quick study

# How do auroras form?

Robert J. Strangeway

**A key element is a current, aligned along Earth's magnetic field, that results from the coupling of plasmas in and above our planet's atmosphere.**

**Bob Strangeway** is a research geophysicist with the Institute of Geophysics and Planetary Physics at the University of California, Los Angeles.

**The aurora** has long been a source of fascination for artists and scientists alike. The immediate cause of the aurora is the precipitation of energetic particles into the atmosphere. Those particles cause excitation of different atmospheric constituents, including constituents in the ionosphere, the part of the atmosphere ionized by solar radiation. As the excited states relax, they emit auroral light. Figure 1 shows the inherent beauty of the aurora.

Depending on the energy of the precipitating particles, the aurora can have a variety of colors. For energies of a few electron volts or so, the aurora is red and emitted at altitudes above 200 km. Particles with energies of about 1 keV penetrate to lower altitudes and are responsible for the dominant yellow-green color of auroras. Even more energetic particles with energies above 10 keV can get to altitudes below 100 km; at such locations auroras are a deep red or purple color. Auroral emissions, in addition to being in the visible spectrum, occur at UV frequencies. Clearly, scientists know the what of the aurora, but the how—the mechanism that launches the particles into the atmosphere—is still a topic of investigation. This Quick Study emphasizes one aspect of the causal chain and highlights the aurora as a manifestation of electromagnetic forces acting between different regions of a plasma.

Auroras may be divided into two classes: diffuse and discrete. The diffuse aurora, as its name implies, is relatively unstructured. Physicists generally regard it to be a result of particles that are initially trapped in Earth's magnetosphere—the region of space dominated by the magnetic field that connects to Earth's surface and bounded by the so-called magnetopause—and then scattered into the ionosphere and atmosphere through wave–particle interactions or through distortions of Earth's magnetic field. The diffuse aurora can become quite intense at times, especially when the magnetospheric plasma's energy density increases during a geomagnetic storm. The discrete aurora, the focus of my discussion, is much more structured than the diffuse aurora and can be highly variable. It is the one that displays the previously described intense colors associated with precipitating electrons that have energies of several keV.

**Figure 1. An aurora's striking form** is revealed in this photograph taken at Fairbanks, Alaska, in March 2001. The dominant color is a characteristic yellow-green, but wisps of reddish purple are also evident. (Courtesy of Jan Curtis.)

## The big picture

Earth and its magnetic field are embedded in the solar wind, a plasma that rapidly flows away from the Sun. At Earth, the solar wind typically has a speed of about 400 km per second, a density of nearly 10 particles per cm$^3$, and an entrained magnetic field of solar origin whose field strength is in the range of 5–10 nT. The solar wind flows more rapidly than any inertial wave in the plasma—that is, any wave that involves bulk motion of the plasma—and a bow shock consequently forms upwind of the magnetosphere. Downwind of the shock, the solar wind impinges on Earth's magnetic field, compressing the field on the dayside and stretching it out into a large "magnetotail" on the nightside (see the article by Jeffrey Love, PHYSICS TODAY, February 2008, page 31). In addition, if the solar wind's entrained magnetic field is appropriately oriented, magnetic field lines from the solar wind can connect to field lines from Earth in a process known as reconnection.

Reconnection takes place at the dayside of the magnetopause. One consequence of that reconnection is the addition of magnetic flux from the solar wind to the magnetosphere. And because the solar wind and magnetospheric plasma are magnetohydrodynamic (MHD) fluids, the flux addition implies downwind flows in the magnetospheric plasma. Now suppose that reconnection occurred only on the dayside of the magnetosphere. Eventually the entire dayside magnetic field would be connected to the solar wind, and the wind would blow against Earth's ionosphere. That doesn't happen. Flux must therefore be returned to the dayside via nightside reconnection, and so the MHD magnetospheric plasma also contains flows moving upwind from nightside to dayside.

Reconnection on Earth's nightside can be a very impulsive process. Magnetospheric substorms, which involve intensifications of the aurora, are a signature of that impulsive reconnection process. Substorms are not the same as geomagnetic storms. The latter are associated with significant decreases in the magnetic field strength at Earth's surface,

