THE NEAR-EARTH ENVIRONMENT

A Protected Planet

We live, by chance or otherwise, on a highly protected planet. Were it not for the two natural barriers that stand between the Sun and the surface of the Earth, life itself would soon disappear.

These essential shields—the magnetosphere and just beneath it, the gaseous atmosphere of the Earth—protect us from the full fury of the highly variable star with which we live, and from even more energetic atomic particles that arrive from distant cosmic explosions.

The first of these, the magnetosphere, is an invisible cage built of magnetic lines of force that deflect the vast majority but not all of the high energy electrons, protons, and ions that continually assault the Earth on their way outward from the Sun. But this over-arching shell offers no defense at all against a second and equally persistent hazard: the onrushing streams of highly energetic ultraviolet, x-ray and gamma ray photons that pour outward from the Sun and against the Earth, day after day.

Atoms and molecules of oxygen and nitrogen in the air provide this second shield by filtering the electromagnetic radiation that enters the top of the atmosphere. In this wavelength-selective process, solar heat and visible radiation are allowed to pass through the atmosphere from top to bottom to warm and illuminate the surface of the Earth, while solar gamma rays, x-rays, and the most damaging ultraviolet rays are absorbed and thus removed.

The same atoms and molecules of air also serve as a second line of defense against high energy atomic particles, including those of neutral charge.

The incoming solar or cosmic particles that make their way through the magnetosphere and into the atmosphere lose much of their energy in the course of collisions with atoms and molecules of air, and in the process are de-fused: just as the same all-protective blanket of air shields the surface of the planet from the force of all but the largest meteors that enter the top of the atmosphere. It also keeps us comfortable by holding in much of the heat released from the Sun-warmed surface of the Earth.

Beyond the outer edges of these two trusted bulwarks—the atmosphere and the magnetosphere—lies a world that is far more hazardous: the airless, lifeless heliosphere, or realm of the Sun, that reaches outward from the star in all directions for ten billion miles and more.

The Air Above Us

Ulf Merbold, a German astronaut, may have best described the Earth's atmosphere when late in 1983, from the vantage point of space, he first looked out at it from above, through a window in the Columbia space shuttle:

> *"For the first time in my life I saw the horizon as a curved line. It was accentuated by a thin seam of dark blue light—our atmosphere. Obviously, this was not the ocean of air I had been told it was so many times… I was terrified by its fragile appearance."*

The blue blanket of air that warms and protects the planet is indeed fragile and surprisingly thin, given all that it must do. The atmosphere of the Earth extends above the surface in the form of air to a nominal height of roughly 100 miles, but not as a uniform or homogeneous layer. More than half of the air that it contains is found within the first ten miles of that distance, and when one is less than a third of the way to the top so much of the molecular air is gone that the daytime sky is no longer blue, but black.

Moreover, along the way, with ever-increasing height above the ground, the atmosphere undergoes major transformations: not only in density and temperature, but in composition and function as well. Were it a country, the atmosphere would be described by geographers and demographers as a loose federation of vertically-separated ethnic regions, each quite different in population density, demographic characteristics, prevailing climate, and the work that gets done there.

Changes on the Way to the Top

Obvious to anyone who has ascended a mountain is the change in the density of the atmosphere with altitude: how much lighter, thinner and headache-ier the air becomes the higher up we go.

In Denver, but a mile above the sea, the density of the air—and hence the amount of oxygen we take in with each deep breath—has already dropped by 15% when compared with air in Boston or San Francisco. On top of Pike's Peak, at 14,100 ft, the oxygen in each breath has been reduced by more than a third from sea-level values, and atop Mt. Everest—at 29,035 ft, the highest point on the planet—by more than half.

The flash of the rising Sun as it appeared in space, above the darkened night-side of the Earth, illuminating the thin, protective shell of our atmosphere (curved blue line). At lower left we see the first glint of the Sun on the vertical stabilizer and other parts of the Space Shuttle Discovery, from which this picture was made. Due to the speed of its orbit around the Earth, astronauts witness a sunrise and sunset every 90 minutes.

So much of the Earth's atmosphere is so vacuous, diaphanous and diffuse that were we to compress its 100 miles or so of thickness into a layer of uniform density, like the air near the surface, we would be left with a sky but five miles high with no air to breathe above it. In squeezing the atmosphere down in this way, the peaks of at least a dozen Himalayan mountains—no longer covered with air—would protrude above it like a chain of islands in mid-ocean.

The *temperature* of the ever-thinning atmosphere also changes with altitude, but neither uniformly nor always in the same direction. A concentrated layer of ozone in the middle atmosphere, created there by the effects of incoming solar ultraviolet radiation on molecules of oxygen, alters what would be a much simpler profile of temperature vs height.

Were the ozone layer not there we would expect the temperature of the air to cool with height above the warm surface of the planet. At some level part way to the top, as the air becomes thinner and thinner, the cooling trend would slow and reverse direction, now heating up with height, due to the absorption of solar short-wave radiation in the upper atmosphere. In that simplest case the profile of temperature versus height above the surface of the Earth would trace out a smooth curve that looked like a big "C" when plotted with warmer temperatures to the right and colder temperatures to the left.

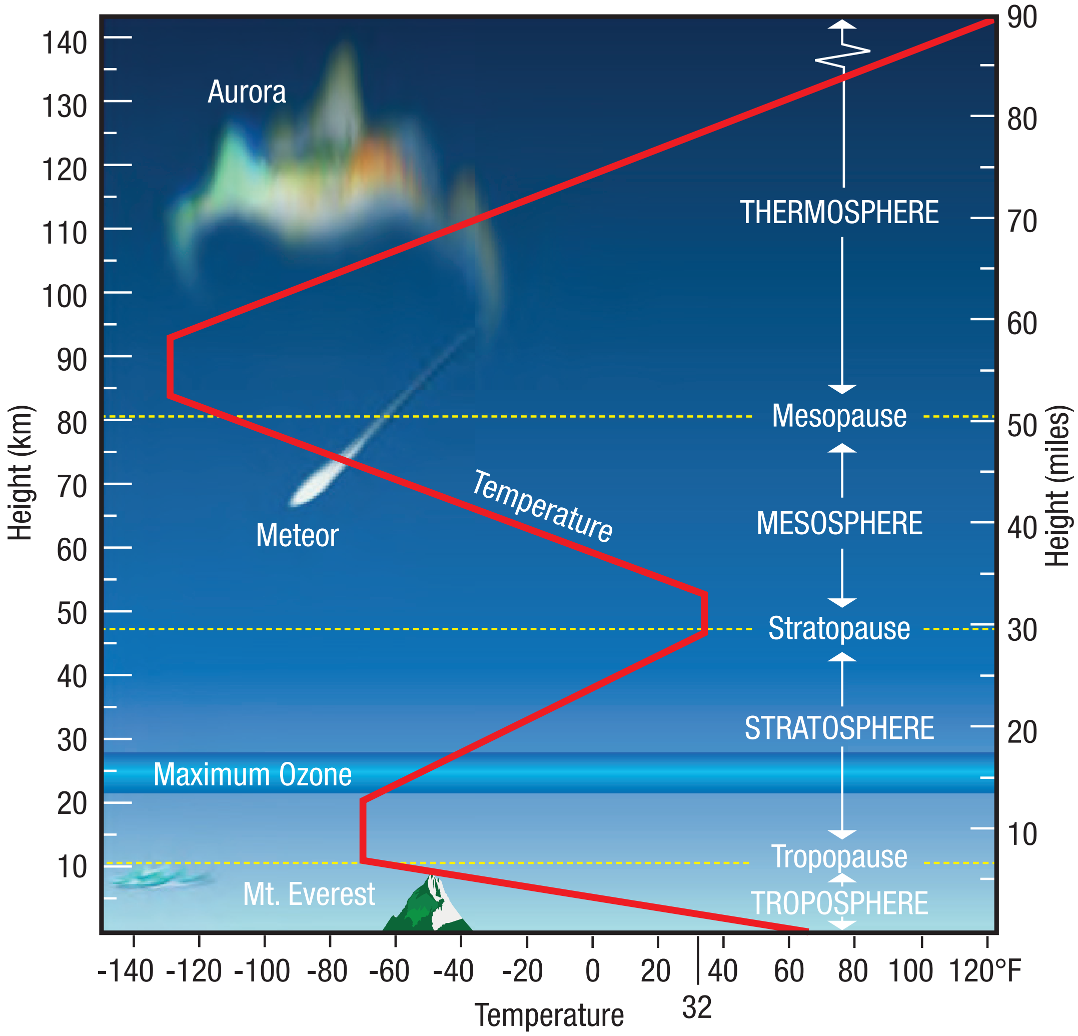
But the stratospheric ozone layer alters this simplified relationship by adding heat to the middle atmosphere, which interrupts the expected cooling with height at an altitude of about ten miles above the surface and reverses its direction. What is then a *warming* with height continues through a run of another 20 miles in altitude, where, at the end of the ozone layer, the atmosphere reverts once more to its initial trend of cooling with height. Much higher, as we approach the top of the atmosphere the cooling trend slows to a stop and reverses itself one more time, now heating monotonically with height, in this case driven by the absorption of solar EUV and x-ray radiation.

