History of Geophysics 7

Discovery of the Magnetosphere

C. Stewart Gillmor
John R. Spreiter
Editors

American Geophysical Union
Washington, DC

History of Geophysics Volume 7
Published under the aegis of the AGU History Committee

Library of Congress Cataloging-in-Publication Data
Discovery of the magnetosphere / C. Stewart Gillmor, John R. Spreiter, editors.
p. cm -- (History of geophysics ; 7)
Includes bibliographical references
ISBN 0-87590-288-X
1. Magnetosphere. I. Gillmor, C. Stewart, 1938- .
II. Spreiter, John R. III. Series: History of geophysics ; v. 7.
QC809.M35D58 1997
538'.766--dc21 97-11447
CIP

ISSN: 8755-1217
ISBN: 0-87590-288-X

Printed in the United States of America

American Geophysical Union
2000 Florida Avenue, N.W.
Washington, D.C. 20009

CONTENTS

CONTENTS

PREFACE

The beginnings of magnetospheric physics were the beginnings of space physics, of the marvelous discoveries made from in situ measurements from rockets and satellites and from increasingly sophisticated ground-based measurements and computer-assisted theoretical and empirical research. The beginnings of magnetospheric physics are also intimately connected with the International Geophysical Year 1957-58, the greatest world-wide cooperative scientific event in history. From the period following World War II until the late 1960s, the United States, and world physics and engineering in general, entered a new level of large-scale research epitomized by "space physics."

Covering the period roughly 1958-1967, this volume contains personal accounts from those pioneers whose pathfinding research initiated and solidified the field of magnetospheric physics. Here are accounts of the first rocket and satellite studies, of the discovery of the magnetosphere and Van Allen belts, of early models of the physics of the space around our Earth and of the Earth's environment within the Sun's plasma. Studies of the magnetosphere of the Earth led directly to our knowledge of the plasma environment around other planets and throughout our solar system. The authors of papers in this volume were in at the beginning, pioneers who played a significant role in the early years of magnetospheric physics.

The idea for this volume grew out of a session on the "History of the Discovery of the Magnetosphere," which was organized by Ed Cliver and George Siscoe and held December 15, 1995, during the AGU Fall meeting in San Francisco. At the session, Stewart Gillmor gave a historical introductory presentation, and interesting papers were given by James Dungey, Syun Akasofu, Robert Helliwell, John Spreiter and Alex Dessler. Carl McIlwain pitched-in for James Van Allen, who at the last moment was unable to attend, and Daniel Baker concluded the session with a synopsis of outstanding questions in contemporary magnetospheric physics. Because the session format limited the number of speakers, Cliver and Siscoe proposed that a monograph be produced to include a larger number of contributors. Gillmor and Spreiter agreed to edit such a volume.

We knew that even a volume could not include all those who were important players in the first years of magnetospheric physics. Cliver, Siscoe and others aided us in the selection of contributors. We hoped for a cross-section of authors, a few who were in their youth in the early 1960s, and who thus may have played supporting roles, and others who were early leaders. We aimed to include some contributors from beyond the U.S. borders. Thermal plasma specialists, particle people, theorists, instrumentalists, ground-based, rocket- and satellite-based, representatives from the academy, industry and government: these were some categories we wanted to include. Response was very encouraging, although several potential contributors regretted that they would be unable to join the project. The project was supported solidly by the AGU History Committee.

With one exception, all papers here have not previously been published. We asked authors to tell about specific ideas, actions and events in which they were involved in the late '50s through the late '60s and which directly involved the magnetosphere. We invited them to give to the reader personal perspectives on their work and that of their research groups, including comment on interactions with colleagues in other groups, those with whom they agreed and disagreed. We also asked for details about early education and mentors along the way, and some of the most vivid portions of this volume appear in such passages. Some authors knew early on that they wanted to do geospace physics, but others started out as engineers, mathematicians, and even as political scientists and musicians! In order to assist the reader to learn more about the history of magnetospheric physics, we invited each author to include in the bibliography a list of his or her additional publications pertaining to history, and we asked for a personal "period" photograph taken during the early days of magnetospheric physics.

The literature on history of the magnetosphere and related areas is scattered. We do want to note here some related publications: The *Journal of Geophysical Research-Space Physics* published special sections entitled "Pioneers of Space Physics" in 1994 (Vol. 99, No. A10, pp. 19099-

19212) and again in 1996 (Vol. 101, No. A5, pp. 10477-10585), and these were subsequently published separately as *Pioneers of Space Physics* (ISBN 0-87590-847-0) and *Pioneers of Space Physics 2* (ISBN 0-87590-890-X) and are available from AGU. The first collection contains papers by D. R. Bates, N. Fukushima, H. Friedman, E. N. Parker, J. A. Simpson, C. P. Sonett, J. W. Dungey, and W. I. Axford. The second collection contains papers by J. A. Van Allen, N. F. Ness, H. E. Petschek, F. B. McDonald, S.-I. Akasofu, G. Haerendel, T. M. Donahue, D. M. Hunten, and S. Kato. These collections indeed go beyond the subject of magnetospheric physics and include aeronomy, cosmic rays and matters well past the years of the "early days," but they should be of interest to the inquiring magnetospherist. And we want to note two valuable histories by David P. Stern: "A Brief History of Magnetospheric Physics Before the Spaceflight Era" (*Reviews of Geophysics*, *27*, 103-114, 1989) and "A Brief History of Magnetospheric Physics During the Space Age" (*Reviews of Geophysics*, *34*, 1-31, 1996).

Finally, we thank Ed Cliver and George Siscoe for getting the project started and for continued support during the development of the project.

C. Stewart Gillmor
John R. Spreiter
Editors

The Formation and Early Evolution of Studies of the Magnetosphere

C. Stewart Gillmor

Wesleyan University, Middletown, Connecticut
and Colorado School of Mines, Golden, Colorado

This paper will present a brief overview of observations, experiments and theoretical study of natural geoplasmas including the ionosphere and aurorae, and also of other geophysical and solar phenomena which led to the formation of the discipline of magnetospheric physics. Some of the researches and events only very briefly described here are discussed in greater detail in other papers in this volume.

INTRODUCTION

As well stated by James Van Allen, "The scientific heritage of magnetospheric physics lies principally in studies of geomagnetism, aurorae, and the geophysical aspects of cosmic radiation and solar corpuscular streams" [*Van Allen,* 1983, 9]. In the contemporary international use of the terms, magnetospheric physics is a part of geophysics, although we now know that there are magnetospheres on some of the other planets of our solar system. But today's magnetospheric physics evolved from different scholarly ancestors in various countries. In the United States until recently, geophysics signified only structural studies of the solid Earth. Since radio techniques so influenced work in the United States in the first two-thirds of this century, upper-atmosphere geophysics was studied predominantly in government and industrial laboratories and in university departments of electrical engineering and applied physics. In France, geophysics essentially included geomagnetism, as was the case in Germany and in countries influenced by Germany in the nineteenth century. Japan followed the German model, conducting upper-atmosphere geophysics research from physics and geomagnetism centers until, following World War II and the American occupation, such research shifted to electrical engineering research centers [*Gillmor*, 1973b]. In the United Kingdom and some Commonwealth countries, upper-atmosphere geophysics research was pursued by physicists and quite often by mathematicians in departments of *mathematics and theoretical physics*. In Australia, mixed alliances to American engineering and to British physics from the 1920s and after saw upper-atmosphere research and later radio astronomy emerge both from physics and from electrical engineering departments [*Gillmor*, 1991]. With the arrival of the *space age* in the late 1950s, a considerable portion of upper-atmosphere geophysics research in the United States evolved from physics and applied physics departments performing rocket studies of solar and cosmic-ray physics. And many of the instruments and techniques later used in magnetospheric physics--electronic circuits, computer techniques, remote sensing, even special halogen-quenched geiger counters--[*Friedman*, 1994] were first developed by physicists and engineers working in other areas of science and technology. But tit for tat: In the case of ionospheric research, the first studies in radio astronomy grew out of attempts to survey radio frequency noise sources hampering transat-

Discovery of the Magnetosphere
History of Geophysics Volume 7

lantic ionospherically propagated radio communications. And other ionospheric and magneto-spheric workers developed computer techniques and software and for some of them the techniques and design of the computer became so interesting they never returned to geophysics.

