

Radiation Belts around the Earth

Instruments borne aloft by artificial satellites and lunar probes indicate that our planet is encircled by two zones of high-energy particles, against which space travelers will have to be shielded

by James A. Van Allen

So far, the most interesting and least expected result of man's exploration of the immediate vicinity of the earth is the discovery that our planet is ringed by a region—to be exact, two regions—of high-energy radiation extending many thousands of miles into space. The discovery is of course troubling to astronauts; somehow the human body will have to be shielded from this radiation, even on a rapid transit through the region. But geophysicists, astrophysicists, solar astronomers and cosmic-ray physicists are enthralled by the fresh implications of these findings. The configuration of the region and the radiation it contains bespeak a major physical phenomenon involving cosmic rays and solar corpuscles in the vicinity of the earth. This enormous reservoir of charged particles plays a still-unexplained role as middleman in the interaction of earth and sun which is reflected in magnetic storms, in the airglow and in the beautiful displays of the aurora.

The story of the investigation goes back to 1952 and 1953, before any of us could think realistically about the use of earth satellites to explore the environment of the earth. Parties from our laboratory at the State University of Iowa spent the summers of those years aboard Coast Guard and naval vessels, cruising along a 1,500-mile line from the waters of Baffin Bay, near the magnetic pole in the far northwestern corner of Greenland, southward to the North Atlantic off the coast of Newfoundland. Along the way we launched a series of rocket-

carrying balloons—"rockoons." (The balloon lifts a small rocket to an altitude of 12 to 15 miles, whence the rocket carries a modest payload of instruments to a height of 60 to 70 miles.) Our objective was to develop a profile of the cosmic-ray intensities at high altitudes and latitudes, and thus to learn the nature of the low-energy cosmic rays which at lower altitudes and latitudes are deflected by the earth's magnetic field or absorbed in the atmosphere.

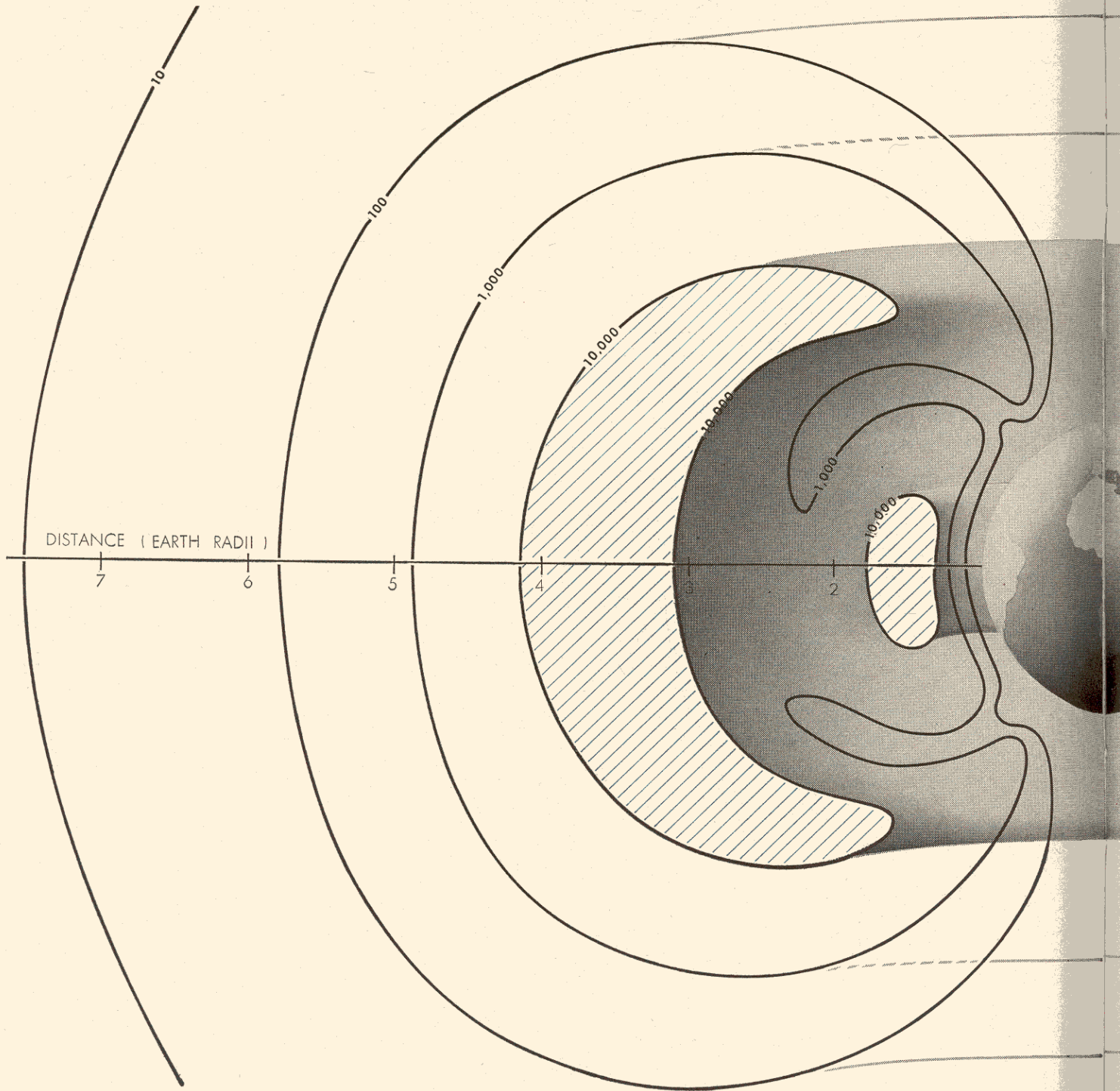
Most of the readings radioed down from the rockets were in accord with plausible expectations. Two rockoons sent aloft in 1953, however, provided us with a puzzle. Launched near Newfoundland by Melvin Gottlieb and Leslie Meredith, they encountered a zone of radiation beginning at an altitude of 30 miles that was far stronger than we had expected. At first we were uneasy about the proper operation of our instruments. But critical examination of the data convinced us that we had unquestionably encountered something new in the upper atmosphere.

Significantly these measurements were made in the northern auroral zone. In this zone, which forms a ring some 23 degrees south of the north geomagnetic pole, the incidence of visible auroras reaches its maximum. Since rockets fired north and south of the zone had revealed nothing unusual, we speculated that the strong radiation played some part in the aurora. Showers of particles from the sun, it was thought, come plunging into the atmosphere along magnetic lines of

force and set off these displays [see "Aurora and Airglow," by C. T. Elvey and Franklin E. Roach; SCIENTIFIC AMERICAN, September, 1955]. But the theory underlying this explanation did not explain satisfactorily why the aurora and the high-intensity radiation we had detected should occur in the auroral zone and not in the vicinity of the geomagnetic pole itself. Nor could it account for the high energies required to carry the solar particles through the atmosphere to such relatively low altitudes.