With these forced reversals from cooling-with-height to warming, to cooling, and back to warming again the expected "C" shape of the temperature profile of the atmosphere (with a single temperature minimum about one-third of the way to the top) is bent into a more complex curve (shown in red on the following image), with minima at about fifteen and eighty miles above the surface and a maximum in between. These diversions and excursions divide the atmosphere, in terms of temperature and function—into five distinct horizontal layers: the troposphere, stratosphere, mesosphere, thermosphere, and exosphere.

⊙ ⊙ ⊙

But in speaking of the "temperature" of the atmosphere we refer not to the effective temperature which you and I would feel were we to ascend in a balloon gondola through these higher regions of the atmosphere, but the physically-defined kinetic temperature: a measure of the speed of random motions of atoms and molecules of air, no matter how thin or densely packed they are.

How much a person would feel these random movements of the sub-microscopic constituents of air—moving slower at lower temperatures, faster at higher—depends almost entirely upon the density of the air, which in the Earth's atmosphere decreases markedly with altitude. Thus an astronaut



The average temperature of the air above our heads, shown as a red line, as it varies from the surface of the Earth to a height of 90 miles (about 145 kilometers.) Also shown are the conventional horizontal layers of the atmosphere, which are defined by systematic changes in temperature; Mt. Everest, at 29,028 feet, the highest point on the surface of the Earth; the maximum concentration of ozone in the lower stratosphere; and the upper-atmospheric region in which energetic particles from the Sun or magnetosphere excite atoms of air to produce aurorae. Incoming solar particles are chiefly blocked or attenuated by atoms and molecules of air in the thermosphere, although some secondary neutrons from cosmic rays make it all the way down into the troposphere.

working outside the space station in the thermosphere—some 250 miles above the surface of the Earth where the kinetic temperature can range from 700 to more than +1000°F—would be immersed in so rarified a medium that to him or her it would feel very cold.

The Troposphere

The lowest of these five layers of the atmosphere is our home, the troposphere. It is also the thinnest layer, extending upward to a height of but about seven

Back to Table of Contents

miles, or about 40,000 ft: roughly half again as high as Mt. Everest. Were this layer of mostly breathable air included as a transparent shell on a 12-inch library globe it would be about as thick as the paper of this page. Yet within this thin veneer is almost all life on Earth, and the stage on which all but the last few years of human history have been played out.

Within the troposphere are almost all of the moisture in the atmosphere, all weather and climate as we know them, and all but the rarest of clouds. The greenhouse gases—principally water vapor, carbon dioxide and methane—that trap escaping terrestrial heat to keep the planet warm are found mostly within the troposphere.

Unlike all the rest of the atmosphere, the troposphere is an integral part of the biosphere, in that the air in this layer is connected interactively with the oceans, the land, and living things through biogeochemical exchanges of carbon, nitrogen, oxygen and other life-essential elements.

The tropo in tropo-sphere was chosen to indicate "turning" or "change," and more specifically, the steady drop in temperature that distinguishes this lowest layer of the atmosphere as one ascends higher and higher above the solar heated surface of the Earth. In seven miles (somewhat more at the equator and less at the poles), the temperature falls through about 130°, from a global annual mean value of about 60° F at the surface to 70° below zero at the very top, where the next layer, the overlying stratosphere, begins.

The Stratosphere

At the transition between the thin troposphere and the more extensive stratosphere—called the tropo*pause*, meaning *end of the troposphere*—the cooling trend reverses direction to follow a smooth and continued warming that persists through the next 25 miles in altitude.

At one time the temperature of the air above the troposphere was thought to remain more nearly constant with altitude, and hence the name *strat*, meaning *stretched-out* or *extended* atmosphere. In fact, in the course of the climb through the stratosphere, the air heats up by almost 100° F. There, about 32 miles above the surface of the Earth, the temperature of the air (now about +20° F) is back within the range of wintertime temperatures on the ground, which, were the air more dense, would be comfortable enough to enjoy in a heavy coat, scarf and stocking cap.

The lowest stratosphere is the region where most commercial jet aircraft fly, at altitudes of six to eight miles above the surface. The instrumented meteorological sounding balloons that are routinely launched, by hand, every day around the world, make their measurements from ground level to an altitude of about 22 miles, more than halfway to the top of the stratosphere.

This distinctive 25-mile layer of heated air exists because within its bounds is almost all of the ozone in the atmosphere.

Molecular ozone is a voracious but highly selective absorber that removes almost all of the incoming solar ultraviolet radiation that penetrates to this depth in the atmosphere, and in the process, warms the air around it. In performing this function the ozone layer—which reaches down to a scant 8 miles above our heads—stands as the last line of defense in the atmospheric shield: to protect life on the planet from the daytime dose of potential lethal ultraviolet radiation from the Sun.

The intimate relationship between ozone and near-ultraviolet radiation is a give-and-take affair. Solar radiation in the near-ultraviolet region of the spectrum *produces* ozone by breaking down molecules of ordinary oxygen in the stratosphere. But the ozone created there by the Sun in turn *absorbs* some of the incoming solar ultraviolet radiation in an adjoining region of the ultraviolet spectrum.

Through a balance of these opposing effects, changes in the amount of ultraviolet energy we receive from the Sun, as in the course of the 11-year solar activity cycle, alter the amount of ozone in the stratosphere, as do the high-energy solar particles that come our way in the course of solar explosions and eruptions. And in these modern times, our own activities—miles below the stratosphere—have clearly reduced the amount of ozone in the ozone layer, through our release of ozone-destroying chemicals, industrially, agriculturally, and personally, at the surface of the Earth.

Our gaseous atmosphere is composed mostly of nitrogen and oxygen in molecular and atomic form: ozone, a “trace” constituent, accounts for less than one part per million, by weight or volume. Moreover, a protective ozone layer is found on no other planet in the solar system, and the little we have is endangered by what we do.

The Mesosphere and Thermosphere

At the stratopause, or *end of the stratosphere*, the scant ozone that still remains can no longer heat the atmosphere, and there—a little more than thirty miles above the surface—the temperature of the air again begins to drop, resuming the kind of cooling with increased altitude that was the hallmark of the

Back to Table of Contents

troposphere. The cooling trend, though not as steep as that near the ground, persists through the next 22 miles, through the overlying *meso*, or *intermediate* atmosphere: called the mesosphere. Where the cooling ends—a little more than 50 miles above the surface, at the mesopause—the air temperature has fallen to minus 135° F, the lowest temperature found anywhere on Earth or in its atmosphere.

