Radiation Belts of the Earth

Dr. James A. Van Allen
Chairman of Physics and Astronomy, University of Iowa

The radiation belts of the Earth were discovered in 1958. The nature of these belts, their causes, and the restrictions that they impose on the safe conduct of space flights are described here.

By 1956 it was expected that it would soon be possible to use military rockets to boost small, scientifically instrumented satellites into durable orbits around the Earth. The enormous improvement that such satellites would provide in cosmic ray and auroral studies led me to propose appropriate instruments for early satellite missions. George Ludwig, then a graduate student, and I devised and built such a radiation detecting instrument at the University of Iowa as part of the International Geophysical Year Program (1957–58). This instrument was carried on the first U.S. satellite Explorer I (launched Jan. 31, 1958). Our similar but more elaborate instrument was carried on Explorer III (launched March 26, 1958). At altitudes below about 400 km (240 mi.) we confirmed earlier measurements of cosmic ray intensity that we had obtained from 1946 to 1957. But as the satellites swung out to higher altitudes in their elliptical orbits, we encountered an enormous increase in intensity, hundreds of times what we expected. We soon convinced ourselves that we had discovered a new phenomenon, namely an enormous population of energetic, electrically charged particles trapped in the Earth's magnetic field (Fig. 1) in the manner visualized by the Norwegian mathematician Carl Störmer in 1907. The geometrical distribution of such particles was found to have the form of a huge doughnut encircling the Earth, with the Earth itself occupying the center of the hole. This distribution soon acquired the name "radiation belt." Later in 1958 with Explorer IV and Pioneer III, we found that there are two distinctly different radiation belts of the Earth.

Electrons of energies up to several million electron volts and protons of energies up to several hundred million electron volts are present in these belts. In the heart of the outer belt, some 100 million electrons of energy greater than 40 kilo-electron volts cross an area of a square centimeter per second.

These early studies have been greatly extended and improved during the two subsequent decades. The radiation belts are only a portion of a much larger and very complex system, the "magnetosphere" (pronounced mag-neat-o-sphere). This is the region around the Earth within which its external magnetic field dominates the motion of electricially charged particles; i.e., it is the "sphere of influence" of the Earth's magnetic field.

The residence times of individual particles range from the order of a day in the outer belt to many years in the inner belt. The principal source of the particles is the "solar wind," a hot ionized gas that is emitted by the Sun and flows outward throughout the Solar System. A small fraction of this gas (principally protons and electrons) is captured in the outer reaches of the magnetosphere. Once captured, the particles are raised and lowered in energy in a random manner by fluctuating magnetic and electric fields caused by the varying solar wind. Particles which, on the average, gain energy diffuse inward into stronger field regions nearer to the Earth and are eventually lost in the atmosphere. Others diffuse outward and are lost from the system. The entire situation is a dynamic one.

There is another weak but important source of quite a different nature that supplies the most energetic (order of 100 million electron volts) protons in the inner radiation belt. Cosmic ray collisions with the nuclei of atoms of gas in the upper atmosphere result in a wide variety of secondary particles, among which are neutrons. Some of these neutrons fly outward without being affected by the magnetic field (because they are electrically neutral). A free neutron has a slightly greater mass than a proton and is radioactive with a mean life of 1013 seconds. Hence, a minute fraction of these secondary (or albedo) neutrons decays into protons, electrons, and neutrinos. Protons and electrons, both being electrically charged, then execute corkscrew-shaped Störmerian orbits in the magnetic field, and a few of them are durably trapped to form one component of the radiation belt. The injection rate by this process is only about 1 particle per cubic centimeter per 10 million years, but the residence time near the Earth is long enough that an important intensity of such particles builds up. There is also evidence for the direct capture of solar energetic particles of energies of the order of 10 million electron volts.

It appears that the dominant source of most of the low energy (less than 10 kiloelectron volts) trapped particles (or plasma) is gas from the ionized layers of the upper atmosphere, called the ionosphere.

The total power required to maintain all known magnetospheric processes is about one percent of that potentially available from the solar wind.

Major features of the magnetosphere are sketched in *(Fig. 2)*. The high latitude boundaries of the outer radiation belt are virtually identical to the low latitude boundaries of the auroral zone. In addition, large fluctuations in the aurora are closely correlated with fluctuations in the population and distribution of particles in the outer belt. These and other facts demonstrate that the aurora and the outer radiation belt are closely related through a common cause, namely fluctuations in the flow of the solar wind, but that neither is the cause of the other.