The mystery deepened when we found in later studies that the radiation persists almost continuously in the zone above 30 miles, irrespective of visible auroral displays and other known high-altitude disturbances. More discriminating detectors established that the radiation contains large numbers of electrons. Our original observations had detected X-rays only; now it turned out that the X-rays had been generated by the impact of electrons on the skin of the instrument package (as if it had been the "target" in an X-ray tube) and on the sparse atoms of the upper atmosphere itself. Sydney Chapman and Gordon Little at the University of Alaska suggested that such a process might well account for the attenuation of radio signals in the lower ionosphere of the auroral zones.

The International Geophysical Year gave us our first opportunity to investigate the "auroral soft radiation" on a more comprehensive scale. During the



STRUCTURE OF RADIATION BELTS revealed by contours of radiation intensity (black lines) is shown schematically by shading

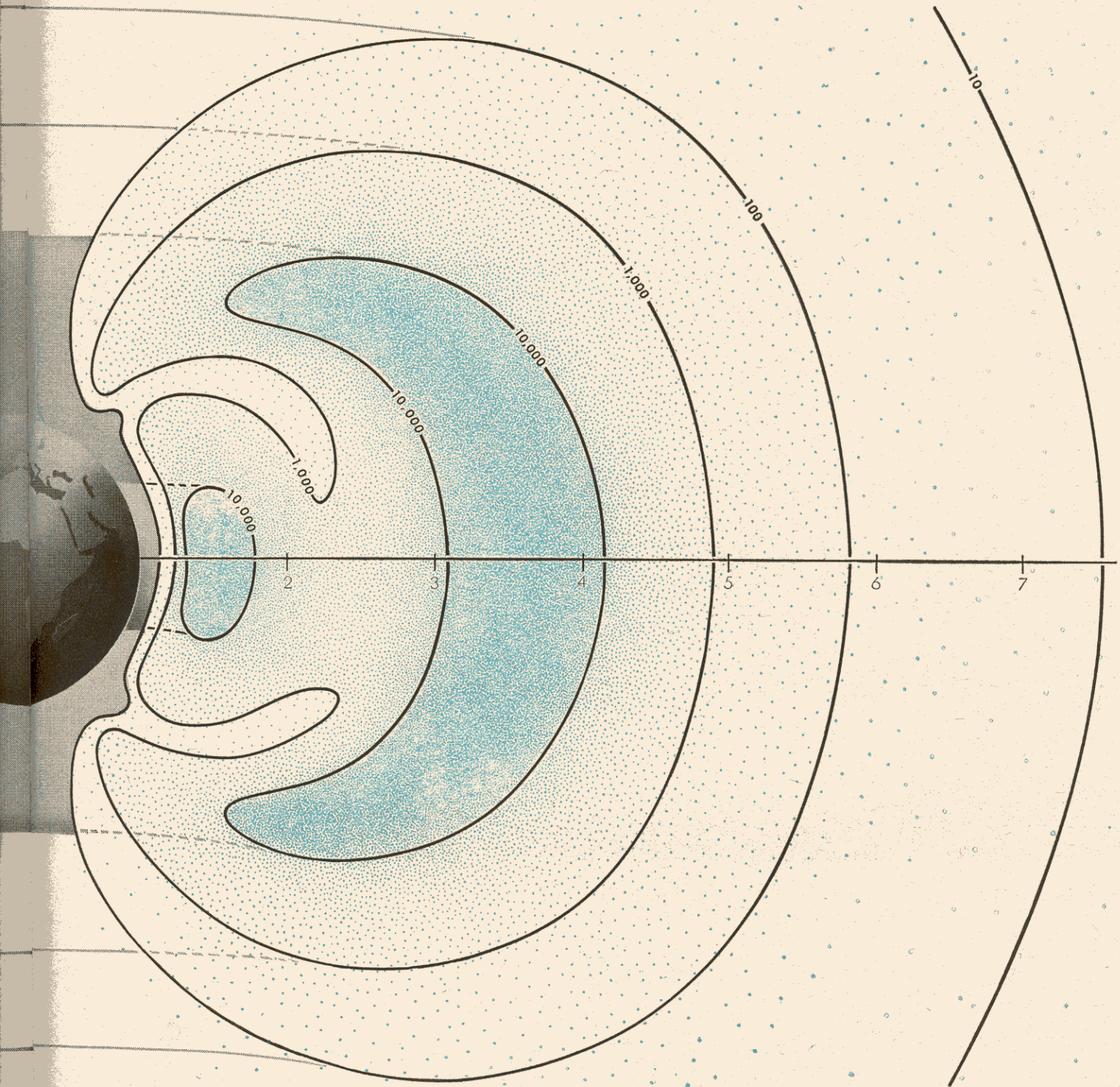
(left); dots (right) suggest distribution of particles in the two belts. Contour numbers give counts per second; horizontal scale

summer and fall of 1957 Laurence Cahill and I launched a number of rockoons off the coast of Greenland and also got off one successful flight in Antarctica. The latter flight established that the radiation exists in the southern as well as the northern auroral zone. In February, 1958, Carl McIlwain fired a series of two-stage rockets through visible auroras above Fort Churchill in Canada, and discovered that the radiation includes

energetic protons (hydrogen nuclei) as well as electrons.

Meanwhile all of us had been pushing a new development that greatly expanded the possibilities for high-altitude research. During the summer of 1955 the President and other Government authorities were finally persuaded that it might be feasible to place artificial satellites in orbit, and authorized an I. G. Y. project for this purpose. In January,

1956, a long-standing group of high-altitude experimentalists, called the Rocket and Satellite Research Panel, held a symposium to consider how the satellites could be most fruitfully employed. At that meeting our group proposed two projects. One was to put a satellite into an orbit nearly pole-to-pole to survey the auroral radiation in both the north and south auroral zones. Such orbits, however, did not appear to be



shows distance in earth radii (about 4,000 miles) from the center of the earth. Particles in the inner belt may originate with the

radioactive decay of neutrons liberated in the upper atmosphere by cosmic rays; those in the outer belt probably originate in the sun.