Molecules of oxygen and nitrogen in the mesosphere absorb a portion of the incoming solar ultraviolet radiation and as a result are broken apart into single atoms of these gases. Beginning about halfway up through the mesosphere, at a little more than 40 miles above the surface, these gases are present in the atmosphere in both molecular and atomic form, with the balance shifting with altitude toward the latter.

Ten miles higher, just above the mesopause—in the thermosphere, or *heated atmosphere*—the profile of air temperature versus height undergoes its third and final reversal, and the air once more begins to warm with increasing altitude. The source of the added heat, as in the stratosphere, is again the absorption of short wave radiation from the Sun, but this time in the more energetic extreme-ultraviolet and x-ray region of the solar spectrum. The heat absorbers in the thermosphere—a vast region that stretches from 50 to more than 600 miles above our heads and far above the "nominal" height of the atmosphere—are the last remaining atoms and molecules of both oxygen and nitrogen, which at these lofty heights provide the first line of defense against the full blast of incoming solar short-wave radiation.

⊙ ⊙ ⊙

The end effect, in the lower half of the thermosphere, is a steep and phenomenal rise in the temperature of the air, becoming ever hotter with altitude due to the increased exposure of the air to the direct short-wave radiation from the Sun. From a starting value of 135° below zero at the base of the thermosphere—about 50 miles above the surface of the Earth—the air warms so steadily with altitude that but 20 miles higher, temperatures in the solar-heated thermosphere are in the range of those experienced in southern Arizona in the summer. Just ten miles higher—some 80 miles above the surface of the Earth—the temperature of the air has reached the boiling point of water, although the air is far too thin to do it.

And on and on it warms, without respite until—at an altitude of about 300 miles—these kinetic air temperatures in the sun-lit thermosphere have flattened out in the range of 900 to 2000° F, depending on the highly variable amount of x-ray and ultraviolet radiation that the Sun happens to be releasing at the time.

The tenuous air in the thermosphere, moreover, has little thermal inertia, or heat-holding power, and as a consequence, the air temperature goes through extreme bimodal swings from night to day, as well as strong seasonal variations, and dramatic changes in the course of changes in solar activity.

⊙ ⊙ ⊙

The thermosphere is the region of the atmosphere where many Earth-orbiting spacecraft fly, and the place where the meteors that find their way into the atmosphere are heated by friction to produce shooting stars and meteor showers. It is also the site of auroral displays: the northern and southern lights which are most often seen at higher latitudes in the nighttime sky.

The Ionized Upper Atmosphere

The absorption of solar EUV and x-ray radiation in the thermosphere provokes major changes in its composition as well as temperature: namely the ionization of atoms and molecules of air, which converts these neutral particles into electrically-charged ions (+) and free electrons (-).

The natural tendency of oppositely charged particles that come in contact with each other is to immediately recombine, and in the process, restore neutral atoms and molecules. But in the rarefied air of the upper atmosphere where encounters are less frequent, the newly-created electrons and ions can survive for a longer time: to about one second in the less diffuse air of the mesosphere to about an hour at the top of the thermosphere.

Although these added times are very short, they are long enough to establish and sustain an electrically charged layer, called the ionosphere, as a permanent feature of the upper atmosphere.

This region of electrically-charged particles extends upward from the middle of the mesosphere, about 35 miles above the surface, through the entire thermosphere and well beyond it, with a maximum density of charged particles in the region between about 125 and 375 miles altitude, which puts it within the lower portion of the magnetosphere. But it stretches on and upward, gradually decreasing in density with height, for thousands of miles.

While always there, the ionosphere is highly variable from place to place and time to time, in both structure and density. Dramatic changes occur each day when the Sun rises or sets. The ionosphere is also highly responsive to changes in the x-ray and ultraviolet emission from the Sun, which vary by more than a factor of two in the course of the 11-year solar activity cycle, and by even

Back to Table of Contents

more on shorter time scales in the course of solar flares and the effect of solar rotation. It is also subject to magnetospheric storms, to the passage of meteors, and to the violent winds that sweep through the thermosphere on the heels of thermospheric temperature changes.

The ionosphere is made up of fairly distinct horizontal layers of differing density. When night falls and the Sun's electromagnetic radiation is turned off, the lower portions of the ionosphere—which are called the D and E *layers*, or *regions*—essentially disappear, leaving only the higher and denser F region—which survives, though much depleted, through the night, chiefly by the return of some of the electrons and ions from the magnetosphere.

During the day the F region is a major *exporter* of ions and electrons, which are carried upward into the realm of the magnetosphere by the thermal expansion of the ionosphere. During the night, as the ionosphere cools and contracts, some of these particles are drawn back to it, helping sustain the now depleted F layer until dawn: when, energized by the rising Sun, it bounds back to full strength again.

⊙ ⊙ ⊙

The existence of an electrically-conducting region in the upper atmosphere, though long suspected, was not confirmed until the age of radio, early in the last century when its reflective properties were employed for practical use in the long-distance transmission of radio waves. The reflectivity is greatest for lower frequency (longer wavelength) radio signals, for which the ionosphere serves as a mirror high in the sky to redirect upward-transmitted waves by a single forward reflection—or multiple reflections between the mirror and the ground—around the curvature of the spherical Earth.

Ionospheric reflection of radio waves was first put to use in 1901, by Guglielmo Marconi, with the successful transmission of a brief wireless signal from England to Newfoundland. But it was not until the 1920s, with the advent of broadcast radio, that the ionosphere was widely acknowledged and given its name. The fuller exploitation of all parts of the radio spectrum since that time has demonstrated that hour-to-hour and day-to-day variations and inhomogeneities in the ionosphere can affect all kinds of radio communications, from the lowest to the highest frequencies, including those used today in connection with the communication satellites that are more and more employed for television, satellite phones, GPS, and national security.

⊙ ⊙ ⊙

And so it was that a very old, unknown and invisible feature near the top of the Earth's atmosphere came to be of immense economic and practical importance on the ground.

The End of the Atmosphere

Unlike the oceans, the Earth's atmosphere holds claim to no fixed upper bound. It comes to an end instead through a process of gradual diminution: with the air growing ever thinner and more diffuse the higher one goes, until it eventually blends into the almost-vacuum of near-Earth space, much as recorded sound is faded into silence at the finish of a song.

The imaginary line that marks the threshold of *space*—like the invisible Equator ceremoniously crossed by cruise passengers—is for astronauts arbitrarily taken to be 60 miles above the surface of the Earth, where the pressure and density of air have already fallen by more than a factor of a million. This puts this dividing line in the lower thermosphere, toward the bottom of the ionospheric E layer where neutral and ionized particles coexist, and within the part of the upper atmosphere where incoming meteors paint their star-like trails. But it is considerably lower than where manned spacecraft fly and far below the last vestiges of neutral air.

It was into this lower region of the thermosphere—from about 100 to 200 miles above the surface—that manned spacecraft were first placed in orbit about the Earth: piloted in 1961 by Yuri Gagarin and in 1962 by John Glenn. To reduce the effects of atmospheric drag, instrumented spacecraft with longer life expectancies were later launched into higher orbits, into the more rarefied air of the upper thermosphere and the region beyond it, the exosphere.

The *neutrally-charged* atmosphere comes to an end at about 600 miles above the Earth's surface, where electrically-charged ions and electrons—created by the action of solar EUV and x-ray radiation on neutral atoms and molecules—have become the dominant species.

The change from a *neutral* gas to a *charged*, plasma atmosphere begins in the high stratosphere, some 40 miles above the surface of the Earth. From that level upward—through the three variable layers of concentrated ions and electrons that make up the ionosphere—electrically-charged constituents are more and more plentiful. Behind this persistent shift is the gradual thinning of air with altitude, which (1) allows more of the full blast of short-wave solar radiation to reach the neutral particles; and (2) extends the time allowed to newly-created ions and free electrons before they re-combine and neutralize each other.

