



Editorial: Topical Collection on Auroral Physics

David J. Knudsen¹ · Joseph E. Borovsky² ·
Tomas Karlsson³ · Ryuho Kataoka⁴ ·
Noora Partamies⁵

Accepted: 10 January 2021 / Published online: 2 February 2021
© The Author(s) 2021

1 Introduction

The term “aurora borealis”, or “northern dawn” dates back centuries and refers to emissions of light from the otherwise-dark nighttime atmosphere, usually occurring in the polar regions except during highly disturbed periods. Aurora is distinct from airglow, which is a weak and relatively unstructured emission in the thermosphere caused by chemical reactions and ionization driven primarily by solar UV illumination during the day. In contrast, auroral emissions are the result of excitation of neutral atoms and molecules in the upper atmosphere by collisions with charged particles, typically which originate in the magnetosphere and precipitate along geomagnetic field lines with energies of hundreds of eV to tens of keV.

Popular descriptions of the aurora including in the media, dictionaries, encyclopedias and even textbooks often claim that the aurora is caused by “particles from the sun striking the upper atmosphere”. It is well established in auroral science that such a description is not accurate except in very limited cases. While it is certainly true that particles from the sun – the solar wind – provide the energy that drives the aurora, the widely varying morphologies and behaviors of the aurora are the result of a complex chain of events that take place within Earth’s magnetosphere or at its boundary, the magnetopause. The end result of this chain – excitation of neutrals by charged particles – is also well-established fact, as are many other

Auroral Physics

Edited by David Knudsen, Joe Borovsky, Tomas Karlsson, Ryuho Kataoka and Noora Partamies

✉ D.J. Knudsen
knudsen@ucalgary.ca

¹ Department of Physics and Astronomy, University of Calgary, Calgary, Alberta, Canada

² Center for Space Plasma Physics, Space Science Institute, Boulder, CO, USA

³ Space and Plasma Physics, KTH Royal Institute of Technology, 10405 Stockholm, Sweden

⁴ National Institute of Polar Research, 10-3 Midori-cho, Tachikawa, Tokyo 185-0031, Japan

⁵ Department of Arctic Geophysics, The University Centre in Svalbard, Longyearbyen, Norway

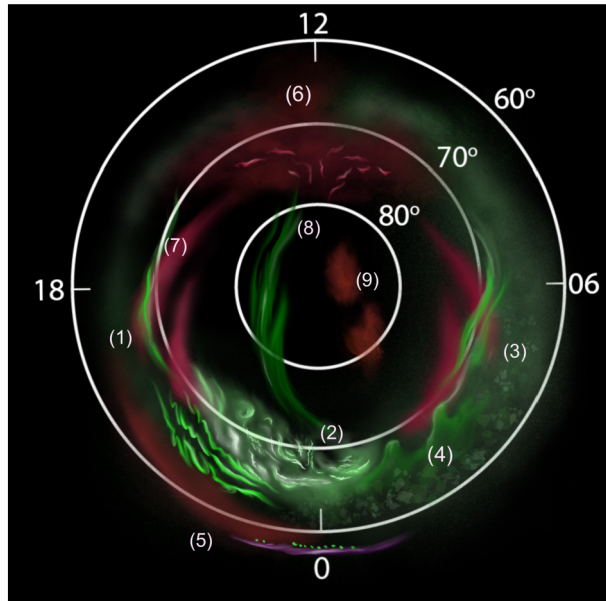
aspects described in this collection. However, due to the vast region of the magnetosphere that is magnetically conjugate to the auroral ionosphere, and the difficulty in sampling it with a single or even multiple spacecraft, the nature of the magnetospheric generator responsible for driving individual auroral forms still remains one of the most elusive aspects of the aurora.

This collection is designed to provide a comprehensive review of our current scientific understanding of the terrestrial aurora, both observational and theoretical, with an emphasis on developments since a previous collection devoted to the terrestrial aurora, “Auroral Plasma Physics” (*Space Science Reviews* volume 103, 2002, by Paschmann, Haaland, and Treumann). That collection emphasizes introductory and background material which is not repeated in the current set.

Organization of this collection is by type of aurora, summarized in Figure 1. State-of-the-art observations of each are summarized and compared to theories, while identifying open questions in obtaining a complete scientific understanding of the mechanisms responsible for their formation and behavior. As summarized below and documented in the individual reviews, while significant progress has been made over the past two decades, the remaining gaps in understanding are significant and in many cases fundamental.