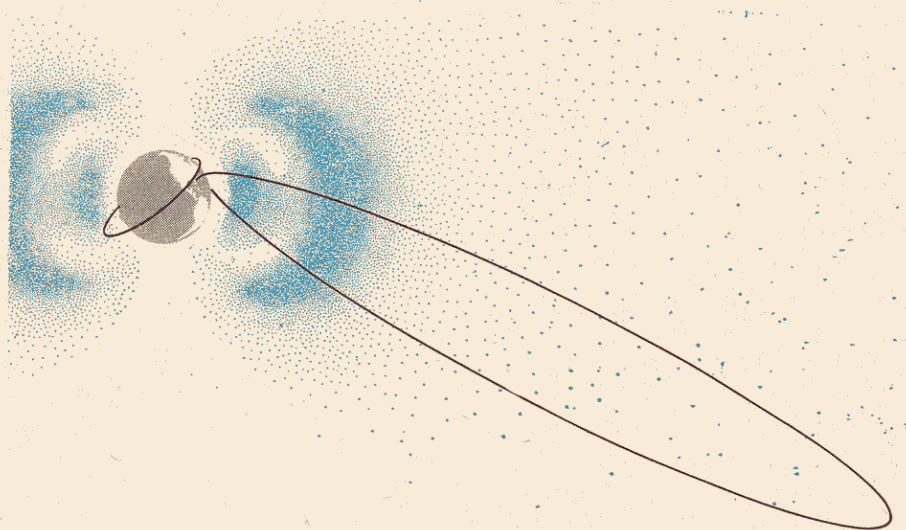
technically feasible in the immediate future. For the time being we were forced to abandon the use of a satellite to probe farther into the auroral soft radiation. We also suggested that a satellite orbiting over the lower latitudes of the earth might usefully be employed in a comprehensive survey of cosmic-ray intensities over those regions. This project was adopted, and we were authorized to prepare suitable experimental

apparatus [see "The Artificial Satellite as a Research Instrument," by James A. Van Allen; *SCIENTIFIC AMERICAN*, November, 1956]. It was planned to place this apparatus on one of the early Vanguard vehicles.

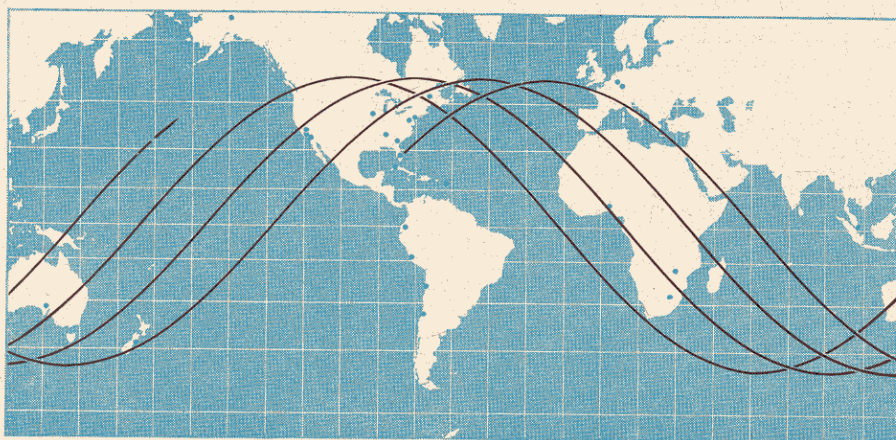
The difficulties and failures of the Vanguard are now history. Sputnik I stimulated some high government officials to accept a proposal that a number of us had been urging for more than

a year: to use the proven Jupiter C rocket as a satellite-launching vehicle. As a result on January 31, 1958, Explorer I went into orbit carrying our simple cosmic-ray detector and a radio to broadcast its readings.

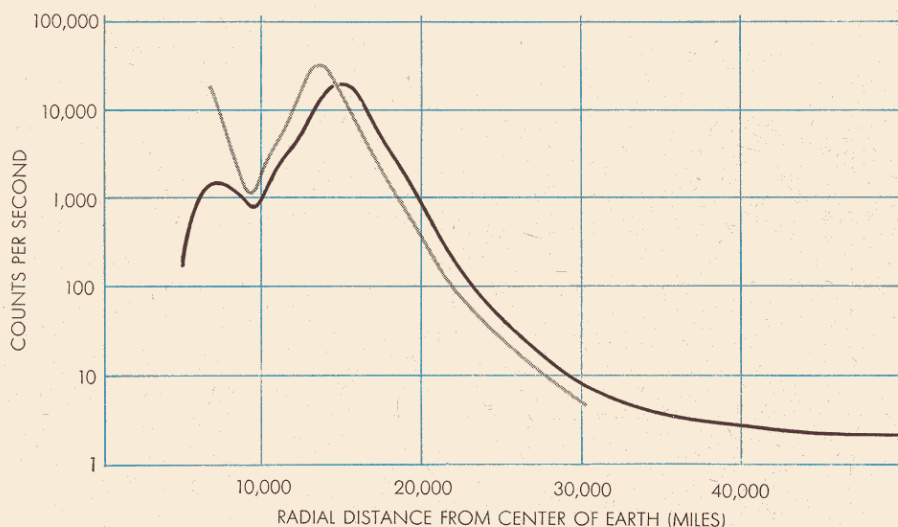
In the first reports from stations located in the U. S. the intensity of radiation increased with altitude along the expected curve. Several weeks later, however, we began to get tapes from stations in



EXPLORER IV AND PIONEER III gave the first detailed picture of the radiation belts. The Explorer IV satellite (*short ellipse*) monitored radiation levels for nearly two months at altitudes up to 1,300 miles. The Pioneer III lunar probe (*long ellipse*) provided data out to 65,000 miles. Its orbit is shown distorted because of the earth's rotation during flight.



EXPLORER IV ORBIT covered the entire region 51 degrees north and south of the equator; the black curve shows a small part of its trace on the earth's surface. More than 25 observation stations (*colored dots*) recorded data from several thousand of the satellite's passes.