Back to Table of Contents

Where the thermosphere ends, the outermost portion of the atmosphere, or exosphere, begins: the last thin vestiges of the atmosphere where all particles—neutral or charged—are on their way out of the atmosphere, and hence the name. The few neutral atoms and molecules that are there escape on the wings of their high thermal velocities, some of them into the magnetosphere.

Into the Magnetosphere

In this vast region of near-Earth space, it is no longer pressure and heat and chemistry, but *magnetic* forces that hold the reins of power and control: trapping and confining electrically-charged particles that have exited the exosphere, while at the same time shielding the planet from many charged particles that come at it from other directions.

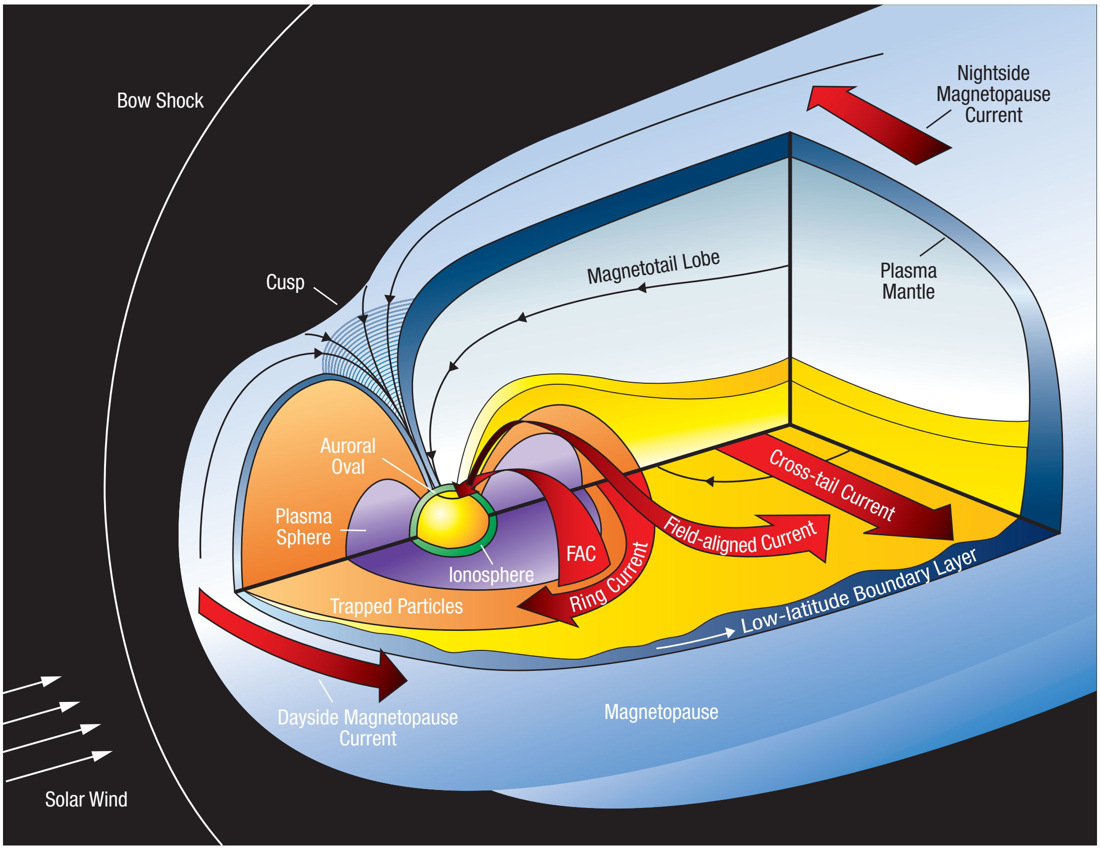
The trapping and confining are accomplished within a grandiose and many-chambered structure made of magnetic lines of force that reach outward from the planet's internal magnetic field.

As our planet's first and outermost line of defense, the magnetosphere deflects away from the Earth most of the charged, energetic atomic particles that stream outward from the Sun, and some that come from other cosmic sources. Its outer boundary, called the magnetopause, is the unseen line, drawn in space, that for billions of years has separated the Earth and its defenses from the hostile world of near-Earth space.

Although we never see it, the mighty fortress of the magnetosphere is truly immense: large enough to dwarf the planet it encloses and protects, extending outward from the Earth toward the Sun for a distance of roughly 40,000 miles, or ten times the radius of the Earth. Moreover in the opposite, anti-solar direction it stretches outward into space for at least 800,000 miles: more than three times the distance to the Moon. Were it as large as its voluminous magnetosphere, our little Earth would replace Jupiter as the largest planet in the solar system.

Inside the inner magnetosphere we are yet within the strong grip of the Earth's gravity, and still in the company of some un-ionized atoms and even molecules of what might be called air, were they not so few and far between. As noted above, these last remnants of oxygen and nitrogen and other elements will continue to be found, though in ever-dwindling numbers, for hundreds of miles above the lower boundary of the magnetosphere.

In this fast-fading part of the gaseous atmosphere, neutral atoms of different elements are sorted by weight based on their response to the pull of the Earth's



A simplified three-dimensional representation of the Sun-facing portion of the Earth's magnetosphere: the volume of space surrounding our planet that is dominated by the geomagnetic field and populated with plasmas from both the ionosphere and the solar wind. In the "upstream" direction, at left, the magnetopause, or outer boundary of the magnetosphere, is compressed by the solar wind on its Sun-facing side, pushing this boundary inward to about 40,000 miles (5 Earth diameters) above the planet's surface. Farther upstream a standing shock wave called the bow shock forms as the supersonic solar wind is slowed and heated by its initial encounter with the still-distant magnetosphere. On the night or "downstream" side, right, interaction with the streaming solar wind stretches the Earth's field into an elongated tail which extends more than 30 times farther: well beyond the orbit of the Moon and what is shown here. The ionized gases that populate the magnetosphere are extremely dilute and wholly invisible to our eyes. Nonetheless, they drive a system of powerful electric currents which during magnetic storms can dissipate well in excess of 100 billion watts of power, comparable to the output of all the electric power plants operating in the United States.

gravity. Through this process, heavier atoms and molecules are held closer and more tightly to the Earth, and the lighter more loosely bound. Thus as one moves higher in the atmosphere there are fewer and fewer heavy elements, until near the top, only hydrogen and helium, the lightest of all the elements, remain.

At that level many of the atoms that are present are no longer the gravitational captives of the Earth. Thanks to the energy and speed they acquired in their passage through the heated thermosphere, these fast-moving atoms are well on their way to freedom: no longer the wards of the little planet that has been their home for more than 4 billion years.

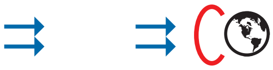
The Form and Function of the Magnetosphere

Long before Columbus, voyagers on land and sea had made use of the fact that the Earth was somehow magnetized: in such a way that one end of a lodestone or an iron compass needle was attracted toward the North Pole of the Earth, and the other to the South Pole.

As more was learned about magnets and the nature of magnetism, it came to be known that the Earth's magnetic field was a lot like that of a simple bar magnet: that is to say, with an oppositely-magnetized pole at either end, connected, one to the other, by curved invisible magnetic lines of force. Each of these was expected to arch outward from its roots in one of the magnetic poles of the planet and follow a gradual C-shaped curve, some distance above the surface of the Earth, to connect at the opposite magnetic pole.

It was also surmised, as early as 1930, that this idealized picture of the Earth's magnetic field—which should take a form similar to the pattern traced by iron filings above a laboratory magnet in a classroom demonstration—would be disturbed and distorted whenever charged particles shot out from the Sun impinged upon it.