A diversity has existed also in the nomenclature for the Earth's upper atmosphere and geospace, a region characterized by its emphasis on naturally occurring plasma. We consider a huge volume, at laboratory vacuum pressure, primarily acted upon by the terrestrial magnetic field and by energetic radiations from the Sun and from cosmic rays. The lower ionized part of the region, from roughly 50 to 1000 km above the Earth's surface was termed the *ionosphere* in 1926, by Robert Watson Watt [*Gillmor,* 1976]. Thomas Gold coined the term *magnetosphere* in 1959 for the region extending from the upper part of the ionosphere out to the extent of the Earth's environment which is dominated by the geomagnetic field. We now know that the tail of the magnetosphere extends out beyond the Moon's orbit. The region of the ionosphere out to several earth radii is also sometimes called the *plasmasphere.* This provides a second name for the inner portion of the magnetosphere, which is confusing, but typical of scientific fields where more than one specialty can claim expertise.

In wishing to unify the physical and chemical study of the upper atmosphere, Sydney Chapman in 1950 [*Chapman,* 1950] termed such study of the ionospheric region, *aeronomy.* In many cases, the ionosphere and magnetosphere were studied by the same people using similar methods. To give us more confusion, the Earth's upper-atmospheric regions have been classified differently depending upon whether temperature, composition, state of mixing, gaseous escape or ionization is deemed more central to the work at hand, or to the scientific community at work. Thus the temperature-classified *troposphere* and *stratosphere* are located lower than, but the *mesosphere* and *thermosphere* overlap, the *ionosphere* and *magnetosphere.* In the 1950s and 1960s, most scientists spoke interchangeably of *space physics* and what has been called here *upper-atmosphere geophysics.* As instrumentation advanced and propulsion vehicles came to be more powerful, space physics came more to identify either those fields of physics and allied areas utilizing rockets and satellites, or those physical locations further and further from the Earth's surface. In 1915, *short* radio waves were waves of less than several hundreds of meters in wavelength; by 1940 *short* radio waves were less than 40 or perhaps less than 10 meters in length. In a similar way, *space* as a defining term moved higher and higher above the Earth's surface, from the *stratosphere* in the early part of this century, to the *ionosphere* in the inter-war period, to the *magnetosphere* by the early 1960s, then to the intra-solar system, then on to *Deep Space* as American television science fiction programs called it beginning in the 1970s [*Brush and Gillmor,* 1995].

Similarly, the community of those scientists studying solar-terrestrial relations, or solar terrestrial physics shares many citizens with the upper-atmosphere and geospace community. And the *magnetosphere* community is squarely in the middle of all this, particularly since magnetospheres are found on some other bodies in our solar system. Although the lower atmosphere sciences have come closer to the upper-atmosphere and geospace field in the last couple of decades, the latter seems always to have had a stronger tie to plasma physics, solar physics and astronomy than to classical neutral air meteorology.

EARLY GEOMAGNETISM

Magnetism is commented upon in ancient Asia, and in the Christian Medieval West but for its full introduction to Europe one usually cites William Gilbert's *De Magnete* [1600] where the loadstone and its analogies to the Earth are discussed in detail, indeed "...the terrestrial globe is magnetic and is a loadstone" [*Gilbert*, 1600, Book I, Chap. XVII]. In the seventeenth and eighteenth centuries some further progress was made in recognizing the variation of the earth's magnetic field and in some correlations with geophysical phenomena such as the aurorae. The great Edmond Halley even magnetically surveyed the Atlantic Ocean from his ship the *Paramore* in observations made between 1698 and 1700 and was the first to adopt isogonic lines [*Gillmor,*1990]. But until the late eighteenth century, the exact physical sciences comprised, at most, theoretical mechanics, astronomy and geometrical optics. The fields of heat, light, electricity and magnetism were descriptive and empirical, rather than analytical. During the course of that century, the empirical physical disciplines began to develop into exact and quantitative sciences. The period from about 1775 to 1825 was the making of natural philosophy into physics and was an exciting time [*Gillmor*, 1971, 42]. Pierre Simon Laplace, Charles Augustin Coulomb, Antoine Laurent Lavoisier and other physicists and chemists of the time recognized this. Joseph Louis de Lagrange recognized it as well when he wrote to Jean Le Rond d'Alembert in 1781 dispiritedly: "Physics and chemistry now offer riches more brilliant and easier to exploit; in addition, the taste of the century appears to be entirely aimed in this direction, and it is not impossible that the chairs for Mathematics [*Géométrie*] in the Academies will one day occupy the same insignificant position that the University chairs in Arabic occupy at present." [*Lagrange*, 1867-1892, XIII, 368]. The discovery of electro- magnetism and its phenomena and laws by Hans Christian Oersted, André-Marie Ampère, William Faraday and others in the first quarter of the nineteenth century would lead to important developments in geomagnetism.

It was not only magnetism, electricity, heat and so forth which blossomed in the late eighteenth and nineteenth centuries. Certainly a major paradigm or model for the physical sciences was Newtonian rational mechanics, particularly as applied to celestial mechanics. But during the eighteenth century, some scientists turned to problems of the earth and its environment for their exemplars (and this continued throughout the nineteenth century), and many physicists and mathematical physicists considered geophysical problems as integral portions of their life work. John Dalton, William Herschel, Hermann Helmholtz, Jean Baptiste Fourier, Lord Kelvin, James Clerk Maxwell and others were inspired to treat numerous problems in geophysics and atmospheric and solar physics [*Gillmor*, 1975a]. Fourier's interest in geophysics was so strong that he wrote "The profound study of Nature is the most fertile source of mathematical discoveries...", "...Mathematical analysis is as extensive as Nature itself..." and "The question of the terrestrial temperatures has always appeared to me as one of the grandest objects of cosmological study and I had it principally in mind when establishing the mathematical theory of heat." [*Fourier*, 1822].

Joseph Banks, perhaps the most influential British scientist of the times around 1800, and Alexander von Humboldt championed exploration as well as geophysics in all its recognized specialties. The British astronomer royal George Biddell Airy counseled Cambridge University to add more studies of geophysics, including geomagnetism, and solar physics to the curriculum and the examination requirements of the university. He was distressed that the B.A. honors exams at Cambridge stressed "pure" mathematics to such an extent that in 1866 he urged that additional subjects including "Magnetism, terrestrial and experimental, and their connection..." be added [*Airy*, 1896, 267-9]. Airy worked constantly to obtain support for observatory and field work in geodesy and terrestrial magnetism. The great polymath, Carl Friedrich Gauss, was inspired in his research in spherical harmonics by his desire to model the Earth's geomagnetic field. Gauss and his colleague Wilhelm Weber established a series of geomagnetic observatories and this work was expanded by Edward Sabine [*Cawood*, 1979; *Stern*, 1989]. Sabine was so impressed by Gauss' work in organizing international cooperation in geophysical measurements that he urged the British Association and the Royal Society to support the extension of geophysical studies to the ends of the British Empire. Sabine was also the first to propose that an expedition be mounted to the southern polar areas [*Gillmor*, 1978]. The expedition Sabine proposed, to be led by James Clark Ross, had various scientific objectives but its main work "beyond all question ... and that which must be considered as, in an emphatic manner, the great scientific object of the Expedition, is that of Terrestrial Magnetism ..." [*Royal Society*, 1840]. This was to include simultaneous observations to be undertaken on certain world term days, determination of the position of the south magnetic pole, and special attention to aurorae.

Gauss showed that the great preponderance of the Earth's magnetic field was from internal origins. But he suggested that some magnetic variation could be from electric currents flowing high above the Earth. Eighteenth-century ideas about geomagnetism and occurences of solar and auroral events intensified in the mid-nineteenth century and correlations between the sunspot cycle, solar flares, the aurorae and geomagnetism and geomagnetic *storms* were established. A notable example of this was the observation and correlation by *Carrington* [1860] of a solar flare, a simultaneous magnetometer deflection and a closely following geomagnetic storm and auroral display. By the last third of the nineteenth century, scientists generally accepted the relationship between the solar cycle and magnetic disturbances on the Earth. Balfour *Stewart* [1882] showed that there are electric currents flowing in the Earth's high atmosphere, that solar action must render the air there electrically conducting, and that the conductivity is higher at sunspot maximum than at sunspot minimum. Not all, however, agreed: Lord Kelvin in his later years, more and more frequently disagreed with scientific opinion, *e.g.*, on the age of the Earth, on the importance of Transatlantic radio, and on the connection between magnetic storms and sunspots [*Pomerantz*, 1974; *Stern*, 1989].