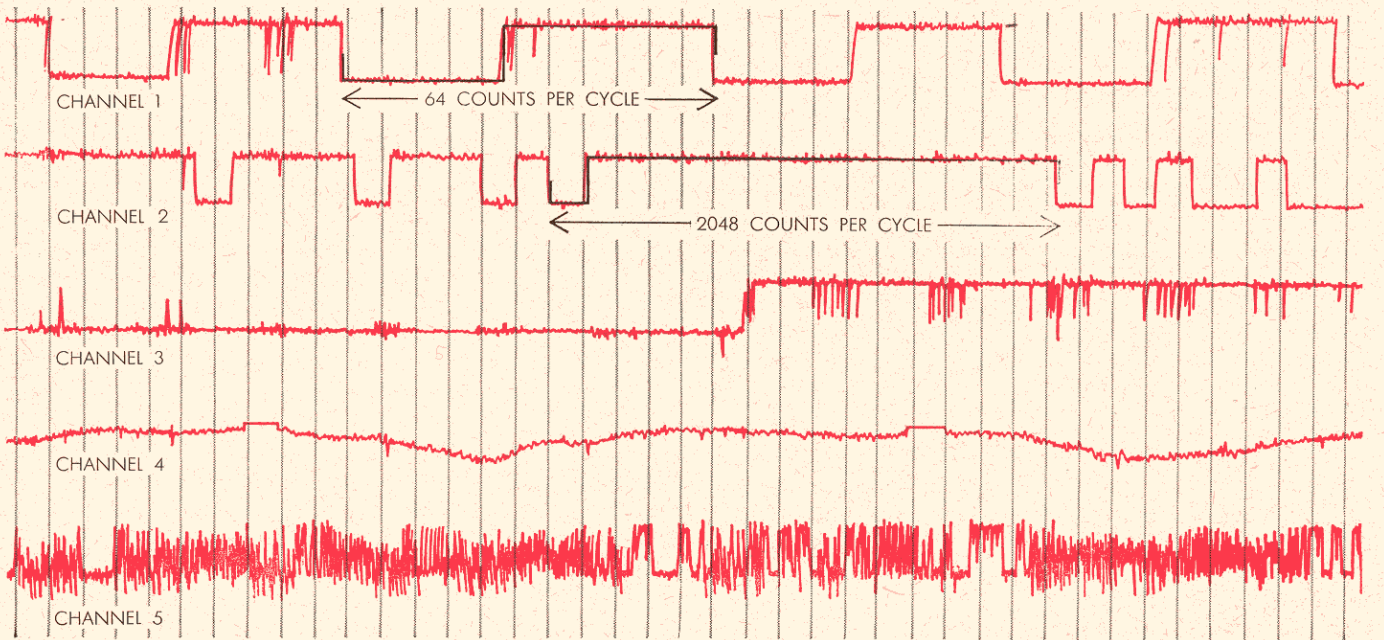
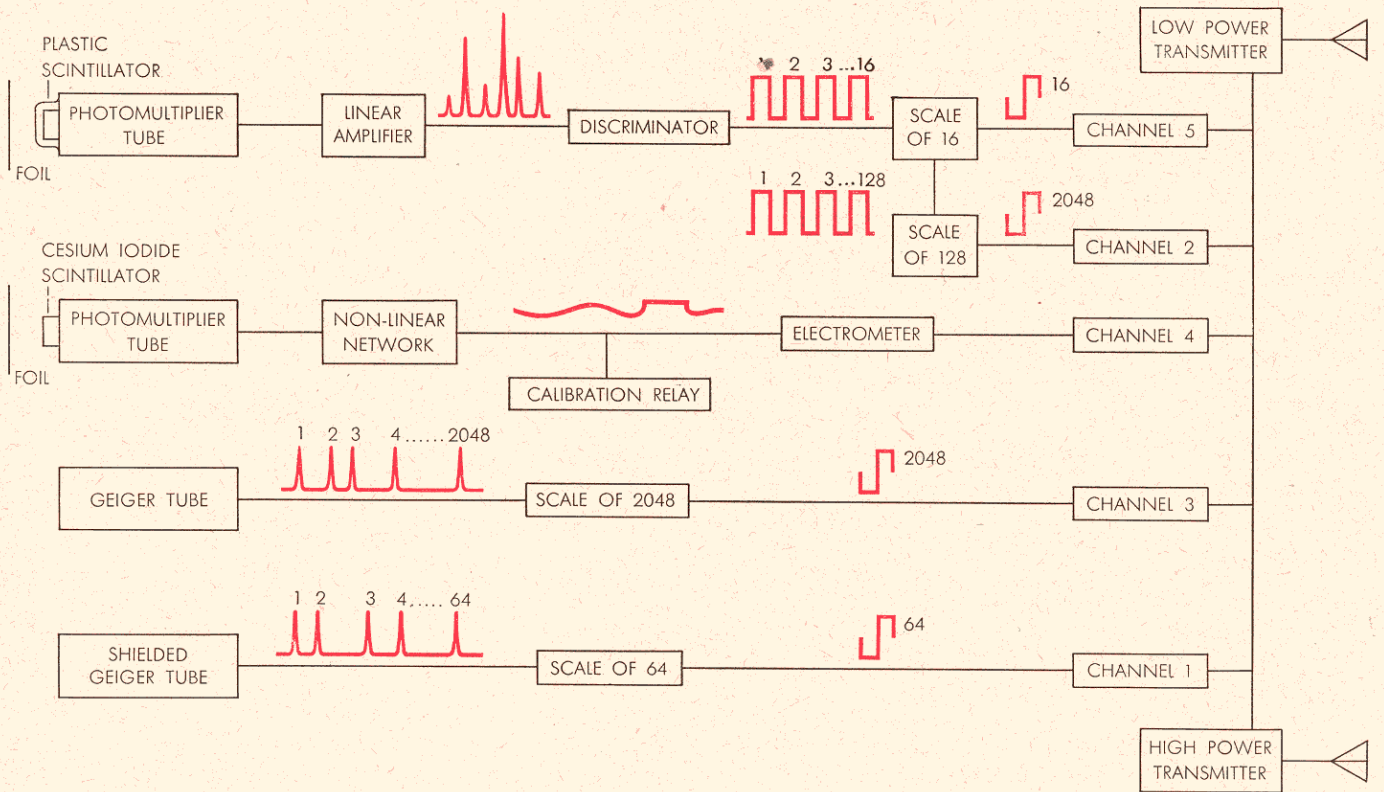
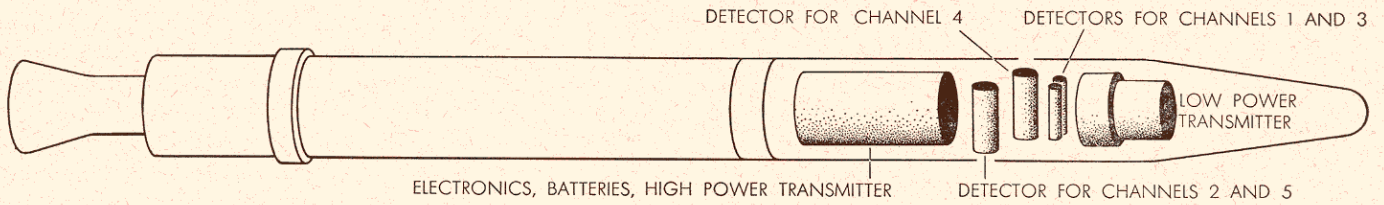
PIONEER III DATA gave the first confirmation of two distinct rings of particles. Counting rates on both the outbound (*black curve*) and the inbound (*gray curve*) legs of the flight showed two peaks. The two curves differ because they cover different sections of the belts.