Further insight into the likely form of the magnetosphere came in the late 1950s and early 1960s with the realization that a continually blowing solar wind would exert unrelenting pressure on the Sun-facing half of the magnetosphere. This was bound to hammer the idealized rounded C-shaped form of field lines on the day-side of the Earth into a flatter C, and in the process push the protective magnetic shield on that side much closer to the surface of the planet.



But it was not until the voyages of exploration of near-Earth space began, in the remaining years of the 20th century, that the actual form, function and great extent of the Earth's magnetosphere—and its ever-changing nature—came to be fully known and understood.

☉ ☉ ☉

In many ways, including the societal value of what was found, these venturesome missions to explore and map the unknown near-environment of the Earth were not unlike the great voyages of discovery of the late 15th and early 16th centuries.

Most striking among their discoveries was the pronounced asymmetry of the magnetosphere: highly compressed and flattened on the side that faced the Sun, and grossly distended on the opposite, down-wind side. There, above the darkened half of the planet, the solar wind drags the lines of force of the Earth's magnetic field in the anti-solar direction into an extended tail, like that of a comet.

Within this grossly lop-sided magnetic shield, shaped by the Sun, the spherical Earth spins on its own axis at the rate of one rotation per day. The result is that the thickness of the magnetosphere above any fixed place on Earth is continually changing, and by vast amounts: thinnest at local noon; increasing through the afternoon until, as darkness falls, it stretches out above one's head beyond the orbit of the Moon, only to shrink back again to but a twentieth of its nighttime extent when at dawn the Sun again appears.

The Paths That Particles Follow

The Earth's magnetosphere, though an essential protector of life on Earth, is not 100% effective in this regard. Plasma and higher energy particles arriving from the Sun or other celestial sources can find their way, if their energies are sufficient or conditions are right, through the outer bulwarks of the magnetosphere and into the ionosphere and upper atmosphere.

As noted earlier, the incoming solar wind plasma, as it approaches the Earth's magnetosphere first comes upon a more distant, then a nearer barrier. The initial bulwark is a shock wave—known as the bow shock—which forms ahead of the Sun-facing "nose" of the magnetosphere.

There, pressure forces deflect, slow and heat the incoming plasma diverting it away from the direct Sun-to-Earth line. Similar forces produce the swept-back sheath of illuminated dust which forms around the solid head of a comet approaching the Sun.

Slightly farther on, the incoming plasma comes in contact with the magnetopause, the outer boundary of the magnetosphere. There it meets a

Back to Table of Contents

barrier of closed magnetic field lines that are oriented more perpendicularly to the particle's path.

Most of the onrushing plasma is diverted around the magnetopause, like the diversion of a flowing stream by a partially-submerged rock. Some of the diverted flow sweeps over the poles of the Earth, and a fraction of that will find its way into the magnetosphere through unguarded gaps at the poles called the polar cusps.

⊙ ⊙ ⊙

Another means of entry for incoming plasma in either the solar wind or from more energetic particles is available in the vicinity of the equatorial, Sun-facing nose of the magnetopause. Whether that door is open or closed depends upon the polarity and orientation of the imbedded magnetic field that the alien plasma brings with it. When it is the same (or almost the same) as the orientation of the Earth's magnetic field, the particles that impinge upon it will be deflected and redirected as noted above, to continue their downward passage around the outer boundary of the magnetosphere.

If the polarity and orientation of the impinging magnetic field are opposite to that of the Earth (or nearly so) magnetic field lines in the incoming plasma stream will connect to the oppositely-directed field lines of the magnetosphere, which opens up additional routes, some through the polar cusps, into its guarded interior.

Captive Particles In the Magnetosphere

The same lines of force that prevent charged particles from entering the magnetosphere and atmosphere serve as well to entrap and confine any that from whatever source have found their way into the magnetic cage. These include some of the solar and cosmic particles that have worked their way around the bulwarks and into the magnetosphere, as well as those from the ionosphere and exosphere that have entered it from below.

Thus the electrons, protons, and ions that are magnetically confined at any time in the holding cells of the magnetosphere are often strangers: a mixture of immigrants arriving from afar and transients of local origin, who were caught on their way out of our atmosphere in the same magnetic net.

The most energetic are captured solar protons and electrons with energies in the range of a million or more electron volts, and they behave like swarms of angry bees, once free to roam but now confined. These and others less energetic are held in the Earth's magnetic grasp within flat current sheets in the inner magnetosphere that circle the planet in the plane of the Earth's magnetic equator. Others are held in the stretched magnetotail, in the radiation belts, or held within a well-worn set of tightly prescribed paths defined by the arching form of individual lines of magnetic force that are rooted in the north and south auroral ovals.

Constrained in this way, captive particles in the inner magnetosphere trace out a ritual dance of helical spins, bouncing reversals, and longitudinal drifts which take them back and forth from pole to pole, over and over again, while drifting around and around the Earth: and all of it frenetically, at break-neck speed.

⊙ ⊙ ⊙

In the first of the prescribed motions, a moving charged particle that comes into the vicinity of a line of magnetic force is drawn by its influence to gyrate (or circle) around it, as fast as a thousand turns per second, while following its full length from one pole to the other. As the tightly-spiraling particle approaches either end of a closed field line, at high magnetic latitudes, it meets an ever stronger and more crowded polar field, where adjacent field lines are packed more closely together.

There, at the end of the line, the second step is taken. Faced with stronger and more closely-packed field lines, the downward-moving gyrating particle is reflected and repelled, as from a magnetic mirror, and spiraled back from whence it came. When the now-reversed particle approaches the other end of the same closed field line—at the opposite magnetic pole—the same thing happens again. And on and on it goes, on a wild ride: bouncing back and forth from one pole of the Earth to the other, tightly spinning as it goes, completing each pole-to-pole trip in but a few minutes time.

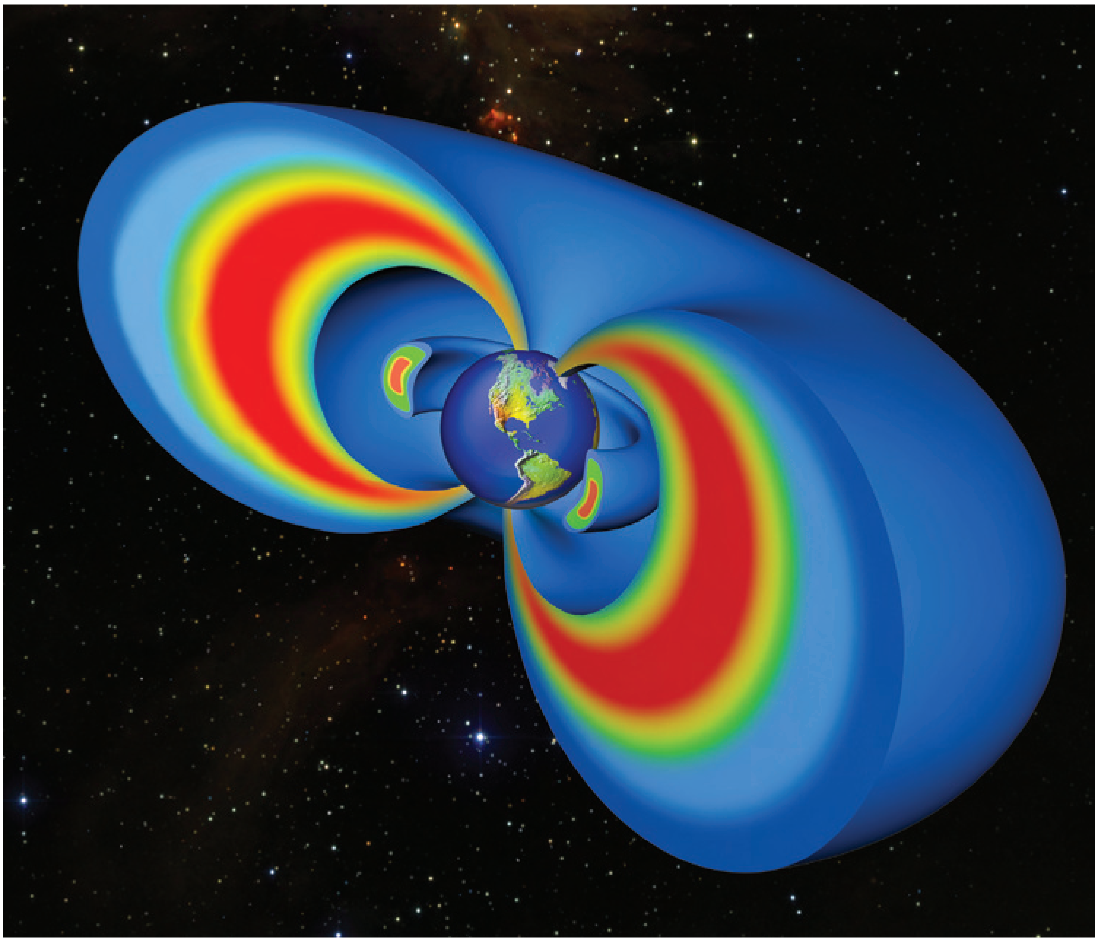
The third obligatory motion is an induced, slower drift in longitude that results from the curvature of the magnetic field lines and the diminished strength of the field with distance above the surface of the planet. The effect is to nudge the gyrating particle a little bit in longitude—an electron in the eastward direction, protons or other positive ions westward—each time it bounces. Repeated nudging pushes it, bit by bit, around the Earth, such that the pole-to-pole motions of the particle sweep over the entire surface of the planet, all in about one hour.