- (1) *Quiet discrete arcs* are featured in three separate reviews in this collection covering observations (Karlsson et al.), acceleration mechanisms (Lysak et al.), and generator theories (Borovsky et al.). These distinctive curtains are frequently visible in the evening and midnight sectors of the auroral zone ($\sim 60\text{--}70$ degrees magnetic latitude), and are perhaps the most common and recognizable auroral form. Quiet arcs have characteristic time scales of several to tens of minutes and are associated with electric fields both perpendicular and parallel to the geomagnetic field, as well as field-aligned electric currents above the ionosphere and strong horizontal currents within it. At intermediate distances above the auroral ionosphere (up to ~ 2 earth radii), electrons are accelerated downward along the geomagnetic field to energies of a few to tens of keV, powered by quasi-static electric fields and currents imposed from still higher altitudes, by the so-called auroral generator. Presently, the magnetospheric auroral generator is the weakest link in terms of our understanding of this phenomenon, owing to the vast region of space magnetically conjugate to these arcs and the difficulty in sampling it with a single – or even multiple – spacecraft.
- (2) *Small-scale dynamic aurora* (Kataoka et al. in this collection) is characteristic of auroral breakups, where energy stored in the stretched magnetotail is released episodically into the auroral zone, initially close to midnight but spreading rapidly poleward and to the west and east. Aurora associated with these transient events has characteristic spatial scales of the order of 0.1 – 10’s of km, time scales of 0.1 – 10’s of seconds, and includes rapidly evolving arcs and flickering aurora. Rapid developments in camera technology in the past two decades have brought vast improvements in our ability to resolve and study these structures and their causes. Electric and magnetic field fluctuations associated with these spatial and temporal scales are in the domain of inertial Alfvén waves (IAW), which can account for many properties of electron acceleration observed in situ. However, as with the case of quiet discrete arcs, the origin of the IAW and the basis of the distinctive morphologies they produce remain important open questions. In this paper, a number of movie examples are included and are important to demonstrate the complexity and coherence of small-scale dynamic auroral features.
- (3) *Diffuse and pulsating aurora* (Nishimura et al. in this collection) is generally less structured and dimmer than the discrete aurora but can be more widespread in local time, and in principle may deposit more energy into the neutral atmosphere overall. Unlike

Fig. 1 A sketch of auroral forms reviewed in this collection, including (1) quiet discrete auroral arcs, (2) small-scale dynamic aurora, (3) diffuse and pulsating aurora, (4) mesoscale dynamical auroral forms, (5) proton aurora and sub-auroral forms, (6) dayside aurora, (7) field-line resonance arcs, (8) polar cap aurora, and (9) polar cap patches. Figure courtesy of D. Megan Gillies with contributions from Eric Donovan. After the original concept by Akasofu (*Space Science Reviews*, 19, 169, 1976)



their counterparts in the discrete aurora, electrons causing the diffuse aurora show no evidence of field-aligned acceleration, but rather are the result of pitch-angle scattering by plasma waves into the atmospheric loss cone by electron-cyclotron-resonance (ECH) or chorus whistler waves in the inner magnetosphere, on mostly dipolar field lines. This review documents significant advances in the past two decades including a direct connection of waves measured in situ with individual pulsating patches, providing, for the first time, definitive measurements of an auroral form as seen from the ground made simultaneously with measurements deep inside the magnetospheric source region. Still, many open issues remain, including the physical mechanisms that determine the spatial and temporal structure of pulsating auroras, the relative role of ECH and chorus waves, and the cause of small-scale sub-structure within the diffuse aurora.

- (4) *Mesoscale dynamical auroral forms* (Forsyth et al. in this collection) have spatial scales of tens to hundreds of km, and include poleward boundary intensifications (PBI), streamers, beads, omega bands and torches, and giant undulations. These structures can be interpreted as visible maps of mesoscale and large-scale dynamics of the nightside magnetosphere. PBIs and streamers are associated with the poleward boundary of the auroral oval, and are interpreted in terms of reconnection in the distant magnetotail in the case of PBIs, and reconnected flux tubes and bursty bulk flows propagating earthward in the case of streamers. The connections between these phenomena have been strengthened considerably in the past twenty years due to an increase in the number of conjunctions between satellites deep in the magnetosphere, and closer to the ionosphere, along with ground-based and space-based imagers. Omega bands and giant undulations are structures that form in the equatorward part of the auroral zone, associated with the diffuse aurora. Kelvin-Helmholtz instability from shear flows in the inner magnetosphere likely plays a role, however many complexities are not yet explained, for example irregular evolution and the formation of poleward-pointing fingers (torches). The story is similar with beads, which are periodically-spaced structures that form along the thin “breakup arc” situated on the equatorward edge of the auroral zone immediately prior