South America and South Africa which gave us counting rates for much higher altitudes, due to the eccentricity of the satellite's orbit. These records brought us a new surprise. At high altitudes over the equatorial region the apparent counting rate was very low; in some passes it dropped to zero for several minutes. Yet at lower altitudes the rate had quite "reasonable" values—from 30 to 50 counts a second. Again we were uneasy about the trustworthiness of the instruments. The only alternative seemed to be that cosmic rays do not strike the uppermost layers of the atmosphere over the tropics, and we were quite unable to accept this conclusion.

Our uneasiness was increased by the incompleteness of our early data. The Explorer I apparatus broadcast its observations continuously, but its signals could be picked up only intermittently, when the satellite came within range of a ground station. Our original apparatus, designed and developed by George Ludwig for the Vanguard satellites, included a magnetic-tape recorder which could store its observations for a complete orbit around the earth and then report them in a "burst" on radio command from the ground.

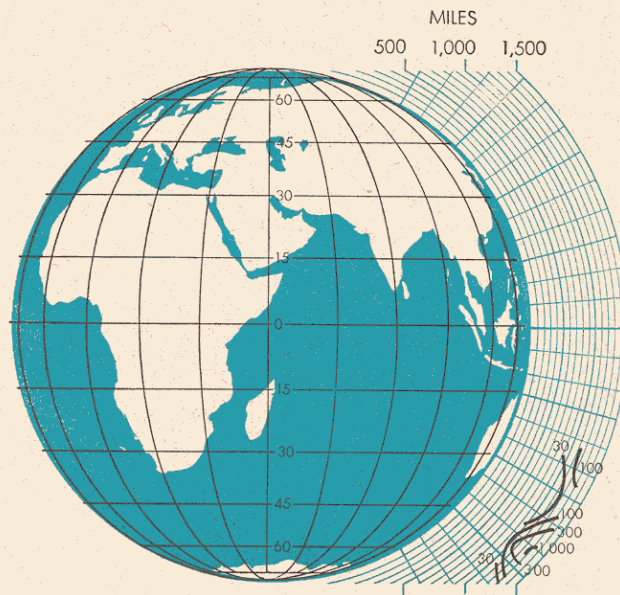
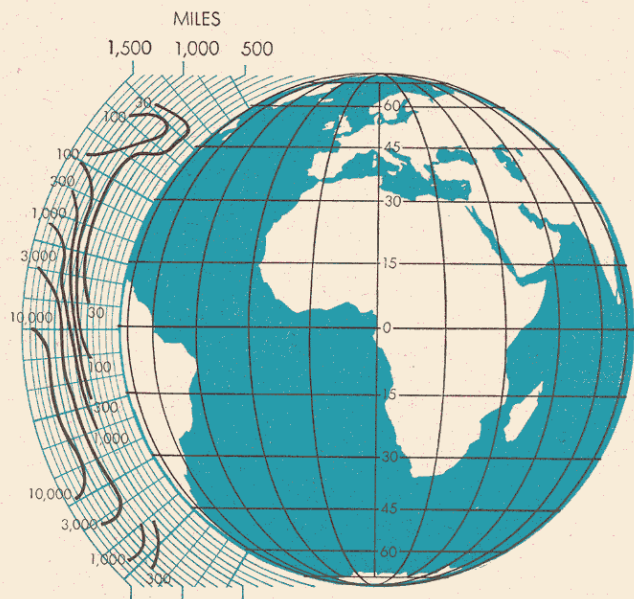
By early February, working with the Jet Propulsion Laboratory, we had converted this apparatus for use in the Explorer II satellite. The first attempt to get it into orbit failed. A second rocket placed Explorer III, carrying identical apparatus, in orbit on March 26. This satellite fully confirmed the anomalous results of Explorer I. At altitudes of 200 to 300 miles the counting rate was low. When the satellite went out to 500 to 600 miles, the apparent rate ascended rapidly and then dropped almost to zero. One day, as we were puzzling over the first tapes from Explorer III, McIlwain suggested the first plausible explanation for their peculiar readings. He had just been calibrating his rocket instruments, and called our attention to something that we all knew but had temporarily forgotten: A sufficiently high level of radiation can jam the counter and send the apparent counting rate to zero. We had discovered an enormously high level of radiation, not a lack of it. As Ernest Ray, a member of our group, inaccurately but graphically exclaimed: "Space is radioactive!"

During the next two months Explorer III produced a large number of playback records, every one of which showed the same effect. At low altitudes the counting rate was reasonably attributable to



EXPLORER IV INSTRUMENTS were designed to give a detailed picture of the nature and intensity of the radiation. Plastic scintillator counted only charged particles above certain energies; two different scaling factors adapted it to both high and low counting rates. Cesium-iodide scintillator measured the total energy input

rather than individual particles. Shielded and unshielded Geiger tubes could be compared to estimate the penetrability of the radiation. Radio signals suggested by the red curves in upper drawing were recorded by ground stations and later played through a multichannel oscillograph to yield records like that shown below.



TWO SETS OF CONTOURS from readings on opposite sides of the earth (left and center) show the northern and southern "horns"

of radiation, which point toward the auroral zones; the contour numbers show radiation intensity in counts per second. The "tipped"

cosmic rays. At higher altitudes—the precise height depended on both latitude and longitude—the count increased to very high values. Up to the points at which the counter jammed, it showed counting rates more than 1,000 times the theoretical expectation for cosmic rays. From the rate of increase and the length of the periods of jamming we judged that the maximum count probably went to several times this level. Since the radiation appeared to resemble the auroral soft radiation, we would not have been surprised to find it in the auroral zone or along the magnetic lines of force that connect these zones. But in the equatorial latitudes these lines of force lie much farther out in space than the altitudes attained by the satellites.

On May 1 of last year we were able to report with confidence to the National Academy of Sciences and the American Physical Society that Explorers I and III had discovered a major new phenomenon: a very great intensity of radiation above altitudes of some 500 miles over the entire region of their traverse, some 34 degrees north and south of the equator. At the same time we advanced the idea that the radiation consists of charged particles—presumably protons and electrons—trapped in the magnetic field of the earth.