Confidence in the future of geomagnetic studies was shown by the courage of Louis Agricola Bauer when he founded the journal *Terrestrial Magnetism* in 1896 (which became *Terrestrial Magnetism and Atmospheric Electricity* in 1899, and the *Journal of Geophysical Research* in 1949). Bauer was impressed with the importance of terrestrial magnetism both for geophysics and solar-terrestrial physics in general, and for broad applications to mapping, mining and navigation. The first article to appear in *Terrestrial Magnetism* was "On Electric Currents Induced by Rotating Magnets, and Their Application to Some Phenomena of Terrestrial Magnetism" by Arthur Schuster of Owens College at Manchester [*Schuster*, 1896]. Schuster had worked with James Clerk Maxwell and with Lord Raleigh in Cambridge and succeeded Balfour Stewart at Manchester in 1887. Schuster's broad research interests included spectroscopy, electron and X-ray studies, and the electrical behavior of gases. Schuster demonstrated that a small potential could maintain an electrical current, once a gas was ionized. He is best remembered by geophysicists for extending Stewart's ideas on geomagnetism and upper atmosphere physics. He believed that the upper atmosphere layer was conducting due to solar ionizing radiation and that convective motion was cause for the upper atmosphere electrical currents. He calculated the ionization and conductivity as a function of the solar zenith angle and, by 1908, had calculated an estimate for the specific conductivity of the layer as 10^{-13} emu and estimated the layer thickness as 300 km. He resigned his chair at Manchester in 1907 so that a young Ernest Rutherford could gain it, an act of generosity reminiscent of Isaac Barrow resigning the Lucasian Chair at Cambridge in favor of young Isaac Newton. [*Gillmor*, 1990, ix-x].

IONIZED GASES

Electrical researches by Rutherford's former teacher and Schuster's colleague, Joseph John Thomson, on electrically charged particles would spark interest in upper atmosphere geophysics in two major ways: It stimulated the existent research areas of geomagnetism and the aurorae, and the newly emerging area of radio communications and propagation. Thomson himself had been a student of Stewart and was grounded in the ideas that "...the only conceivable cause capable of operating in such [high atmospheric] regions must be an electric current... we know from our study of the aurora that there are such currents in these regions...convective currents established by the Sun's heating influence in the upper regions of the atmosphere are to be regarded as conductors moving across lines of magnetic force, and are thus the vehicle of electric currents which act upon the magnet." [*Stewart*, 1882, 181]. Perhaps some of the phenomena of geomagnetism and the high atmosphere of the earth could be related to the sun through streams of electrical particles such as those hypothesized in connection with the new experiments with electrical discharges in vacuum tubes and vessels.

Thus it was that laboratory work by Thomson and others on electrical discharge phenomena in gases doubly stimulated upper atmosphere geophysics: It provided a thematic base for building from the Scandinavian interest in magnetism and the aurorae, and, it similarly stimulated the new technical field of radio communication. In one sense, then, basic physics stimulated geophysics. But again, tit-for-tat: Change within geophysics allowed it to exploit research contributions of general physics while at the same time contributing to the development of physics as a whole, for example, with the work of Johann Elster, Hans Geitel, and Charles T. R. Wilson leading from geophysics to the future of the fields of cosmic ray and high-energy physics. The considerable importance of the nineteenth-century tradition of observational measurement in "terrestrial magnetism and atmospheric electricity" is seen also in the role of measurement, standardization, and international cooperation in geophysics. The establishment of the international geophysical networks contributed to the movement to establish national bureaus of scientific standards and measurement beginning around 1900 and to the broadening of the scope of international scientific expeditions [*Gillmor*, 1975a].

SOLAR STREAMS

Beginning in 1896, the Norwegian Kristian Birkeland considered that a stream of electrons from the Sun could reach the geomagnetic field with enough density to produce geomagnetic disturbances. He caused cathode rays to be projected towards a spherical magnet, which he called a "terrella". *Birkeland* [1901, 1908, 1913] demonstrated with beautiful photographs that a toroidal space was observed around his "terrella" and that many of the cathode rays (electrons) were directed towards the poles of the "terrella", leaving an equatorial belt and the toroidal space relatively free of the particles. Birkeland's theory was criticized by Arthur *Schuster* [1911] and by others. For one thing, Schuster argued that rays of electrons, being of the same charge, would be electrostatically dispersed. But Birkeland's ideas were influential on the Scandinavian school of auroral and magnetic researchers and he was instrumental in inspiring Carl Størmer. Størmer was a recognized expert in mathematics and in experimental physics. For example, he had gone further than anyone else in attempting to quantify auroral data using photography. Robert Scott, Douglas Mawson and Ernest Shackleton, over several expeditions, all attempted to photograph aurorae in the Antarctic and were unsuccessful until 1912 or 1913 [*Gillmor*, 1978, 239]. Edward Wilson noted in his Antarctic diary in 1912, "No one has ever got any results except one Norwegian (Størmer) who was consulted before we left England..." [*Wilson*, 1972, 126].

Størmer developed mathematically Birkeland's theory and experiments. He dealt with the possible paths of a single charged particle and showed that such a particle from the Sun can reach the Earth only in two narrow zones centered near the polar regions of the Earth. In 1907 Størmer published a diagram of the meridian projection of the trajectory of an electrically charged particle in a magnetic dipole field [*Størmer*, 1907]. As an example of the labor involved in such hand calculations, Størmer's female assistant required almost two years to produce the plot. Størmer devoted many years to studying the motion of electrically charged particles near magnetic dipoles. His work and that of Birkeland would provide inspiration for studies of solar and cosmic-ray particles and their interactions with the

Earth's magnetosphere. As Stern has indicated [*Stern*, 1989], Størmer succeeded to the extent of showing that a wide class of orbits in the dipole field were trapped, and that low energy particles arriving from a distant source were either steered to the polar zones or were turned away. But Størmer's theory worked better for high energy particles than for auroral particle energies. Both Sydney Chapman and Hannes Alfvén were strongly influenced in their theories of the Sun's effects on geomagnetic storms by the earlier auroral studies of Birkeland and of Størmer.

SOLAR LINKS TO GEOPHYSICS

From the late nineteenth century then, networks of stations had concentrated on studying the very small changes in the surface geomagnetic field which occur nearly continuously in time. These changes were most often recorded with ink on paper rolls or on photographic paper. Some of the features of the tracings were given the names "bays", "crochets", and "spikes", since they resembled to the human eye the outline, for example, of a marine bay coastline on a map. Study of the changes indicated that "storms" occurred and that they could affect large portions of the Earth. They seemed to re-occur roughly at periods of 27 days, the synodic period of rotation of the Sun. Certain phenomena, such as the sudden radio fade-outs to be discovered later by the ionospheric researchers in the 1930s, occurred in daylight shortly after some solar flares. Most geomagnetic storms, however, seemed to lag solar flares by about two days. In spite of the criticisms of the idea of streams of charged particles proceeding from the Sun to the Earth, it was assumed that these streams of ions or electrons travelled towards the Earth from time to time, at speeds of some hundreds of kilometers per second. Carl Størmer posited these solar streams coming to the Earth and causing the aurorae [*Størmer*, 1930a,b]. Thomas L. Eckersley, a pioneer in the demonstration in 1921 of the height of the so-called Kennelly-Heaviside ionospheric layer, of the development of the magneto-ionic theory, and of early quantitative studies of whistlers and VLF, thought at one time that whistlers as well as long-delayed radio echoes could be caused by radio waves interacting with Størmer's clouds of charged particles in space [*Eckersley*, 1929]. (The long-delay radio echo is a phenomenon that was noted in the late 1920s, examined for several years and then dropped. Consideration of this problem began again in the late 1950s and has been pursued off and on since that time with no convincing mechanism having yet been accepted.) Edward Appleton wrote to B. van der Pol in 1930 that the two exciting things he looked forward to in his planned Second Polar Year radio expedition to Tromsø, Norway in 1932-33 would be the chance to study the ionospheric layer and also long-delay radio echoes at high latitudes [*Appleton*, 1930]. This, too, was based upon the idea that somehow solar particles were impinging upon the Earth's upper atmosphere and entering in the polar zones. Sydney Chapman thought that solar corpuscular streams were responsible for ionizing what came to be called the E layer in the ionosphere, but Appleton convinced him by 1933 that solar radiation was most likely the cause. Thus, for considerable years, many specific geophysical events were believed to be linked to sporadically or regularly arriving solar streams.