Back to Table of Contents

The paths of the hordes of particles engaged in these and other patterns of motion delineate the form of the enclosures in which captured particles are detained. One follows the framework of closed field lines, reaching from pole to pole, that define the basic form of the Earth's magnetic field. The other is a ring of current flowing around the magnetic equator of the Earth, which defines two bagel-shaped reservoirs that store charged particles in two belts of charged particles.

The Earth's Radiation Belts

These near-Earth reservoirs of magnetically-trapped particles are the inner and outer Van Allen radiation belts, which were unknown and largely unexpected until early 1958, when they were chanced upon by the first U.S. satellite, *Explorer 1*, and named for their discoverer, the late James A. Van Allen.



Schematic depiction, not to scale, of an idealized cross section of the plasmasphere (in blue) and the two doughnut-shaped concentrations of the high energy atomic particles that lie within its bounds, defining the Earth's inner and outer radiation Van Allen belts. Zones where the concentration of particles is greatest are colored red, then yellow and green. Between the inner and outer radiation belts is the (blue) gap or slot region, a zone swept free of particles by physical interactions.

Particles trapped within them are confined within two distinct and clearly separated zones, each shaped much like a bagel, that surround the Earth. The inner radiation belt—containing energetic electrons, protons and heavier ions, all in frantic motion—extends from the top of the atmosphere to a height of about 12,000 miles: a distance, measured from the *center* of the planet, of about three Earth radii, or 3 R_E.

The outer belt, of similar shape but larger diameter, reaches from about 16,000 to 24,000 miles (and at times as much as 36,000 miles) above the surface of the Earth, or between about 4 and 6 R_E, and at times as much as 10 R_E. Within it are the lightest and least energetic of charged atomic particles, which are weaker electrons with energies in the range of 10,000 to about one million electron volts. They come largely in lower energy plasmas that have entered the magnetosphere from the solar wind and the Earth's ionosphere.

Between the inner and outer belts—from about 12,000 to 16,000 miles above the surface of the Earth—is an empty gap or slot about 4000 miles wide: a region of near-Earth space in which protons but few if any electrons are found. This empty region between the highly-populated radiation belts is the result of a loss mechanism involving the interaction of charged particles and electromagnetic waves, that is particularly effective in that zone.

⊙ ⊙ ⊙

One of the ways by which moving charged particles can escape the radiation belts is through a chance encounter with any force that sufficiently alters its energy or disturbs its path and manner of movement. This can occur in the presence of electromagnetic radiation of a frequency that happens to resonate with the particle's own motion. This wave-particle interaction, which has maximum effect at a distance of 3 to 4 R_E, creates the gap between the inner and outer radiation belts: sweeping out newly-bound electrons as fast as they are supplied.

⊙ ⊙ ⊙

Few if any of the charged particles that were chanced upon in the Earth's radiation belts in 1958 are still there today, for although some can indeed remain for hundreds of years, following the same well-worn paths, most—in time—will find their way out of captivity, through the wave-particle process just described, or other mechanisms.

Among the other processes that continually deplete the particle belts are collisions between the charged particles that spiral downward with atoms and molecules of air at the foot points of magnetic field lines high in the atmosphere, other forms of downward diffusion, and almost any externally-caused disturbance of the magnetosphere. Particles to replace the departed continually arrive from three quite different sources: the Sun, the Earth's ionosphere, and some of the energetic particles called cosmic rays.

Cosmic rays that come from outside the solar system arrive with energies per particle of up to a billion or more electron volts: sufficient to speed right through the magnetosphere as though it weren't there. Those found in the inner radiation belt are their less energetic "daughter particles"—a second generation of ions created in the Earth's atmosphere when the original primary cosmic ray particles collide with atoms and molecules of air. Some of these secondary cosmic rays can move upward by diffusion through the upper atmosphere and into the magnetosphere, where, if they carry an electric charge, they are entrapped.

These upwardly-mobile secondary cosmic rays serve also as a primary source of particles for the inner radiation belt.

Charged particles are also created in the mesosphere and thermosphere when neutral atoms and molecules of air absorb far ultraviolet and x-ray radiation from the Sun. Upwardly-mobile ions and electrons from the upper thermosphere are subject to the same fate of magnetic capture and confinement as those produced by incoming solar particles and galactic cosmic rays.

* * *

Test explosions of nuclear weapons in near-Earth space—at a time when they were allowed—were also found to add energetic charged particles to the Earth's radiation belts. Best remembered is the *"Starfish Prime"* test conducted by the U.S. over Johnston Island in the Pacific Ocean on July 9, 1962: when an awareness of the radiation belts was but four years old.

THE EARTH'S RADIATION BELTS

FEATURE	GEOCENTRIC DISTANCE IN R_E	DISTANCE ABOVE EARTH'S SURFACE IN MILES
Inner Radiation Belt	1.2 - 3	650 – 12,000
Slot Region	3 - 4	12,000 – 16,000
Outer Radiation Belt	4 - 6+	16,000 – 24,000+
Plasmasphere	1.2 - 6	1000 – 20,000
For Reference: Orbits of GPS Satellites	5	≈16,000
Geosynchronous Orbit	6.6	22, 200

Exploded in the thermosphere at an altitude of 250 miles (about the same height at which the *International Space Station* now operates), the *Starfish* 1.4 megaton thermonuclear device filled the radiation belts with highly energetic electrons and positrons, some of which were still there five years later. The consequences of this apparently unanticipated happening were more than academic, for these entrapped particles of human origin crippled about one-third of all spacecraft in low-Earth orbit at the time, including *Telstar*, the first commercial communications satellite.

The Plasmasphere

Co-existing with the swarms of high energy particles held captive in the inner and outer radiation belts is another more extensive torus, depicted in blue on the figure on page 88, called the plasmasphere, which is filled with charged particles whose energies are many orders of magnitude weaker. Particles found in the plasmasphere are electrons, protons and singly ionized helium, along with a small amount of singly-ionized oxygen. These are for the most part remnants of the splitting apart of these atoms by solar short-wave radiation far below in the thermosphere and mesosphere. Atmospheric diffusion carries them along magnetic field lines into the magnetosphere, where they are captured and held.