We could rule out uncharged particles and gamma and X-rays because they would not be confined by the magnetic field, and so would be observed at lower altitudes. The possibility that the earth's

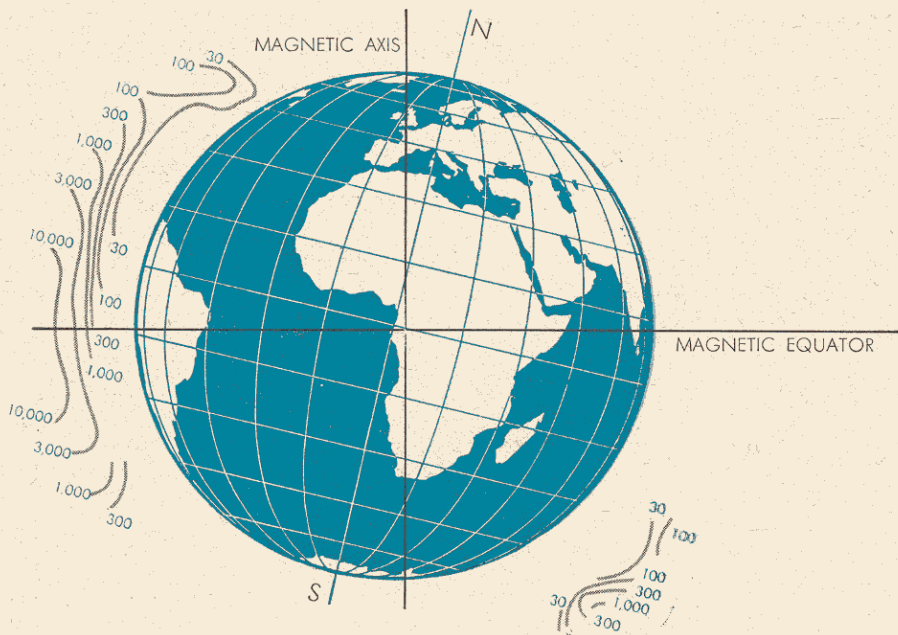
magnetic field might act as a trap for charged particles was first suggested by the Norwegian physicist Carl Störmer in a classical series of papers beginning some 50 years ago, and there was a considerable body of evidence for the existence of low-energy charged particles throughout our solar system and specifically in the vicinity of the earth. But there had been no indication that these particles would possess the high energies we had detected.

From Störmer's theoretical discussion and our own observations we evolved a rough picture of the trapping mechanism. When a fast-moving charged particle is injected into the earth's magnetic field, it describes a corkscrew-shaped trajectory, the center line of which lies along a magnetic line of force. The turns of the helical path are quite open over the equator but become tighter as the particle reaches the stronger magnetic field toward the poles [see illustration at bottom of opposite page]. At the lower end of its trajectory the particle goes into a flat spiral and then winds back along a similar path to the other hemisphere, making the transit from one hemisphere to the other in a second or so. During this time its line of travel shifts slightly, so that the particle drifts slowly around the earth as it corkscrews from hemisphere to hemisphere. An electron drifts from west to east; a proton, in the opposite direction. At each end of its path the particle descends into regions of higher atmospheric density; collisions

with the atoms of atmospheric gases cause it gradually to change its trajectory and to lose energy. After a period of days or weeks the particle is lost into the lower atmosphere.

There was obviously an urgent scientific need to extend these observations with equipment of greater dynamic range and discrimination. In April of 1958 we persuaded several Federal agencies to support further satellite flights of our radiation equipment as an adjunct to the I. G. Y. program, and we received the enthusiastic support of the National Academy of Sciences for the continuation of our work. We also persuaded the Army Ballistic Missile Agency and the Cape Canaveral Air Force Base to try to place the satellite in an orbit more steeply inclined to the equator; at an inclination of about 50 degrees to the equator it would cover a much greater area of earth and skim the edges of both auroral zones.

Working night and day, we set out at once to build new apparatus of a more discriminating nature. We retained the Geiger tube, which we had used in previous satellites, as a basic "simple-minded" detector. To be ready for the highest intensities of radiation, however, we used a much smaller tube that would yield a lower count in a given flux of radiation, and we hooked it into a circuit that would scale down its count by a much larger factor. To obtain a better idea of the penetrability of the radiation



drawing at right shows the essential symmetry of the radiation around the earth's magnetic axis. The structure of the radiation zone was built up from hundreds of observed points.

we shielded a similar Geiger tube with a millimeter of lead. As a more discriminating particle detector we adopted a plastic scintillator and photomultiplier tube to respond to electrons with an energy of more than 650,000 electron volts and to protons of more than 10 million electron volts. Finally we glued a thin cesium-iodide crystal to the window of another photomultiplier tube; the light emitted by the crystal when it was irradiated would measure the over-all input of energy rather than the arrival of individual particles. To keep out light when the crystal faced the sun, we shielded it with thin, opaque nickel foil. A special amplifier gave this detector a large dynamic range extending from about .1 erg per second to 100,000 ergs per second.

Explorer IV carried this apparatus into orbit on July 26, and sent down data for almost two months. Magnetic tapes from some 25 observing stations flowed in steadily from late July to late September; altogether we obtained some 3,600 recorded passes of the satellite. A typical pass was readable for several minutes; some of the best were readable for up to 20 minutes, a large fraction of the time required for the satellite to make a turn around the earth. We are still analyzing this mass of data, but the preliminary results have already proved to be enlightening.

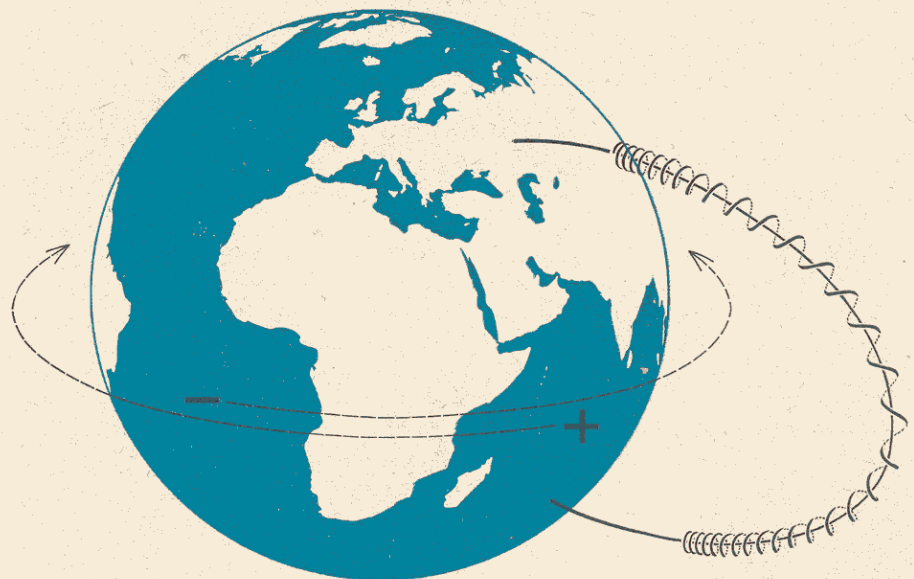
The readings have confirmed our earlier estimates of the maximum levels of radiation. Moreover, we have extended

our observations to more than 50 degrees north and south of the equator and have been able to plot the intensity of the radiation at various latitudes and longitudes for altitudes up to 1,300 miles. The intensity contours follow the shape of the earth in the equatorial region, but as they approach high northern and southern latitudes they swing outward, then inward and sharply outward again to form "horns" reaching down toward

the earth near the auroral zones [see illustrations at the top of these two pages]. The entire picture so far is completely consistent with the magnetic-trapping theory.