What about theories, then, for the action of such solar streams as they intersected the Earth and possibly caused geomagnetic storms? In 1926, E. A. Milne argued that neutral or ionized atoms moving upward through the solar atmosphere might continually be accelerated away from the Sun by selective radiation pressure and achieve velocities up to 1600 km s^{-1}. But how could atoms at such relatively low velocities then penetrate the Earth's upper atmosphere? Sydney Chapman had been considering such things for some time, aware that Frederick Lindemann had pointed out that the large quantity of solar corpuscules in a stream necessitated a largely neutral plasma. In 1927 Chapman continued his researches with the assistance of Vincent Ferraro, and in a series of papers [*Chapman and Ferraro*, 1931,1932,1933] suggested a neutral plasma cloud leaving the Sun and impacting upon the geomagnetic field, compressing the field on the sunward boundary at a rather sharp interface. Chapman and Ferraro explained the beginning of the geomagnetic storm (the *sc* or sudden commencement) as resulting from the impact of the plasma cloud. The initial phase of the geomagnetic storm was a result of the compression of the geomagnetic field by the continuing pressure of the impacting plasma cloud. They suggested that the main phase of geomagnetic storms was caused by a ring current which formed around the Earth at several earth radii and which then was dissipated through collisions with the ambient atmosphere [*Parker*, 1997]. Thus their ring current was a transient phenomenon. The geomagnetic field had carved out a hollow space around the earth, but the solar particles could reach the upper atmosphere along two "horns" extending into the polar regions. During the main phase of geomagnetic storms, the resultant radio interference, auroral phenomena, etc, would be produced by charged layers induced on and escaping from the surface of the hollow. Qualitatively their theory fit most of the facts, though Chapman wrote in 1940 "Unless some undiscovered mechanism exists which imparts much greater velocities to the solar corpuscles, possibly only in the near neighborhood of the Earth, we must conclude that the Earth's atmosphere is more penetrable, down to 100, 80, or even 70 km, than is indicated by our present information..." [*Chapman and Bartels,* 1940, 810]. Chapman wrote for more than four decades on the causes of geomagnetic storms, a subject which continues to this day as an active field of research.

Notable alternative models for geomagnetic storms were produced contemporaneously with Chapman's by Hannes Alfvén. Alfvén also posited the development of a ring current around the Earth and envisioned a weak magnetic field accompanying the solar stream [*Alfvén,* 1939, 1940]. The solar particles in the stream entered the Earth's geomagnetic field convected by an electric field due to the cloud's motion. [*Stern,* 1989]. Alfvén disagreed with Chapman's theory. Chapman disagreed with Alfvén, as indeed he had disagreed with much of Birkeland's ideas, even though influenced by Birkeland. The Scandinavians fared not too well: Alfvén's work was little appreciated for some years, and was published in a relatively unknown Swedish journal [*Dessler,* 1970]. Its reception was similar to that of the work of the Danish electrical engineer and radio physicist P. O. Pedersen who published the fundamental modelling of the theory of ionospheric layer formation four years before Chapman produced his so-called "Chapman Layer" theory [*Chapman,* 1931]. Pedersen's work [*Pedersen,* 1927] was published as a monograph in an obscure Danish engineering series, and although praised privately by Appleton, was otherwise ignored. Appleton later recognized Pedersen's work in general by naming the Pedersen Ray in his honor [*Gillmor*, 1986b].

RADIOPHYSICS

It was stated above that laboratory studies of electrical discharges in gases performed about 1900 stimulated the new field of radio communications. Indeed, I would argue that from about the time just after World War I until perhaps 1960 (roughly four decades), geophysics and radio communications mutually grew and benefitted. Geophysics (and more recently, radio and radar astronomy, and plasma physics) held onto radio communications and propagation as a tool to study the Earth's ionized atmosphere and near space. Syncretically, radio comunications techniques needed the knowledge from geophysics: What was the extent of the plasma above the Earth? How did it vary in time and in regions around the globe? What were its constituents, ions? electrons? What were the causes of noise outbursts, and radio storms and fadeouts? This partnership began changing in the 1960s: part of the interest in communications has headed toward environmental and remote sensing, and communications in the widest sense, utilizing satellites, fiber optics, etc. while the geophysics interests have become closer to a field concerned with planetary atmospheres and the physics of plasmas [*Gillmor,* 1986a].

Following Marconi's attempts to receive long-wave signals across the Atlantic in 1901 and 1902, various physicists tried to explain the phenomenon using the theory of diffraction as in optics. But diffraction could not nearly account for the bending of a radio wave passing over the ocean from England to Newfoundland. In 1902, Arthur Kennelly, and shortly afterwards, Oliver Heaviside, put forth ideas that the radio waves perhaps were reflected off a conducting layer located perhaps 80 km above the Earth. As discussed above, Gauss, Stewart, Schuster and others had invoked such ideas to explain the aurorae and geomagnetic field variations.

THE IONOSPHERE AND MAGNETO-IONIC THEORY

In 1912 W. H. Eccles made the first attempt to use quantitative physics to study radio waves propagating in a plasma. He assumed the effective charged particles were ions and that the ionized medium acted as a conductor, the waves therefore being reflected at a sharp boundary [*Eccles,* 1912]. A decade later, Joseph Larmor suggested that the collision frequencies of the ionized particles in the conducting layer (by then, called the *Heaviside layer*, or the *Kennelly-Heaviside layer*) were low enough so that the region acted as a dielectric and the radio waves could be refracted back to Earth. In addition, Larmor assumed the effective particles were electrons, not ions [*Larmor,* 1924]. About this time, Edward Appleton had been considering the effect the terrestrial magnetic field would have on rotating the plane of polarization of radio waves passing along field lines in the ionized upper atmosphere. Soon after Eccles' paper in 1912, Lee de Forest (1913) and Leonard Fuller (1915) in the U.S., had noted differential fading on frequency-shift-keyed transmitters of the Federal Telegraph Company and had attributed this to wave interference effects between the surface path and a path reflected off an upper conducting layer located perhaps 62 miles above the Earth [*Villard,* 1976]. T. L. Eckersley (1921) in England and others argued for the existence of an upper ionized layer following studies of fading and deviating of long-distance radio signals. These studies were not successful in bringing a majority to accept that such an ionized layer was responsible for long distance radio propagation. Edward Appleton and his student M. A. F. Barnett did succeed in convincing many scientists of the existence of an ionosphere in experiments beginning in late 1924, where a slowly changing frequency emitted by a BBC broadcast transmitter was used as a Lloyd's mirror interferometer to estimate a layer height of about 90 to 115 km [*Appleton* *and Barnett,* 1925a,b]. A few months later, Gregory Breit and Merle Tuve at the Carnegie Institution of Washington, with the assistance of the U. S. Naval Research Laboratory, successfully demonstrated the existence of the ionosphere using, in effect, pulse radar techniques [*Breit and Tuve,* 1925, 1926]. The pulse sounding technique eventually became the standard ionospheric tool. As to the history of the magneto-ionic theory, based upon the famous work of H. A. *Lorentz* [1909], early efforts by H. W. Nichols and J. C. Schelleng, A. H. Taylor and E. O. Hulburt, H. Lassen, T. L. Eckersley, and others were little recognized and pride of place went to Appleton, assisted by Douglas R. Hartree. The most well-known formulae for the refractive index of a wave propagating in an arbitrary direction in a magneto-plasma were known as the Appleton-Hartree equations and were published by *Appleton* [1932]. Unknown was the fact that Wilhelm Altar, an Austrian mathematical physicist, had worked with Appleton on magneto-ionic theory in England for six months during 1925-1926 and constructed the magneto-ionic equations in the so-called dielectric tensor form and in more general terms presented what would today be called the dispersion relation, or the equation for the wave-normal surface, in cold plasma theory. The magneto-ionic theory continued to be extended during the 1930s, especially with the work of H. G. Booker [*Gillmor,* 1982]. By the beginning of the satellite era, the standard treatments of the magneto-ionic theory included the monographs by J. A. *Ratcliffe* [1959] and K. G. *Budden* [1961].

WHISTLERS AND THE PLASMAPAUSE

In the 1930s and 1940s, the ionospheric workers were closely tied to geomagneticians, as mentioned above, in questions about solar effects on long distance propagation, sudden radio fadeouts, correlations of aurorae and radio propagation, and the effects of the Sun on maintenance of the various ionospheric layers. The previously cited works of P. O. *Pedersen* [1927] and S. *Chapman* [1931] are notable examples. A number of puzzling phenomena littered the playing field of early ionospheric physics: supposed strange propagation anomalies around the electron gyro frequency of about 0.5 Mhz in the Earth's field; the long-delay-echoes (wherein a short-wave radio signal may appear on a receiver after a delay time of say-- 0.1 second and then be followed by an echo or echoes delayed by as much as several seconds); sudden radio fadeouts over the entire sunlit portion of the Earth; reception of radio signals for brief periods at frequencies much higher than thought possible; and whistling and burbling tones heard on telephone lines and audio systems. Some of these problems were later explained, some remain unexplained, some were first noted and then laid aside for some years. The whistling/burbling sounds are in the latter category of early radio "problems", pondered and lain aside, then later reappearing to produce a discovery of the magnetosphere!

