrease in GCR flux during the twentieth century could lead to significant changes in CCN and clouds.

In their paper[1], Pierce and Adams estimate the magnitude of the ion–aerosol clear-air mechanism. They used a global atmospheric model with a detailed treatment of aerosol physics to estimate some limiting values of CCN formation from changes in GCR flux. Their conclusion is clear: CCN concentrations just aren't very sensitive to the changes in GCRs that have occurred during the twentieth century. The authors predict that CCN concentrations will change by less than 0.1% between solar maxima and minima as GCRs change by 15% — about the same as the change seen during the last century. They estimate that this change in CCN translates into a change of 0.005 watts per square metre in solar radiation reflected from clouds, insignificant compared with the greenhouse-gas warming of 2 watts per square metre or more over roughly the same period.

Pierce and Adams's model is quite sophisticated in the way it treats the global lifecycle of aerosols, from formation at nanometre sizes to their eventual growth over days to weeks to CCN sizes. But rather than trying to model the complex ion–aerosol processes in detail (physics that is still incompletely understood), they make an upper-limit assumption that all nucleation is due to ions, thereby circumventing one obstacle to making such a global assessment.

Is this negative result the last word on the ion–aerosol clear-air mechanism? Climate modellers are always quick to point out that predictions can be model-dependent. Certainly CCN may be more sensitive to the ion-induced nucleation rate in a different model or under conditions not explored by Pierce and Adams. But other global-model studies[4,5] of nucleation suggest that CCN are fairly insensitive to the nucleation rate for a simple reason: during the time taken for nuclei to grow to CCN sizes, coagulation depletes particle concentrations — just as raindrops are always fewer in number than cloud drops. Unless there is some as-yet-undiscovered process that accelerates the growth of a few charged nuclei all the way up to CCN sizes, this low sensitivity is likely to be a robust conclusion.

Despite this result[1], it is likely that a cosmic-ray–cloud–climate connection will continue to be explored, for two reasons. First, scientists continue to be intrigued by correlations between cosmic rays, Earth's electrical state and climate variables (clouds, precipitation, drought and so on) on timescales from hours to millennia[6,7]. Because the climate displays a

C. DAVE/SPL

Light on stratocumulus: GCR-cloud physics will remain a hot topic.

multitude of cycles on almost all timescales, detection of a correlation among climate variables usually meets with initial and healthy scepticism. But variations in cloud properties observed on timescales that are unique to GCRs[8] will always prompt a hunt for a plausible mechanism.

The second reason that GCR–cloud physics will remain a hot topic is that we have yet to explore all the possible mechanisms. Attention may now shift to the 'ion–aerosol near-cloud' mechanism[2]. GCR ionization modulates the fair-weather conduction current (about 2 picoamps per square metre) flowing between the ionosphere and Earth, thereby altering the charge that has been observed to accumulate around cloud layers. Just like static electricity, this charge can influence how cloud drops stick to aerosol particles. If the particles are effective nuclei for ice formation, then GCRs may influence cloud glaciation and precipitation. And the charge on some aerosol particles in the near-cloud environment could possibly become large enough to influence the formation of cloud drops directly[9]. But our understanding of the relevant physics is incomplete, and it will be some time before global-impact investigations along the lines of Pierce and Adams's study can be made.

Some would argue that the link between cosmic rays and climate is just too tenuous to be worth pursuing. Others would point out that, by ignoring the fact that the atmosphere is actually a dilute plasma (that is, is weakly ionized), we are missing some potentially important cloud physics — and clouds are a very large lever by which to influence climate. Despite the controversy, it is clear that the study of cosmic rays in our climate system has come of age. Sophisticated models of ion–aerosol processes now exist. They are supported by observations and laboratory studies, which will include the upcoming CLOUD experiment at the CERN laboratory near Geneva, Switzerland, in which a proton beam will generate highly controllable ionization events in an aerosol–cloud chamber[10]. ■

Ken Carslaw is in the School of Earth and Environment, University of Leeds, Leeds LS2 9JT, UK.
e-mail: k.s.carslaw@leeds.ac.uk