It was clear from the contours that Explorers I, III and IV penetrated only the lower portion of the radiation belt. As early as last spring we began to make hypothetical extensions of the observed contours out to a distance of several thousand miles. One of these speculative diagrams showed a single, doughnut-shaped belt of radiation with a ridge around the northern and southern edges of its inner circumference, corresponding to the horns of the contours. Another showed two belts—an outer region with a banana-shaped cross section that extended from the northern to the southern auroral zone and an inner belt over the equator with a bean-shaped cross section [see illustration on pages 40 and 41]. The latter diagram seemed to fit the contours better. In our seminars and after-hour discussions McIlwain held out for the two-belt theory. The rest of us tended to agree with him but preferred to stay with the single "doughnut" because of its simplicity.

To take the question out of the realm of speculation we had to secure measurements through the entire region of radiation. In May, therefore, I arranged to have one of our radiation detectors carried aboard the lunar probes planned for the fall of 1958. On October



TRAPPED PARTICLES spiral rapidly back and forth along a corkscrew-shaped path whose center is a magnetic line of force. At the same time they drift slowly around the earth (broken arrows). Electrons (negative) and protons (positive) drift in opposite directions.

11, 12 and 13 Pioneer I, the first lunar probe, carried our instruments nearly 70,000 miles out from the earth. Though its readings were spotty, they confirmed our belief that the radiation extended outward for many thousands of miles, with its maximum intensity no more than 10,000 miles above the earth.

The next attempted moon shot, Pioneer II, was a fizzle. Pioneer III, however, went off beautifully on December 6. Although this rocket was intended to reach the vicinity of the moon, we were almost as pleased when it failed to do so, for it gave us excellent data on both the upward and downward legs of its flight, cutting through the radiation region for 65,000 miles in two places.

The observations on both legs showed a double peak in intensity [see illustration at bottom of page 42], establishing that there are indeed two belts rather than one. The inner belt reaches its peak at about 2,000 miles from the earth, the outer one at about 10,000 miles. Beyond 10,000 miles the radiation intensity diminishes steadily; it disappears almost completely beyond 40,000 miles. The maximum intensity of radiation in each belt is about 25,000 counts per second, equivalent to some 40,000 parti-

cles per square centimeter per second.

Most of us believe that this great reservoir of particles originates largely in the sun. The particles are somehow injected into the earth's magnetic field, where they are deflected into corkscrew trajectories around lines of force and trapped. In this theoretical scheme the radiation belts resemble a sort of leaky bucket, constantly refilled from the sun and draining away into the atmosphere. A particularly large influx of solar particles causes the bucket to "slop over," mainly in the auroral zone, generating visible auroras, magnetic storms and related disturbances. The normal leakage may be responsible for the airglow which faintly illuminates the night sky and may also account for some of the unexplained high temperatures which have been observed in the upper atmosphere.

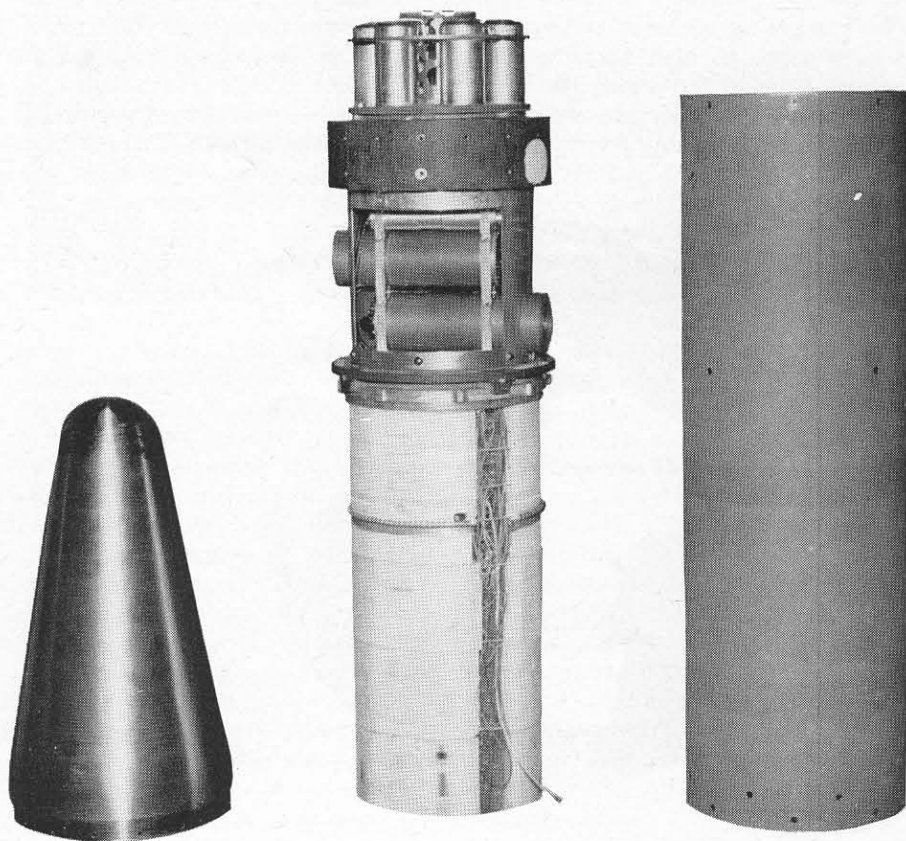
This solar-origin theory, while attractive, presents two problems, neither of which is yet solved. In the first place the energy of many of the particles we have observed is far greater than the presumed energy of solar corpuscles. The kinetic energy of solar corpuscles has not been measured directly, but the time-lag between a solar outburst and the consequent magnetic disturbances

on earth indicates that the particles are slow-moving and thus of relatively low energy. It may be that the earth's magnetic field traps only a high-energy fraction of the particles. Alternatively, some unknown magnetohydrodynamic effect of the earth's field may accelerate the sluggish particles to higher velocities. Some such process in our galaxy has been suggested as responsible for the great energies of cosmic rays. The second problem in the solar-origin theory is that it is difficult to explain how charged particles can get into the earth's magnetic field in the first place. We believe that neither problem is unsolvable.