Whistlers are, in fact, naturally occurring non-acoustic waves in the *plasmasphere* and have descending (and rising) frequencies in the audio radio wavelengths at Very-Low- and Extremely-Low-Frequencies (VLF and ELF). They were first reported at the end of the nineteenth century by maintenance workers on telephone line systems and subsequently by a monitor eaves-dropping on Allied telephone conversations during World War I [*Barkhausen,* 1919, 1930; *Helliwell,* 1965, 1997]. The telephone lines were acting as receiving antennas for the waves. Similar reports [*Burton and Boardman,* 1933] even influenced John N. Dyer to monitor and record whistler activity on aluminum records during his stint as radio engineer for Richard Byrd's Second Antarctic Expedition (1933-1934) [*Gillmor,* 1978]. In 1935, Thomas L. Eckersley gave an explanation [*Eckersley,* 1935] of the whistling sounds in terms of the newly expounded magneto-ionic theory (on which he had composed several manuscripts by 1935 but was unable to get published until after World War II). Recall that in 1929 Eckersley had suggested that both whistlers and long-delay echoes could be caused by radio waves perhaps interacting with clouds of charged particles in space. Here, however, Eckersley noted that the magneto-ionic theory gave the required type of dispersion in the right frequency range for whistlers. He suggested that whistlers were produced by lightning discharge energy that had traveled great distances from the Earth (to a distance of some Earth's radii), along the lines of the geomagnetic field. This didn't make much sense at the time, since ionospheric workers and other physicists assumed the electron density fell virtually to zero some hundreds of kilometers above the Earth. Thus, the whistler would have had to emerge from a lightning stroke, proceed upwards along a magnetic field line and beyond the ionosphere into empty space, according to ideas at the time [*Booker,* 1974]. So Eckersley's extremely fruitful paper lay fallow for more than fifteen years. L. R. O. Storey, one of J. A. Ratcliffe's graduate students at Cambridge University, began some ingenious experiments and theoretical work on whistlers in the late 1940s and discovered the *plasmasphere,* the *thermal* magnetosphere, some years before the satellite discoveries [*Storey,* 1953]. Ratcliffe presented Storey's findings at an URSI (International Radio Science Union) meeting in 1952. Storey showed that the whistler data fit Eckersley's model if the whistler was assumed to travel out to a distance of about 3 or 4 Earth radii and that the electron density in that region was several hundred cm^{-3}! [*Helliwell,* 1997].

THE BEGINNINGS OF MAGNETOSPHERIC PHYSICS

Now at the end, a beginning: Cosmic ray studies in the 1930s and after World War II brought people expert in rockets and balloons ready to apply their knowledge and experience to what would be magnetospheric physics. R. A. Millikan, A. H. Compton and others, building upon the earlier work of V. F. Hess, L. Myssowsky, L. Tuwim, and the theoretical work of G. Lemaitre and M. S. Vallarta had made the basic nature of cosmic rays and their occurrence above and on the earth subjects of major scientific interest [*Gillmor*, 1978; *DeVorkin*, 1989] as will be discussed by first-hand participants in several other papers in this volume. In addition, geomagneticians and radiophysicists would benefit from new tools, new colleagues, and new means of support. The military, and in the U.S., especially the Navy and Air Force, were highly supportive, for rockets and radar could be used just as radio had been used decades before to pursue jointly geophysics and defense. Just after World War II, international geophysics took a quite new turn. Step functions occurred in sociometric terms: Data extracted from 3,485 articles on ionospheric and magnetospheric physics published from 1925 through 1960 in 195 different journals reveal that from 1946 through 1948, the percentage of articles identified by institutional address following the name of the author(s) on the papers jumped from 5% of the time to 85%. Similarly, U.S. authors of papers in ionospheric and magnetospheric physics publishing in the *Journal of Geophysical Research (JGR)* from 1945 through 1981 showed a step function in acknowledgement of outside funding for research, from about 30% in the immediate post-war years to 70% by 1950, rising steadily to almost 100% by 1981. The increase in multi-authored papers on ionospheric and magnetospheric physics published in *JGR* and in the *Journal of Atmospheric and Terrestrial Physics* showed a similar step function increase in the years 1958-1960. All these indices, strange as they may seem to physicists, indicate that "Big Science" was arriving along with the magnetosphere.

The National Science Foundation was organized in 1950, formed initially along Naval Research funding models; NASA followed in 1958 [*Gillmor*, 1986a]. The discovery of the magnetosphere arrived with the IGY, with the satellite, and with the computer.

Particularly effective was the International Geophysical Year 1957-58 (IGY), in this author's opinion the greatest cooperative scientific venture in history. The IGY provided several years of planning and introductory phases. The IGY was planned to occur during a solar maximum period, and it opened up the Antarctic for extensive and continuous study and extended geophysical measurements in space via satellites and rockets. It also established the philosophy and practice of World Data Centers (WDC) for the international collection, comparison, and sharing of data, and became a model for large-scale cooperative scientific programs. U.S. satellite work was planned to coincide with the IGY and was the largest single budget item; but ionospheric (and magnetospheric) physics was the largest ground-based specialty in terms of funding [*IGY*, 1965]. Who were the prime movers of the IGY? They included Lloyd Berkner, polar radio explorer, ionospheric physicist and science adviser to Presidents; Sydney Chapman; James Van Allen; Fred Singer; and others included among the authors in this present volume [*Nicolet*, 1983; *Van Allen* , 1983].

When I was an undergraduate at Stanford University in the late 1950s, I recall the largest computer as being the IBM 360, with a 2k memory drum. One of my ionosphere professors, Allen Peterson, was one of the first on campus to utilize that machine. It would be a very few short years before the days would be gone, of Carl Størmer's assistant spending two years by hand calculating a particle orbit in 1907 or of Edward Appleton's "IYLs" (Ionospheric Young Ladies) in the 1930s and 1940s who performed mathematical calculations on hand crank machines [*Clark*, 1971].

And, the *in situ* measurements which would be provided by satellites would virtually transform geospace physics. Surveys conducted by this author in 1971 and again in 1984 of ionospheric and magnetospheric physicists [*Gillmor*, 1986a], underpin the statements in this section.

This author now invites the reader to proceed to the papers written by Discoverers of the Magnetosphere.

Post Script:

I have been only the scribe of the efforts in this volume rather than one of the actors. And I am on the trailing edge of the pulse in the time domain, being about Fairfield's age and a little older than Southwood. Still, it has been suggested to me by several contributors that I might add a short piece about my own early years in ionosphere and magnetosphere studies, even though my research for the past thirty years has been largely in the *history* of ionospheric and magnetospheric studies.

I was interested in science from an early age, especially astronomy, although this metamorphosed into interests in chemistry and then radio. My father was a physician and thus I was able to collect a large amount of laboratory glassware and produce noxious odors and fumes from about the age of seven. I built a crystal radio set as a Cub Scout and then got my Radio merit badge as a Boy Scout. I became a ham radio operator by the age of thirteen and was particularly interested in antennas and wave propagation. I didn't know much about the difference between DC and RF and impedance matching, so that some of my home-made stuff passed the radio signal from thin wires to coaxial cable to wide-spaced feedlines and I modulated lots of things around the house. This resulted in my being banned from the air when *I Love Lucy* or *Rasslin' with Russ from the Chicago Amphitheater* came on our TV set in the early 1950s. I did lots of things as a kid, moving from Independence, Missouri to Texas, to Arkansas, to New York and to Massachusetts and then back to Missouri during and just after World War II as we followed my father in the Army Medical Corps. Music, sports and then girls also interested me. I became convinced that I wanted to study radio wave propagation in college, and I was fascinated with polar expeditions. An opportunity to visit Antarctica appeared when the Boy Scouts had a national competition to choose a boy to accompany a US expedition to Antarctica in the pre-IGY years. I got into the finals but didn't win.