- to substorm onset. Beads are clearly a signature of a plasma instability in the magnetosphere associated with substorm onset, however there remains no scientific consensus on the specific mechanism or mechanisms responsible.
- (5) *Proton aurora and subauroral forms* (Gallardo-Lacourt et al. in this collection) reviews morphologies just equatorward of the auroral zone that include detached arcs and spots (both ion and electron), stable auroral red (SAR) arcs, and the newly-identified phenomenon “STEVE”. These phenomena map to and beyond the outer boundary of the plasmasphere, where dipolar field lines begin to transition to a more stretched configuration, and energetic ring current ions interact with low-energy plasma. These processes give rise to proton scattering and subsequent precipitation into the upper atmosphere, as well as heating of ionospheric electrons that is believed to be the direct cause of the sub-visual SAR arcs. A possible cause of STEVE’s distinctive, mauve-coloured emission is the interaction between neutral particles and ions driven to hypersonic speeds by large convection electric fields originating in the inner magnetosphere. In this sense STEVE does not fit the usual definition of aurora in terms of charged particles precipitating from the magnetosphere, however it is often accompanied by a bright, spatially periodic feature in green-line emission known as the “picket fence”, which may be associated with electron precipitation. Remaining open questions in this area include the specifics of proton scattering mechanisms, the source of electron heating in SAR arcs, and excitation and emission mechanisms within STEVE.
 - (6) *Dayside aurora* (Frey et al. in this collection) includes structures near and surrounding the dayside cusp region observed with space-borne auroral UV imagers capable of suppressing the strong background of solar illumination, or with ground-based cameras operated during the dark polar winter. Auroral phenomena in this region include high-latitude dayside auroral structures (HiLDAS), poleward-moving auroral forms (PMAFs), travelling convection vortices (TCVs), throat aurora, dayside diffuse aurora, and diffuse auroral spots. Many of these phenomena are associated with or driven directly by dayside magnetic reconnection and are used as a valuable diagnostic to complement in situ studies of the dayside magnetopause. As with other auroral types, significant progress in the past twenty years has been driven by advances in camera technology and availability of in situ observations. Remaining questions include the specific relation between dayside auroral forms and the morphology of the dayside X line (e.g. spatial extent, pulsed versus steady), and the role of the interplanetary magnetic field and other parameters.
 - (7) As reviewed by Rankin et al. in this collection, interhemispheric *field-line resonances* (FLRs) were proposed in the 1990’s as a mechanism for the generation of discrete auroral arcs, however this explanation was not widely appreciated due to the relatively low energies (predominantly sub-keV) predicted by the theory of dispersive Alfvén waves, especially at such low frequencies, and because of the fact that only a minority of arcs were observed to oscillate at relevant frequencies. More recently, observations using a new generation of sensitive red-line (630 nm) cameras have been used to identify hundreds of examples of oscillating arcs. Many aspects of the overall morphology and behavior of FLR arcs can be explained by fluid simulations, however more sophisticated models of field-aligned electron acceleration are needed, and are being developed, in order to explain observed electron energies.
 - (8) *Polar cap aurora* (PCA) (Hosokawa et al. in this collection) tends to be sun-aligned on average, can drift in the dawn-dusk direction, and at times can extend across the entire extent of the polar cap to form the so-called theta aurora. PCA occurs almost exclusively during geomagnetically quiet periods when the interplanetary magnetic field is

northward, and is generally interpreted as a tracer of large-scale magnetospheric topology. Recent scientific progress has been made possible due to rapid advances in both ground and space-based camera technology, along with MHD-based modeling. Outstanding questions include the relation of PCA to open versus closed field lines, magnetospheric source regions, generation mechanisms, cause of dawn-dusk motion, and the possible role of substorm triggering by polar cap arcs that encounter the poleward boundary of the nightside auroral zone.

- (9) *Airglow Patches in the Polar Cap Region: A Review*, by Hosokawa et al. in this collection. Historically, the term “airglow patch” was based on the assumption of no particle precipitation, since polar cap patches, typically defined as two-fold increases in ionospheric plasma density, are formed by sunlit plasma on the dayside that convects antisunward into open field line regions, and in principle may not require an active source. However, recent findings based on conjunctions of the Swarm satellites with ground-based cameras have demonstrated that some polar cap patches are in fact associated with enhanced convection and field-aligned currents, and therefore presumably with electron precipitation. We note that the study of active driving of polar cap patches through precipitation is still in its early stages.

Acknowledgements Planning for this Topical Collection took place during a workshop hosted in August 2018 by the International Space Science Institute ISSI-Bern, and attended by more than 40 members of the international scientific community.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.