The energy of any one of these more lethargic or "cold plasma" particles is at most a few electron volts, which for comparison corresponds to the thermal motion imparted by being heated to 20,000° to perhaps 40,000° F. In contrast, the electrons and ions in the hotter plasma of the inner radiation belt are suprathermal particles with energies of up to a million or more electron volts, corresponding to almost unimaginable temperatures of billions of degrees.

The plasmasphere extends from just above the ionosphere—about 1000 miles above the surface of the Earth—to a distance of up to 20,000 miles, or 6 R_E, which is almost a tenth as far as the Moon. As such, the plasmasphere overlaps the full extent of the inner radiation belt (about 1¼ to 3 R_E) and a good part of the outer one (4 to 6 R_E).

The cold plasma in the plasmasphere is held there by the same lines of magnetic force that confine hotter plasma in the radiation belts. But since these particles are less energetic, they are bound by a different set of constraints.

One difference is that they are spread continuously and without interruption from the exosphere to about 6 R_E, for they are unaffected by the "zone of avoidance" mechanism that sweeps energetic electrons from the gap in the radiation belts between 3 and 4 R_E.

As a result, the extent of the plasmasphere on any day, and how densely it is filled, are more variable than the radiation belts. Particle density can vary from about 1000 to 20,000 captive protons per cubic inch. The outer boundary of the plasmasphere, called the plasmapause, can reach all the way out to the magnetopause. And like the radiation belts, the plasmapause retreats inward toward the surface of the planet at those times when the Earth's magnetic field becomes connected to that of the Sun, moving at times from a distance of about 20,000 to but 4000 miles above the surface of the Earth. Like the magnetosphere of which it is a part, the plasmasphere is also asymmetric, bulging outward on the nighttime side.

Because of their low energies, particles in the plasmasphere pose no hazard to spaceflight, in contrast to the swarms of charged particles that encircle the planet in the radiation belts which can affect manned or instrumented spacecraft that enter or pass near them. Manned spacecraft that routinely travel in orbits nearer the surface of the Earth—like those that took the first cosmonauts and astronauts around the world in ninety minutes, or the well-worn trails of the *International Space Station*—are largely exempt, for these venture no farther than a few hundred miles above the surface of the Earth, well within the thermosphere and ionosphere.

The Heliosphere

When spacecraft cross the outer boundary of the magnetosphere they leave behind all natural protection against high-energy atomic particles from the Sun or the cosmos. From this point onward, and for at least the next 15 billion miles, spacecraft—manned or instrumented—will travel in a much harsher environment: the vast untamed realm of the much extended Sun, called the heliosphere.

This immense zone in interstellar space is the region dominated by the solar wind and the remnant parts of the Sun's magnetic field that are carried with it, and it stretches outward from the Sun to a distance well beyond the world of planets and planetesimals. Beyond its distant outer boundary—called the heliopause—lies the broader realm of other stars and other galaxies.

The heliosphere is commonly described as an oversize "bubble" surrounding the Sun and all the planets: a protected zone within the interstellar plasma, carved out by the solar wind and sustained by the Sun's extended magnetic field. It "protects" in the sense that the solar plasma that flows continually outward from the Sun is strong enough to fend off most of the plasma that comes in stellar winds from other stars, and to keep out all but the most energetic cosmic rays.

In this sense, the heliosphere is a much larger version of the Earth's magnetosphere, which provides a smaller bubble of magnetic protection around our tiny planet, protecting life upon it from most incoming charged atomic particles, both from the Sun and from more distant cosmic sources. The mighty Sun would seem to require no such protection for itself from incoming atomic particles—whatever their energy—although the planets, all of which are found within the innermost core of the heliosphere, benefit from the presence of this distant outer rampart and the additional magnetic shielding that it provides.

Within the heliosphere the solar wind blows radially away from the Sun and in all directions. It does so until the solar plasma has traveled so far that the outward pressure it exerts, determined by its speed and its ever decreasing density, no longer exceeds the competing, combined pressure of similar stellar winds from all the other stars, near or far.

⊙ ⊙ ⊙

The distance at which the solar wind first meets and is overcome by the pressure of incoming stellar winds varies with direction. It is nearest on the side that faces the direction in which the Sun and the heliosphere are moving through the interstellar medium: which indeed they are, taking the Earth and all the other planets with them. The velocity of the Sun relative to the local interstellar medium in our region of the Galaxy is a little more than 12 miles per second, or 45,000 miles per hour. And its direction—should you want to check it out on a starry night—is toward the bright star *Vega* in the summer constellation *Lyra*, just north of the Milky Way.

At the place where the solar wind first begins to feel the competing force of stellar winds a shock wave forms, much like the bow shock that forms near the outer boundary of the Earth's magnetosphere, or the bow wave ahead of the bow of a moving ship. In passing through this termination shock, the solar wind slows down from supersonic (about a million miles per hour) to subsonic speeds, and is partially deflected away from the direction of the Sun's motion through the interstellar medium.

The heliopause—marking the true outer limit of the heliosphere—lies beyond the termination shock and is separated from it by the sheath of weakened and diverted solar wind. At its closest point, directly ahead of the moving Sun, the distance between the termination shock and the heliopause is a billion miles or so. This places the end of the heliosphere at a minimum of perhaps 10 billion miles from the Sun, or roughly 30 times more distant than Pluto. Thus the Earth and all the other planets and planet-like objects occupy but the inner core of the far larger heliosphere.

Back to Table of Contents

The shape of the heliosphere, like that of the magnetosphere, most closely resembles an extended windsock, with a long tail that stretches behind the moving Sun as it travels through space. It is also thought that the extent of the heliosphere may respond to solar activity in that some *in situ* evidence suggests that the termination shock lies farther from the Sun in years when the Sun is more active and closer in years of minimum activity.

In this case the response at the termination shock to changing conditions on the Sun will be necessarily delayed by the year or so needed for the solar wind to reach that far-distant point.

Cruising the Heliosphere

More than 3,000 spacecraft—an average of more than 100 each year—have been successfully lifted into near-Earth space since the launch of *Sputnik* in the autumn of 1957. Only a handful of these, however, have ventured beyond the magnetopause: the line that marks the outer limit of our world and the beginning of the heliosphere. Most of the rest have been placed in more mundane, circular orbits around the Earth at varied angles of inclination to the plane of the Earth's equator, including many whose orbits take them in polar orbits over the North and South Poles of the planet.

Most of these Earth-bound spacecraft, including John Glenn's *Mercury 6* and today's *International Space Station*, have followed what are now familiar courses around the planet while clinging closely to it: remaining within the bounds of the upper thermosphere at altitudes of several hundred miles and completing each pass around the Earth in about ninety minutes time.

Some—including almost all communications satellites and the weather satellites that provide the familiar images of cloud cover for the evening news—are placed in higher, geosynchronous or geostationary orbits for which the time needed to circle the Earth is precisely twenty-four hours: the same as the rotation rate of the Earth itself. Satellites that orbit the Earth in this way appear fixed in the sky—unlike the ever-circling Sun and Moon and stars—remaining night and day above the same fixed geographical region on the surface of the Earth.

To achieve a geosynchronous orbit, a spacecraft must be lifted well above the thermosphere to an altitude of 22,200 miles: which is almost a tenth of the distance to the Moon. At this height, it will circle the Earth within the outer radiation belt, still within the magnetosphere but tens of thousands of miles below its outer boundary.

Among those that have gone beyond that line are the manned and unmanned spacecraft that have traveled to the Moon, instrumented explorers of the solar system and its planets, and "*in situ*" samplers and monitors of the solar wind. With very few exceptions, however, all of these ventures into the open space that lies beyond the Earth's protective shell have been limited to the two dimensions of the ecliptic plane in which the Earth orbits the Sun. This restriction has kept these venturers within a very thin slice of the true extent of the three-dimensional heliosphere.