I entered Stanford University in the Fall of 1956, planning to be an electrical engineering major, since all the radio and solar system work at Stanford was done in E.E. in those days. I soon met and came to work for or took courses with a very interesting group of faculty in the "RPL", the Radio Propagation Lab (later named Radioscience Lab, then STARLab): Ronald N. Bracewell, Von R. Eshleman, Robert A. Helliwell, Lawrence A. Manning, Allen M. Peterson, and O. G. "Mike" Villard. These contacts led me to become employed at Stanford Research Institute as well as on several projects at the University, for summer and part-time jobs. I participated in the *Argus* nuclear tests in 1958, working in the US and in the Azores Islands and got my introduction to the *riometer* ionospheric absorption recorder and to whistler recording equipment. While in the Azores I spent my spare time monitoring the radios and noted what I thought were broad-band HF noise bursts. When I returned to Stanford the next year, the RPL people suggested I investigate solar and planetary radio noise. They kindly loaned me lots of stuff, really a great amount of steel poles, motors, multi-element antennas, cameras, tubes, oscilloscopes, and a trailer and I built a 25-35 MHz sweep-frequency radio observatory, with a large Log Periodic Antenna array on a rotating 60-foot pole. Particularly helpful was L. H. "Bud" Rorden, an SRI engineer. I began to take data from my observatory in 1960, won some student research contests, and wrote a paper which was accepted for publication [*Gillmor,* 1962a]. Too many hours spent at the Stanford student radio station KZSU, renting our off-campus house out to dormitories for campus parties, and working on my research project didn't do a lot for my academic grades. But Bob Helliwell, a scoutmaster himself, knew of my interest in Antarctica and in early November 1960 he asked me if I'd like to go to Antarctica with the Sixth Soviet Antarctic Expedition. I left Stanford within three hours and flew for an interview with the National Bureau of Standards (NBS) in Boulder, Colorado. C. Gordon Little, Chief of the Upper Atmosphere and Space Physics Division at NBS had secured a grant from NSF to supply a "US Exchange Scientist" to accompany the Soviets to Antarctica, but the designated NBS employee became ill at the very last moment. The proposal was to operate riometers in the Antarctic. I succeeded in the interview, was hired by NBS, flew back to Stanford and told Felix Bloch (thermodynamics), Nina Byers (nuclear physics), and three other physics and engineering professors (in introductory quantum mechanics, and fluid mechanics, as I recall) that I couldn't take the mid-term exams that week since I was going to the South Pole with a hundred Soviets.

Philip M. Smith of the NSF's US Antarctic Research Program (USARP) in Washington, D.C. got for me odds and ends of polar clothing and saw me through a confusing couple of days of US State Department interviews and off to New York to join up with all my equipment and to board the weekly Pan American Airways milk-run flight to Africa. I flew to South Africa with 7000 lbs of gear in 41 wooden crates to meet the Soviet ice-breaking freighter *OBb* in Capetown. Only after much argument from me, Pan Am removed 23 passenger seats from the DC-7 so that all my cargo could fit on the plane. When I arrived in South Africa I was greeted by several reporters. It seemed that I had broken the "White Hunter Record", by entering with the largest excess baggage charge ever recorded. My "safari" excess baggage was 7000 pounds and the charge was $11,000. Now I was carrying an official US government passport onto a ship with about 200 people who spoke a language I didn't understand and I had received virtually no preparation for the voyage.

I arrived in the Antarctic in early December 1960 with 30MHz and 50MHz riometers from NBS for the station *Mirnyy* (which would be my primary base), a portable whistler station loaned to me by Helliwell, which I hoped to set up at the very cold inland station *Vostok* (at the South geomagnetic pole), and short-wave listening equipment loaned by Mike Villard. I had operated all these sorts of equipment before but I was somewhat uneasy: I spoke a little French, about 25 words of Russian, and I had not seen the inside of most of the 41 wooden crates of gear. It would be a wonderful 14 months and I was helped a lot by kind people, from tractor drivers to cooks to geophysicists. The scientific leader of the expedition, Valentine Driatsky, specialized in polar radiophysics. My best pals and helpmates on the expedition were an East German radiophysi-

est to the military had been mostly ignored (except for WWII radar and the "bomb"). In 1971 I began to visit geospace workers in many labs in the US, Europe and Asia. I corresponded with hundreds, visited archives, conducted an international mail survey, and converted a great amount of published bibliography to digital form for computer analysis. I have reported on this research in numerous articles [*Gillmor and Terman,* 1973; *Gillmor*, 1973; *Gillmor and Gran*, 1974; and *Gillmor,* 1975a, 1975b; 1976; 1978; 1981; 1982; 1984; 1985; 1986a; 1986b; 1987; 1989; 1990; 1991; 1994; and *Brush and Gillmor*, 1995]. I still have a ways to go.

In terms of community service to AGU, I was History Editor of *Eos* for three years and I founded and was Editor for the first five volumes of this series, *History of Geophysics.* I've enjoyed the help and support of historically oriented physicists such as David Stern, Martin Walt, George Siscoe and Ed Cliver. It has been a pleasure to work with John Spreiter and with AGU in producing the volume.

REFERENCES

Airy, G. B., *Autobiography of Sir George Biddell Airy*, ed. W. Airy, Cambridge, England, 1896.

Alfvén, H., A theory of magnetic storms and of the aurorae, *K. Sven. Vetenskapakad. Handl., Ser. 3, 18(3),* 1939.

Alfvén, H., A theory of magnetic storms and of the aurorae, II, The aurorae; III, The magnetic disturbances, *K. Sven. Vetenskapakad. Handl., Ser. 3, 18(9),* 1940.

Appleton, E. V., Letter to B. van der Pol, July 11, 1930. Appleton Room, Edinburgh University Library.

Appleton, E. V., Wireless studies of the ionosphere, *Proc., Inst. Elec. Engrs., 71,* 642-650, 1932.

Appleton, E. V., and M. A. F. Barnett, Local reflections of wireless waves from the upper atmosphere, *Nature, 115,* 333-334, 1925a.

Appleton, E. V. and M. A. F. Barnett, Wireless wave propagation. The Magneto-ionic theory-- The part played by the atmosphere-- The effect of diurnal variation, *Electrician, 94*, 398, 1925b.

Barkhausen, H., Zwei mit Hilfe der Neuen Verstärker entdeckte Erscheinungen, *Phys. Z., 20,* 401-403, 1919.

Barkhausen, H., Whistling tones from the Earth, *Proc. IRE, 18,* 1155-1159, 1930.

Birkeland, K., Expédition Norvegienne de 1899-1900 pour l'étude des aurores boréales, *Vidensk. Skrifter I. Mat. naturv. Kl. 1901, 1,* 1901.

Birkeland, K. , *The Norwegian Aurora Polaris Expedition, 1902-1903, vol. 1,* 1st and 2d sections, Aschehoug, Christiana, 1908 and 1913.

Booker, H. G., Fifty years of the ionosphere. The early years-- Electromagnetic theory, *J. Atmos. Terr. Phys., 36,* 2113-2136, 1974.

Breit, G., and M. A. Tuve, A radio method for estimating the height of the conducting layer, *Nature, 116,* 375, 1925.

Breit, G., and M. A. Tuve, A test of the existence of the conducting layer, *Phys. Rev., 28,* 554-573, 1926.

Brush, S. G., and C. S. Gillmor, Geophysics, pp. 1943-2016, in *Twentieth Century Physics*, 3 vols., American Institute of Physics (New York) and Institute of Physics (London), 1995.

Budden, K. G., *Radio Waves in the Ionosphere,* Cambridge University Press, 1961.

Burton, E. T., and E. M. Boardman, Audio-frequency atmospherics, *Proc. IRE, 21,* 1476-1494, 1933.

Carrington, R. C., Description of a singular appearance seen in the sun on September 1, 1859, *Mon. Not. R. astr. Soc., 20,* 13-15, 1860.

Cawood, J., The magnetic crusade: Science and politics in early Victorian Britain, *Isis, 70,* 493-519, 1979.

Chapman, S., The absorption and dissociative or ionizing effect of monochromatic radiation in an atmosphere on a rotating earth, *Proc. Phys. Soc. London, A, 43,* 26-45, 1931.

Chapman, S., Upper atmospheric nomenclature, *J. Atmos. Terr. Phys., 1,* 121-124, 1950.

Chapman, S., and J. Bartels, *Geomagnetism,* 2 vols., Oxford Press, Oxford and New York, 1940.

Chapman, S., and V. C. A. Ferraro, A new theory of magnetic storms, I., The initial phase, *J. Geophys. Res., 36,* 77-97, 171-186, 1931.

Chapman, S., and V. C. A. Ferraro, A new theory of magnetic storms, I, The initial phase (continued), *J. Geophys. Res., 37,* 147-156, 421-429, 1932.