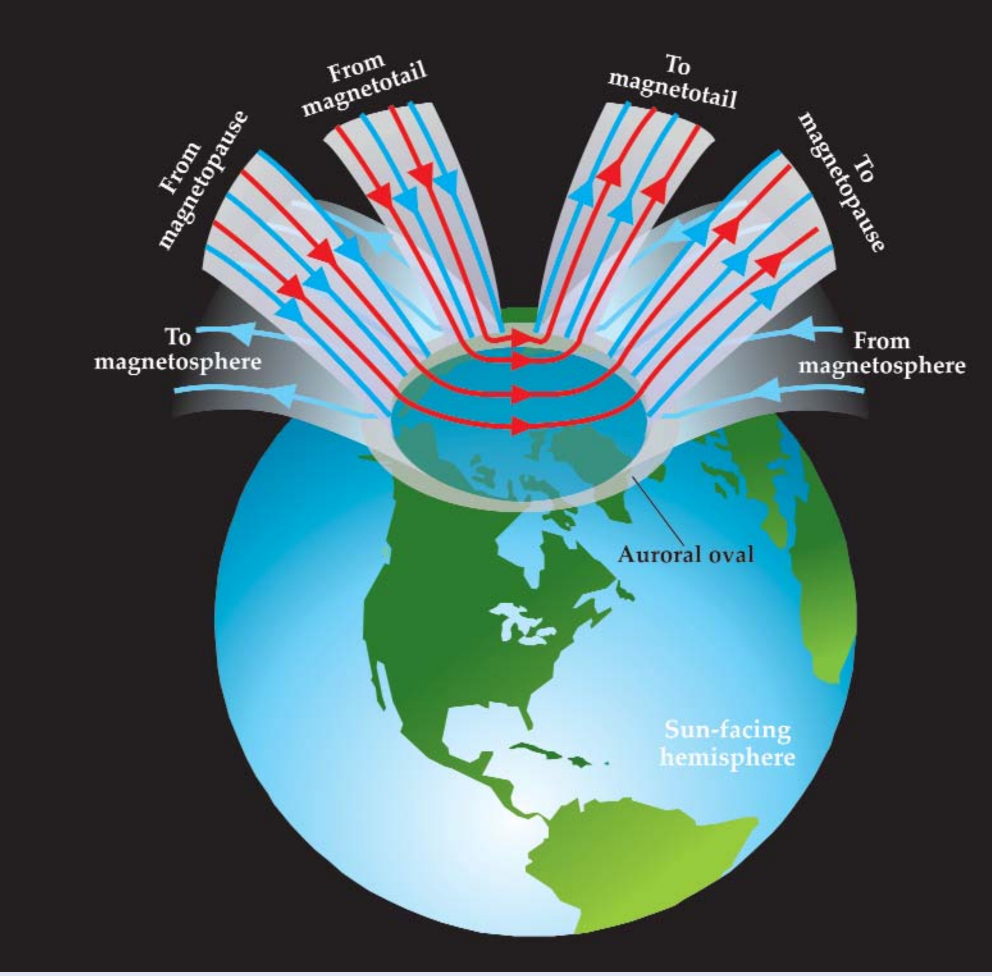


**Figure 2. Field-aligned currents flow** along Earth's magnetic field into and out of the high-latitude ionosphere. This dayside view shows how the field-aligned currents close in the polar ionosphere, where horizontal currents impart a Lorentz force to the ionospheric plasma; that force balances the frictional drag imposed by collisions with the neutral atmosphere. The balance of forces ensures that the ionospheric plasma convection matches that of the plasma in the magnetosphere, the region of space dominated by Earth's magnetic field. The field-aligned currents also close in the magnetosphere, where processes such as magnetic reconnection drive magnetospheric convection. As indicated, some of the currents close in the magnetopause, the boundary of the magnetosphere, and others close in the elongated nightside magnetotail. If the high-latitude field-aligned currents indicated by upward-pointing vectors in the right half of the figure become sufficiently strong, then local acceleration of the precipitating electrons occurs; the result is a bright auroral display. The current lines are colored to make it easier to follow their complicated paths. The auroral oval depicts the region of most frequent and intense auroral activity.

caused by injection and energizing of plasma in the magnetosphere. Geomagnetic storms have time scales of several hours or days; substorms don't last nearly as long, but they generate the brightest auroral displays.

## Enter the ionosphere

Through reconnection, the solar wind imposes flows on the magnetosphere. The magnetosphere in turn imposes those flows on the ionosphere, which, like the magnetosphere, is an MHD fluid. The ionosphere, however, is frictionally coupled to the neutral atmosphere via collisions; nonetheless, the ionospheric flow matches that of the magnetosphere. The match arises because when two MHD plasmas have different velocities, the magnetic field, $\mathbf{B}$, is distorted by the differential flow. That distortion induces horizontal currents in the ionosphere that provide a $\mathbf{j} \times \mathbf{B}$ Lorentz force (more precisely, a force density; here $\mathbf{j}$ is a current density) to balance the atmospheric drag force and guarantee a matched flow. The horizontal currents are coupled to the magnetosphere through field-aligned currents, those flowing along the magnetic field. Figure 2 shows the field-aligned currents and the horizontal ionospheric currents that flow across the magnetic field. The current loop is closed by magnetospheric currents not indicated in the figure.

Why do the field-aligned currents result in an aurora? Because of their low mass, electrons generally carry those field-aligned currents. And because of their negative charge, electrons precipitating into the atmosphere carry the currents out of the ionosphere. Not all the electrons gyrating along a field line at high altitudes will enter the ionosphere. As an electron moves into a region of increasing magnetic field strength, it senses an induced electric field that increases its gyrational energy. A magnetic field, however, cannot increase the total energy of a particle, and so the gain in gyrational energy must come at the expense of the parallel energy. Thus an electron moving into a region of increasing magnetic field strength will be reflected at that point where the gyrational energy equals the initial total energy, the so-called magnetic mirror.

Only the electrons that have very low gyrational energies at high altitudes will avoid being reflected before they hit the atmosphere. Those electrons occupy a narrow "loss cone" about the magnetic field direction and carry a so-called thermal current. If the Lorentz force imposed on the ionosphere by the horizontal currents requires field-aligned currents much larger than the thermal current, then a parallel electric field will arise that accelerates the electrons into the loss cone. Hence, one observes the discrete aurora as subsequent collisions occur.

The aurora is complex and fascinating. This Quick Study emphasizes one aspect of its formation: the imposition of force by one plasma onto another to generate flow and the required field-aligned currents. In the simplest case, magnetospheric flows drive ionospheric flows. But other forces, such as pressure gradients, can also drive field-aligned currents.

In the future, physicists may fully understand how the aurora works as an electromechanical system. But as anyone who has ever gazed on an aurora will attest, Nature's dancing lights will always be a source of wonder.

## Additional resource

▶ NASA THEMIS (Time History of Events and Macroscale Interactions during Substorms), education and public outreach site, http://ds9.ssl.berkeley.edu/themis/mission_mystery.html. ■

*The online version of this Quick Study links to an essay describing auroras on other planets.*