⊙ ⊙ ⊙

The most venturesome of those that have left the ecliptic plane is the *Ulysses* spacecraft, launched in the autumn of 1990, and first dispatched to distant Jupiter. There, almost two years later and far from home, it succeeded in harnessing the gravitational pull of that most massive of the planets to hurl itself, as from a slingshot, out of the ecliptic and into a polar orbit around the Sun, passing alternately over the north and south rotational poles of the star: where no spacecraft from the Earth had ever gone before. And there it operates today.

This truly epic voyage of discovery takes *Ulysses* about as close to the Sun as the orbit of Mars, and as far away as the orbit of Jupiter: in an elliptical orbit that takes more than six years to complete. Along the way, month after month and year upon year, it gathers and sends back to Earth a host of *in situ* measurements that tell of conditions in the solar wind, probing the three-dimensional inner heliosphere through all seasons of the Sun's 11-year cycle of activity. These data have proved to be of inestimable value in defining the velocities and the solar sources of both high- and low-speed streams in the solar wind, the three-dimensional nature of the Sun's extended magnetic field, and the spatial distribution of cosmic rays in the inner heliosphere.

⊙ ⊙ ⊙

Two other pioneering spacecraft, *Voyager 1* and *Voyager 2*, launched about two weeks apart in 1977—thirty-two years ago—have now made their way to the outer limits of the heliosphere. Within perhaps the next ten years the most distant of the pair, *Voyager 1*, is expected to achieve the long-sought goal—like that of finding the source of the Nile, or first reaching the frigid Poles of the Earth—of finding the end of the heliosphere: the heliopause. And crossing it, to enter the interstellar medium.

These identical spacecraft were sent like two ships of discovery on extended voyages through the solar system. On the way *Voyager 1* first visited the giant

Back to Table of Contents

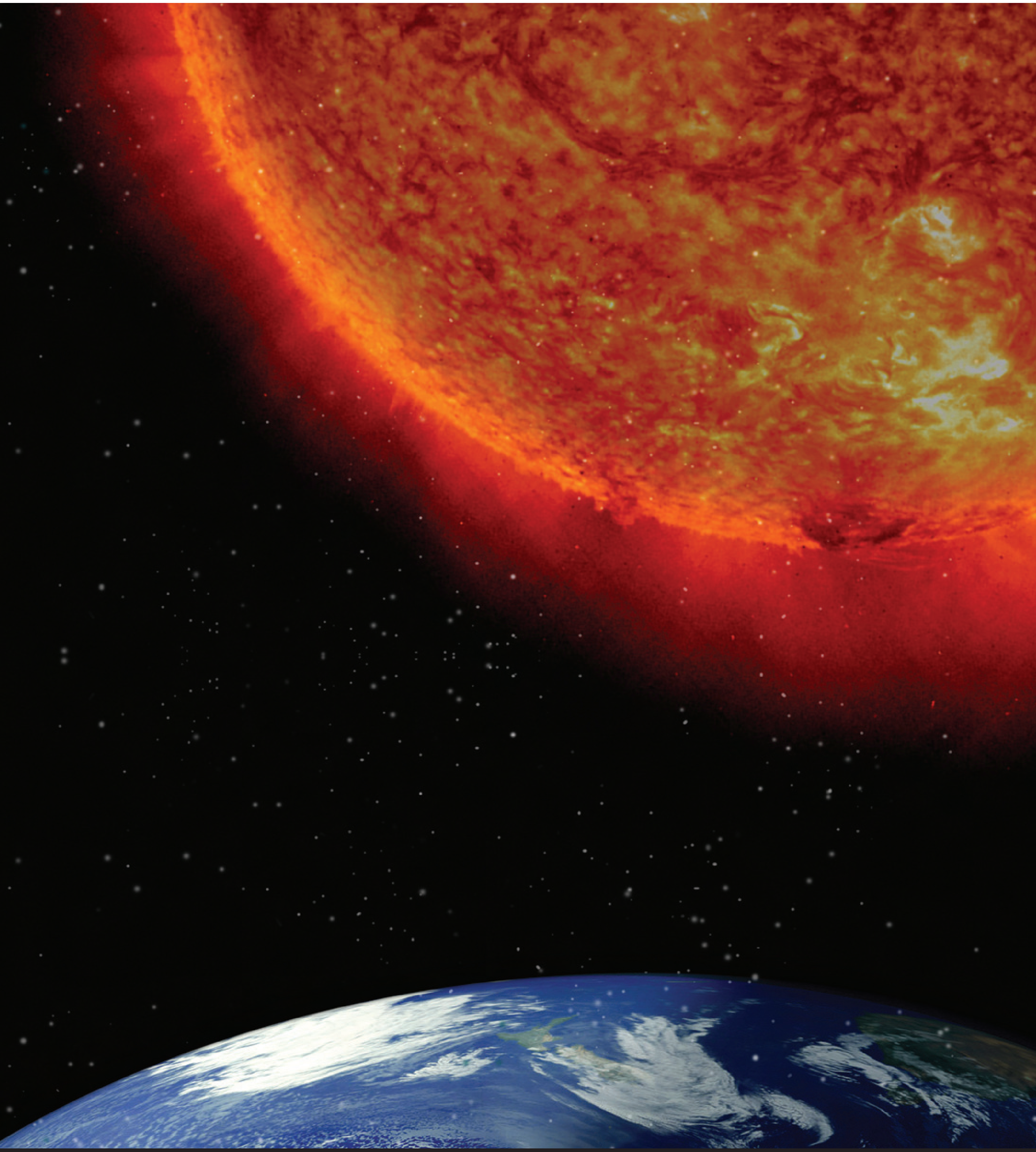
planet Jupiter, and *Voyager 2*, not far behind, gaseous Saturn, Uranus and Neptune, as well as their moons, in part to take the measure of their magnetic fields. By 1989 they had made their last planetary encounter; accomplished all their initial tasks and more; and sailed on to explore the farthest limits of the solar system. Because the radiation from the Sun is so feeble at these great distances—where it appears as but a point of light in a black sky—the Voyagers were both powered by batteries continually charged by the decay of plutonium, which are expected to last until about the year 2020.

On the 16th of December of 2004, after twenty-seven years of voyaging, *Voyager 1*, then about 8.7 billion miles from the Sun and the Earth, came upon and passed through the termination shock of the heliosphere: where the solar wind first experiences effects of opposing gusts from other stars. Like the reflected swells that told early Polynesian navigators that land lay ahead, somewhere over the horizon, this expected wall of high pressure in the farthermost heliosphere served as a clear sign that the heliopause—the true end of the heliosphere—lay out there, somewhere, dead ahead.

The news of this historic crossing, coded in solar wind and magnetic data that were transmitted—like all radio waves—at the speed of light—took thirteen hours to reach the Earth. In the first year of its journey, as *Voyager 1* passed the Moon, the transmission time was less than two seconds, and from Jupiter, forty minutes.

Voyager 1 now travels onward within the heliosheath of diminishing solar winds that stands between the termination shock and the end of the heliosphere, with miles to go before it sleeps. The distance it must still travel to make it there is not known with certainty, but thought to be another billion miles or so. At its present rate of travel it will require almost another decade—until about 2015—to reach that goal.

Voyager 2, following the same course at the same speed is about two years behind it, soon to make its own passage through the termination shock, and to cross the finish line in about 2017. About three years later—after sailing another half billion miles into the Great Unknown—the atomic batteries which by then will have kept the little spacecraft alive for more than forty years will finally fail, leaving the then voiceless ships to sail on alone in silence, into the dark of space.



Radiation in the form of heat and light provides almost all of the energy transferred between the Sun and the Earth, shown here in close proximity. The diameter of the Sun is actually about 100 times larger than that of the Earth, and the distance that separates them—about 93 million miles—is equal to well more than 100 solar diameters. Thus the Earth seen from its constant benefactor would appear little larger than a dot in the sky.