Chapman, S., and V. C. A. Ferraro, A new theory of magnetic storms, II, The main phase, *J. Geophys. Res., 38,* 79-96, 1933.

Clark, R., *Sir Edward Appleton,* Pergamon, Oxford, 1971.

DeVorkin, D. H., *Race to the Stratosphere: Manned Scientific Ballooning in America,* Springer, New York, 1989.

Dessler, A., Swedish iconoclast recognized after many years of rejection and obscurity, *Science, 170,* 604-606, 1970.

Eccles, W. H., On the diurnal variations of the electric waves occurring in nature and on the propagation of electric waves round the bend of the earth, *Proc. R. Soc. London, A, 87,* 79-99, 1912.

Eckersley, T. L., An investigation of short waves, *J. Inst. Elec. Eng., 67,* 992-1032, 1929.

Eckersley, T. L., Musical atmospherics, *Nature 135,* 104-105, 1935.

Eriksen, K. W., C. S. Gillmor and J. K. Hargreaves, Some observations of short-duration cosmic noise absorption events in conjugate regions at high magnetic latitude, *J. Atmos. Terr. Phys., 26,* 77-90, 1964.

Fourier, Jean Baptiste, *Théorie analytique de la chaleur (1822),* in vol. 1 of *Oeuvres de Fourier*, ed. G. Darboux, 2 vols., Paris, 1888-1890.

Friedman, H., From ionosonde to rocket sonde, *J. Geophys. Res., 99,* 19143-19153, 1994.

Gilbert, William, *De Magnete* (London, 1600), trans. P. F. Mottelay, Dover, New York, 1958.

Gillmor, C. S., High-frequency extraterrestrial noise investigations, *Electrical Engineering, 81,* 22-26, 1962a.

Gillmor, C. S., Izmerania ionosfernovo pogloshenia metodom kosmicheskovo radioshuma, *Bulletin Sovetskoi Antarkticheskoi Expeditsii, No. 36,* 22-25, 1962b

Gillmor, C. S., Ionospheric Absorption at Mirnyy, Antarctica, U. S. National Bureau of Standards Report 7291, 33 pp., 1962c.

Gillmor, C. S., The day-to-night ratio of cosmic noise absorption during polar cap absorption events, *J. Atmos. Terr. Phys., 25,* 263-266, 1963a.

Gillmor, C. S., Cosmic-noise absorption at conjugate stations on the polar caps, *Eos, Trans. AGU, 44,* 10-15, 1963b.

Gillmor, C. S., *Coulomb and the Evolution of Physics and Engi-*

neering in Eighteenth-Century France, 328 pp., Princeton University Press, Princeton, New Jersey, 1971.

Gillmor, C. S., Aspects of the History of Ionospheric Physics in the Asian-Pacific Area, *Proc., Second General Assembly, International Assoc. of Geomagnetism and Aeronomy*, Kyoto, September, 1973. Abstract.

Gillmor, C. S., The Place of the Geophysical Sciences in Nineteenth Century Natural Philosophy, in *Eos, Trans. AGU, 56*, 4-7, 1975a.

Gillmor, C. S., Citation characteristics of the *JATP* literature, *J. Atmos. Terr. Phys., 37*, 1401-1404, 1975b.

Gillmor, C. S., The history of the term *Ionosphere*, *Nature, 262*, 347-348, 1976.

Gillmor, C. S., Early history of upper atmospheric physics research in Antarctica, pp. 236-262 in *Upper Atmosphere Research in Antarctica*, eds. L. J. Lanzerotti and C. G. Park, Antarctic Research Series, vol. 29, American Geophysical Union, Washington, D.C., 1978.

Gillmor, C. S., Threshold to space: Early studies of the ionosphere, pp. 101-114, in P. A. Hanle and V. del Chamberlain, eds., *Space Science Comes of Age*, Smithsonian Institution Press, Washington, D.C., 1981.

Gillmor, C. S., Wilhelm Altar, Edward Appleton, and the magneto-ionic theory, *Proc. Amer. Philos. Soc., 126, #5*, 395-440, 1982.

Gillmor, C. S., Aging of geophysicists, *Eos, Trans. AGU, 65*, 353-354, 1984.

Gillmor, C. S., L'Utilisation en géophysique du radiotélescope décimetrique de Nançay, *Le Journal des Astronomes Français, No. 25*, 16-17, 1985.

Gillmor, C. S., Federal funding and knowledge growth in ionospheric physics, 1945-1981, *Social Studies of Science, 16*, 105-133, 1986a.

Gillmor, C. S., S. K. Mitra's *The Upper Atmosphere*: The role of monograph and text literature in the evolution of ionospheric physics, *Indian Journal of radio and Space Physics, 15, 1*, 171-181, (Oct. and Dec. 1986b).

Gillmor, C. S., Aging of geophysicists reconsidered, *Eos, Trans. AGU, 68*, 802-803, 805-806, 1987.

Gillmor, C. S., Geospace and its uses: The restructuring of ionospheric physics following World War II, pp. 75-84 in *The Restructuring of Physical Sciences in Europe and the United States -- 1945-60*, eds. M. De Maria, M. Grili, and F. Sabastiani, World Scientific Publishers, Singapore, 1989.

Gillmor, C. S., Introduction, pp. ix-x, in *History of Geophysics, 4*, ed. C. S. Gillmor, American Geophysical Union, Washington, D.C. 1990.

Gillmor, C. S., Ionospheric and Radio Physics in Australian Science Since the Early Days, pp. 181-204 in *International Science and National Scientific Activity*, ed. R. W. Home and S. G. Kohlstedt, Kluwer, Netherlands, 1991.

Gillmor, C. S., The big story: Gregory Breit, Merle Tuve and ionospheric physics at the Carnegie Institute of Washington, pp. 133-141 in *History of Geophysics, 5*, ed. G. Good, American Geophysical Union, Washington, D.C., 1994.

Gillmor, C. S. and Gran, D. Research in ionospheric physics, *Actes of the XIIIth International Congress of the History of Science, Moscow, 6*, 160-164, 1974.

Gillmor, C. S., and J. K. Hargreaves, The occurrence of short-duration cosmic noise absorption events inside the southern auroral zone, *J. Atmos. Terr. Phys., 25*, 311-317, 1963.

Gillmor, C. S. and C. J. Terman, Communication modes of geophysics: The case of ionospheric physics, *Eos, Trans. AGU, 54*, 900-908, 1973.

Helliwell, R. A., *Whistlers and Related Ionospheric Phenomena*, Stanford University Press, 1965.

Helliwell, R, A., Whistlers, (This volume, below), 1997.

IGY. *International Geophysical Year General Report, No. 21* (Washington, D.C.: National Academy of Sciences) 1965.

Lagrange, Joseph-Louis de, *Oeuvres de Lagrange*, ed. J. A. Serret, 14 vols., Paris, 1867-1892.

Larmor, J., Why wireless electric rays can bend round the earth, *Phil. Mag., 48*, 1025-1036, 1924.

Lorentz, H. A., *The theory of Electrons*, B. G. Teubner, Leipzig, 1909, 2d ed. 1916.

Nicolet, M., Historical Aspects of the IGY, *Eos, Trans. AGU, 64*, 369-370, 1983.

Parker, E. N., Adventures with the Geomagnetic Field, (This volume), 1997.

Pedersen, P. O., *The Propagation of Radio Waves Along the Surface of the Earth and in the Atmosphere* Danmarks Naturvidenskabelige Samfund, A. Nr. 15a and 15b, Copenhagen, 1927.

Pomerantz, M., The ancestry of solar-terrestrial research, *Eos, Trans. AGU, 55*, 955-7, 1974.

Ratcliffe, J. A., *The Magneto-Ionic Theory and Its Applications to the Ionosphere*, Cambridge University Press, 1959.

Royal Society, Report of the president and council of the Royal Society on the instructions to be prepared for the scientific expedition to the antarctic regions, London, 1840.

Schuster, A., On electric currents induced by rotating magnets, and their application to some phenomena of terrestrial magnetism, *J. Geophys. Res., 1*, 1-17, 1896.

Schuster, A., The origin of magnetic storms, *Proc. Roy. Soc. London, A, 85*, 44-50, 1911.

Stern, D. P., A brief history of magnetospheric physics before the spaceflight era, *Rev. Geophys., 27*, 103-114, 1989.

Stewart, B. Terrestrial Magnetism, in *Encyclopaedia Britannica*, 9th ed., (U.S. ed., 1883, *16*, 159-184), 1882.