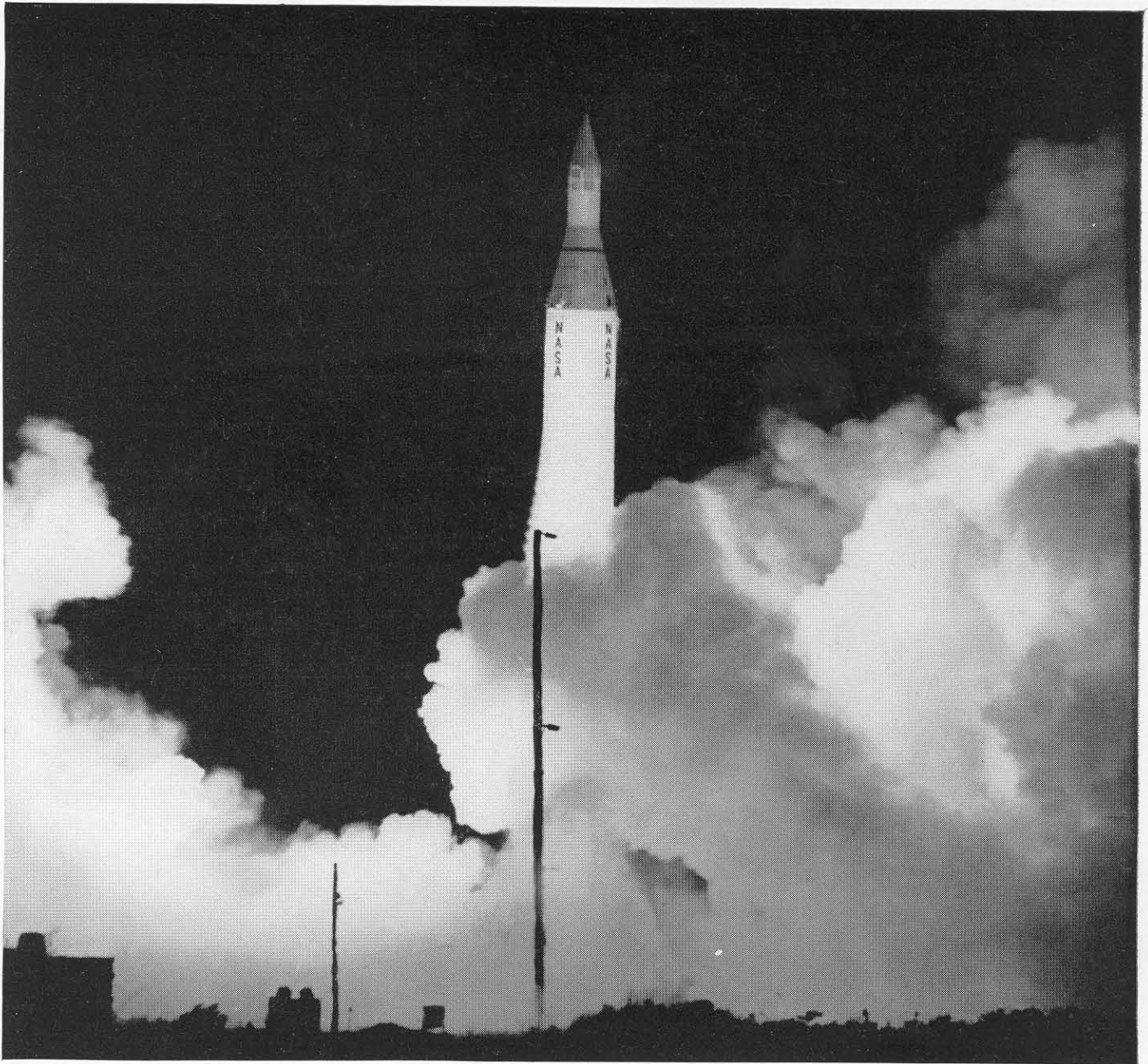
Nicholas Christofilos of the University of California and the Soviet physicist S. N. Vernov have suggested an entirely different theory of how the radiation originates. They note that neutrons are released in large numbers in the earth's upper atmosphere by the impact of cosmic rays. These neutrons, being uncharged, can travel through the magnetic field without deflection. In due course some of them decay there into electrons and protons, which are trapped.

Our group agrees that particle-injection of this sort is going on, and at a rate which can be easily calculated; but we feel for a number of reasons that it cannot be the main source of radiation-belt particles. If we are right in supposing that the radiation belts provide the "reservoir" for the aurora, the neutron hypothesis cannot account for more than one 10,000th of the auroral energy output. Even if the association between the radiation belts and the aurora turns out to be fortuitous, preliminary indications both from our work and from the Russian experience with Sputnik III suggest that most of the particles in the radiation belt have much lower energies than those of particles that would be produced by neutron decay. A full knowledge of the energy distribution of the particles will aid greatly in clarifying their origin.

Neither theory explains why there should be two belts rather than one. It is tempting to combine the two theories and suppose that the inner belt originates with "internal injection"—*i.e.*, neutron-decay products—and the outer one with "external injection" of solar corpuscles. The two-belt configuration may of course be a transitory phenomenon, though the data from Explorer IV and Pioneer III indicate that the separate belts persisted in essentially the same form for at least five months. We should bear in mind, however, that 1958 was a year of great solar activity. Three years



HEAD OF EXPLORER IV includes nose cone (left), instrument "payload" (center) and protective shell (right). Payload includes four detectors, two radio transmitters, batteries and associated electronic circuitry. The outer shell is approximately six inches in diameter.



FOUR-STAGE ROCKET launched the Pioneer III moon probe on December 6, 1958. Though the flight failed to reach the moon, its

outbound leg gave a continuous record of radiation out to 65,000 miles; the inbound leg gave data between 30,000 and 10,000 miles.

from now we may well find a much lower over-all intensity and perhaps a different structure altogether.

In addition to these possible long-term changes, there may be short-term fluctuations in the belts. While we feel sure that the influx and leakage of particles must balance in the long run, a major solar outbreak may temporarily increase the intensity of the radiation many-fold. If we were to detect such fluctuations and were to find that they coincide with solar outbursts on the one hand and with terrestrial magnetic disturbances on the other, we would have a plain lead to the origin of the particles. Before long we hope to launch a satellite

that will monitor radiation levels for at least a year.

Our measurements show that the maximum radiation level as of 1958 is equivalent to between 10 and 100 roentgens per hour, depending on the still-undetermined proportion of protons to electrons. Since a human being exposed for two days to even 10 roentgens would have only an even chance of survival, the radiation belts obviously present an obstacle to space flight. Unless some practical way can be found to shield space-travelers against the effects of the radiation, manned space rockets can best take off through the radiation-free zone over

the poles. A "space station" must orbit below 400 miles or beyond 30,000 miles from the earth. We are now planning a satellite flight that will test the efficacy of various methods of shielding.

The hazard to space-travelers may not end even when they have passed the terrestrial radiation belts. According to present knowledge the other planets of our solar system may have magnetic fields comparable to the earth's and thus may possess radiation belts of their own. The moon, however, probably has no belt, because its magnetic field appears to be feeble. Lunar probes should give us more definite information on this point before long.