1. Pierce, J. R. & Adams, P. J. *Geophys. Res. Lett.* doi:10.1029/2009GL037946 (2009).
2. Carslaw, K. S., Harrison, R. G. & Kirkby, J. *Science* **298,** 1732–1737 (2002).
3. Curtius, J., Lovejoy, E. R. & Froyd, K. D. *Space Sci. Rev.* **125,** 159–167 (2006).
4. Spracklen, D. V. *et al. Geophys. Res. Lett.* doi:10.1029/2007GL033038 (2008).
5. Pierce, J. R. & Adams, P. J. *Atmos. Chem. Phys.* **7,** 1367–1379 (2007).
6. Harrison, R. G. & Ambaum, M. H. P. *Env. Res. Lett.* **4,** 014003 (2009).
7. Kirkby, J. *Surv. Geophys.* **28,** 333–375 (2007).
8. Harrison, R. G. *Proc. R. Soc. A* **464,** 2575–2590 (2008).
9. Harrison, R. G. & Ambaum, M. H. P. *Proc. R. Soc. A* **464,** 2561–2573 (2008).
10. http://www.cloud-itn.uni-frankfurt.de/

GALAXY FORMATION

Anatomy of elliptical galaxies

Luca Ciotti

The family of elliptical galaxies is remarkable for the structural regularity of its members. Inspecting irregularities in this regularity could help in understanding how these galaxies form.

One of the most-debated subjects in modern astrophysics is how elliptical galaxies, which are among the oldest known objects in the Universe, formed. Among the various likely formation mechanisms, merging is the most popular. According to this theory, different galaxies are the aftermath of merger events between progenitors of different morphologies and of varying encounter geometries. But observations indicate that there is room for other mechanisms. Despite great endeavour in trying to match the regularities observed in the structures of elliptical galaxies with theoretical models, there is still no consensus view of how they formed. Writing in *The Astrophysical Journal Supplement Series*, Kormendy and colleagues[1] report a meticulous study of all known elliptical galaxies in the Virgo cluster (one of the clusters of galaxies nearest to Earth) that investigates how departures from the observed regularities can be diagnostic of the processes that triggered the formation of these galaxies (Fig. 1, overleaf).

The most striking property of elliptical galaxies is that their brightness profiles — that is, the way in which the combined luminosity of their stars varies with distance from the centre — depend in a regular way on their total luminosity (Sérsic's law). Other properties of elliptical galaxies that correlate with their total brightness include size, mean star velocity and metal content. Another trait shared by these stellar systems is a supermassive black hole, with a mass of the order of one-thousandth of the galaxy's stellar mass, at their centre[2].

In their study of the Virgo cluster of galaxies, Kormendy *et al.*[1] report galaxy luminosity profiles over large radial ranges and argue that elliptical galaxies are not arranged as a continuous sequence of objects with properties that scale well with their total luminosity. Instead, elliptical galaxies seem to branch out into two families according to a threshold value for the total luminosity. This dichotomy manifests itself in two kinds of departure from the Sérsic law at small radii. Luminous ellipticals have 'cuspy' cores — that is, their luminosity profiles are characterized by 'missing light' at small radii, because their brightness at such radii drops below the Sérsic-fitted, larger-radii profile. By contrast, less-luminous ellipticals are all 'coreless' — their central luminosity profiles seem to have 'extra light' at small radii (but see Graham *et al.*[3] for a different interpretation of the central-light profiles).

Kormendy and colleagues' results add weight to other observations that have hinted at a dichotomy in the properties of elliptical galaxies. Luminous-core galaxies are known to be slowly rotating; to be relatively anisotropic (properties such as stellar velocities depend on direction); to have triaxial shapes (they have different diameters in all three directions); to have quite 'steep' Sérsic profiles; and to have stars that are mostly very old and that formed on comparatively short timescales. Conversely, low-luminosity coreless ellipticals rotate rapidly; are more isotropic; have mostly oblate-spheroidal shapes; have quite