An important practical aspect of the radiation belts is that they impose strict limitations on the regions where prolonged flight of astronauts can be conducted safely. The corridor for safe orbital flight is below an altitude of about 400 km (250 mi.). By contrast, during a 20-hour period in an equatorial orbit at an altitude of 2500 km (1550 mi.), the heart of the inner belt, the cumulative radiation exposure of an astronaut inside a space shuttle's cabin would produce serious radiation sickness. A flight of several days duration there would be fatal. Even electronic equipment has a limited lifetime because of radiation damage, despite the fact that such equipment can tolerate much greater exposure to radiation then can living organisms.

In principle, the radiation can be absorbed and rendered harmless by heavy external shields, but the mass of adequate shields is impractical in the present state of space flight technology. Rapid traversals of the radiation belts, as in the Apollo missions to and from the Moon, result in a total dosage of radiation that can be readily tolerated by an astronaut. Satellites for communications and other purposes have necessary orbital radii of about 42,164 km (26,200 mi.) in order that their orbital periods are synchronous with the rotational period of the Earth. Fortunately, radiation at this altitude is of much lower intensity than is that in the inner zone and is also much more easily absorbed by shielding. Finally, it may be noted that a series of U.S. and Soviet atomic bomb bursts at high altitudes from 1958 to 1962 showed that it is possible to produce artificial radiation belts of enormous intensity. Such bursts are now forbidden by international treaty but remain a military threat to space flight around the earth. ■

Suggested Additional Readings

Eather, Robert H. 1980. *Majestic Lights, The Aurora in Science, History and the Arts.* American Geophysical Union. Washington, D.C.

Beatty, J. K., Chaikin, A., and O'Leary, B., eds. 1981. *The New Solar System.* Cambridge, Mass.: Sky Publishing Corp.

Van Allen, James A. 1959. Radiation Belts Around the Earth. *Scientific American* Vol. 200, No. 3: 39-47.

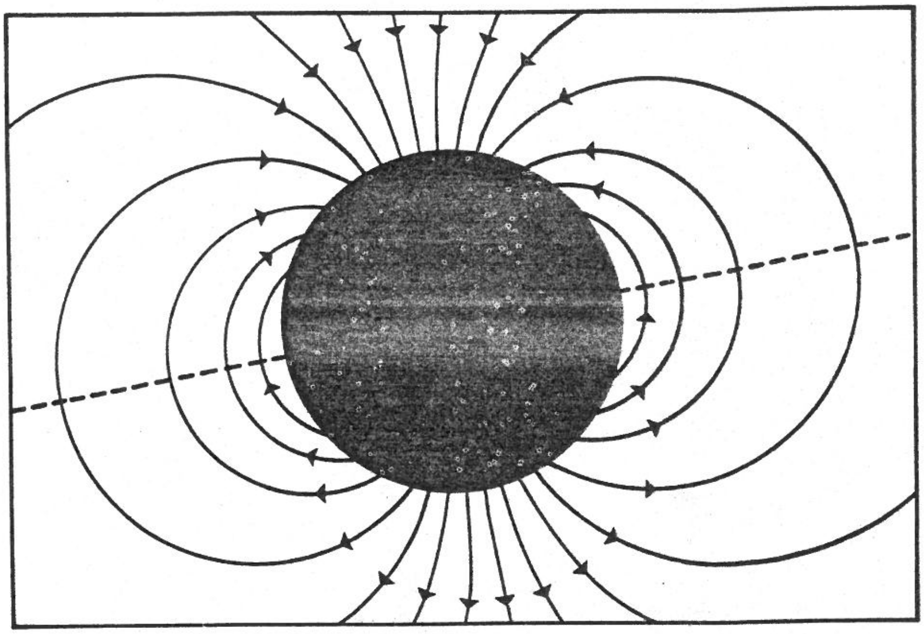

Fig. 1 The black circle represents a meridian cross section of the Earth, and the curved lines represent magnetic lines of force in the region external to the Earth if the solar wind were not present. The broken line is the magnetic equator, tilted at 11° to the geographic equator.

Fig. 2 This is a schematic view of a meridian cross section of the Earth's magnetosphere as actually observed. The Earth is represented by the small circle near the center of the diagram. Major features of the magnetosphere are labeled.