Storey, L. R. O., An investigation of whistling atmospherics, *Phil. Trans. R. Soc. London, A, 246*, 113-141, 1953.

Størmer, C., Sur des trajectoires des corpuscles electrisés dans l'espace sous l'action du magnetisme terrestre, *Arch. Sci. Phys. Nat., Ser. 4, 24*, 317-64, 1907.

Størmer, C., Periodische Elektronenbahnen im Felde eines Elementarmagneten und irhe Anwendung auf Brüches Modellversuche und auf Eschenhagens Elementarwellen des Erdmagnetismus, *Z. Astrophys., 1*, 237-274, 1930a.

Størmer, C., Twenty-five years' work on the polar aurora, *J. Geophys. Res., 35*, 193-208, 1930b.

Van Allen, J. A., *Origins of Magnetospheric Physics*, 144 pp., Smithsonian, Washington, D.C., 1983.

Van Allen, J. A., Genesis of the International Geophysical Year, *Eos, Trans. AGU, 64*, 977, 1983.

Villard, O. G., Jr., The ionospheric sounder and its place in the history of radio science, *Radio Science, 11*, 847-860, 1976.

Wilson, E., *Diary of the 'Terra Nova' Expedition to the Antarctic, 1910-1912*, Blanford Press, London, 1972.

C. Stewart Gillmor, Professor of History and Science, Wesleyan University, 238 Church St., Middletown, CT 06459. e-mail: sgillmor@wesleyan.edu

Aurora Research During the Early Space age: Personal Account

S.-I. Akasofu

Geophysical Institute, University of Alaska Fairbanks

The progress of a ground-based study of auroral morphology and of geomagnetic disturbances in the 1960s is described as a personal account. It is emphasized that although many of the conclusions are stated simply in a few sentences in modern monographs, there was a decade of history involved in reaching those conclusions against pre-existing dogmas. Some of those conclusions are results of a naive question by a graduate student, which opened a new discipline. This may be a lesson for the field of space physics which is now becoming more mature and increasingly quantitative. It does not necessarily mean that the field is coming to an end, however; often, when the ability of researchers is limited by pre-existing dogma, a field may only appear to become mature. Thus, it is important to encourage young researchers to develop the field in the ways they believe in, rather than only making future research more quantitative. It is also emphasized that as a result of several recent advances in ground-based research, a new advance can be made by integrating ground-based and satellite-based observations, together with theoretical/modeling research.

1. INTRODUCTION

The development of auroral science during the 1960s may provide a hint about how to advance magnetospheric physics during the first few decades of the 21st century. There are always periods when many scientists in a particular field tend to feel that there is nothing left to be done. Such is the time when a particular dogma or paradigm matures and prevails, although there might exist many fundamental unsolved problems which are tacitly believed to be understood or solved. A new era in such a field could be opened as a result of a naive question or challenge to those problems by a newly arriving graduate student. This article begins with a personal account of the progress of auroral morphology in the 1960s. Although satellite-based space research has been more emphasized in magnetospheric physics for many years, ground-based space physics has recently reached a very interesting stage in contributing to space physics on an equal basis with satellite-based space physics.

Discovery of the Magnetosphere
History of Geophysics Volume 7

2. AURORAL ZONE TO AURORAL OVAL

It was *Loomis [1860]* who assembled the first extensive collection of auroral appearance and found that the aurora tends to appear as a fairly narrow belt centered around a point at the northwestern tip of Greenland, not at the geographic pole. *Fritz [1873]*, using much more data covering the period from 503 B.C. to A.D. 1872, confirmed Loomis' findings and

night hours and deviates greatly from this latitude at the other local times. However, I could not determine the distribution on the dayside because of the lack of data at that time.

It was *Feldstein [1963]* who determined completely the distribution of the aurora at all local times, using the films from Heiss Island and others, which can observe mid-day auroras. His distribution showed that the belt of the auroral zone is located at about 78° during mid-day hours, instead of 67° (Figure 1). Further, the center of the belt is shifted by about 3° from the geomagnetic pole toward the midnight sector. This belt is called the *auroral oval.* Since the results obtained by Feldstein were basically the same as mine in the dark hours, I supported his results immediately.

It was a sort of golden age of auroral spectroscopy at that time. All-sky cameras were not considered to be a scientific instrument, compared with the then sophisticated spectroscopic instruments. In fact, some of my senior colleagues advised me by saying that the aurora should be the same in Alaska, Siberia, Canada, Norway, that physics of the aurora should be the same everywhere, that the distribution of the aurora is thus not a major issue and thus that it is wasting time to work on the distribution of the aurora. I recall I objected to such an argument by saying that the fact that auroral arcs appear in a very specific belt called the auroral oval, and not all over the polar region, tells us something about their origin, so that it is important to accurately determine the actual distribution.

In such atmosphere, Feldstein's results got little attention from the scientific community. Worse, since the auroral zone had been believed to be the belt of auroral arcs for more than 100 years, it was difficult for us to convince our colleagues of the validity and significance of the auroral oval.

In order to convince the scientific community of the validity of the concept of the auroral oval, I planned several projects. The first one was to establish the Alaska meridian chain of all-sky cameras. Taking advantage of the earth's rotation, a meridian chain of all-sky cameras can scan the entire polar sky (like an azimuth scanning radar at an airport) once a day and delineates the auroral oval which is fixed with respect to the sun (Figure 2). I believe that this is the 'largest scanning device on earth'.

Figure 3 shows an example of the results from this observation. If auroral arcs were distributed along the auroral zone, they should appear in a horizontal belt along the latitude of Fort Yukon. Instead, auroral arcs appear at about gm 76-77° at O GMT (14 MLT, magnetic local time) and shift toward the latitude of College (Fairbanks). The line-dot curve shows Feldstein's oval for the magnetic index Q = 3. Therefore, the meridian chain of all-sky cameras could delineate the auroral oval. The width of the oval changes intermittently. This phenomenon will be discussed later.

IMS Alaska Meridian Chain at Different UT

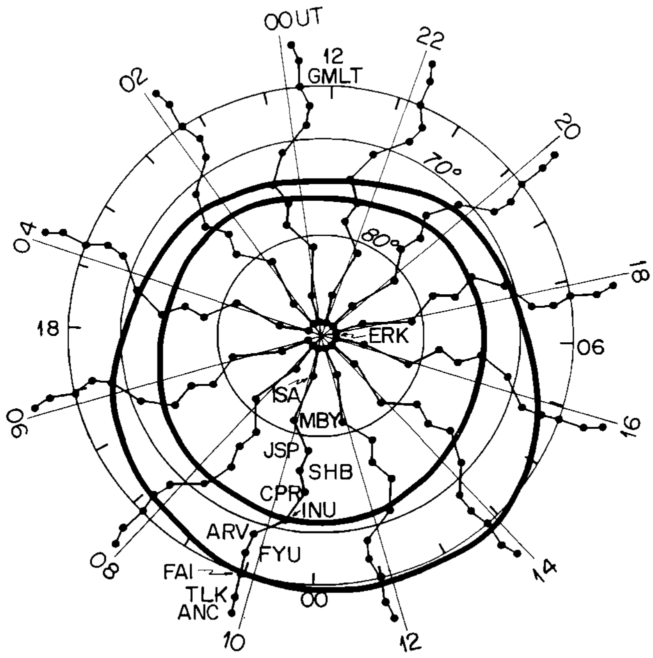


Figure 2. The Alaska meridian chain of the magnetometers and all-sky cameras, scanning the auroral oval once a day as the earth rotates.

The second project was to fly along auroral arcs, since the flight path should be able to delineate the auroral oval. Both a US Air Force jet from Hanscom Air Force Base and a NASA jet from Ames Research Center participated in the operation. The flight paths delineated clearly the auroral oval. George Gasmann, Jurgen Buchau and their colleagues of the Phillips Laboratory were instrumental in accomplishing this task. However, I felt that the scientific community in general was not much interested in such observational results at that time.

The third attempt was to find other geophysical phenomena which have a similar distribution with the auroral oval. Fortunately, I had an opportunity to work with the space physics group of the University of Iowa. I found one day that Lou Frank, James Van Allen and John Craven were plotting the outer boundary of the outer radiation belt onto the earth's surface. It was my great surprise that the boundary thus delineated coincides fairly well with the auroral oval (Figure 1). I remember that I reported the results to Van Allen. It was the time when the initial hope of associating auroral phenomena with the radiation belts had faded, so that it was difficult initially to convince my colleagues of the significance of this finding. Later, *Zmuda et al [1966]* found that field-aligned currents flow in or out from the auroral oval. I recall that Al

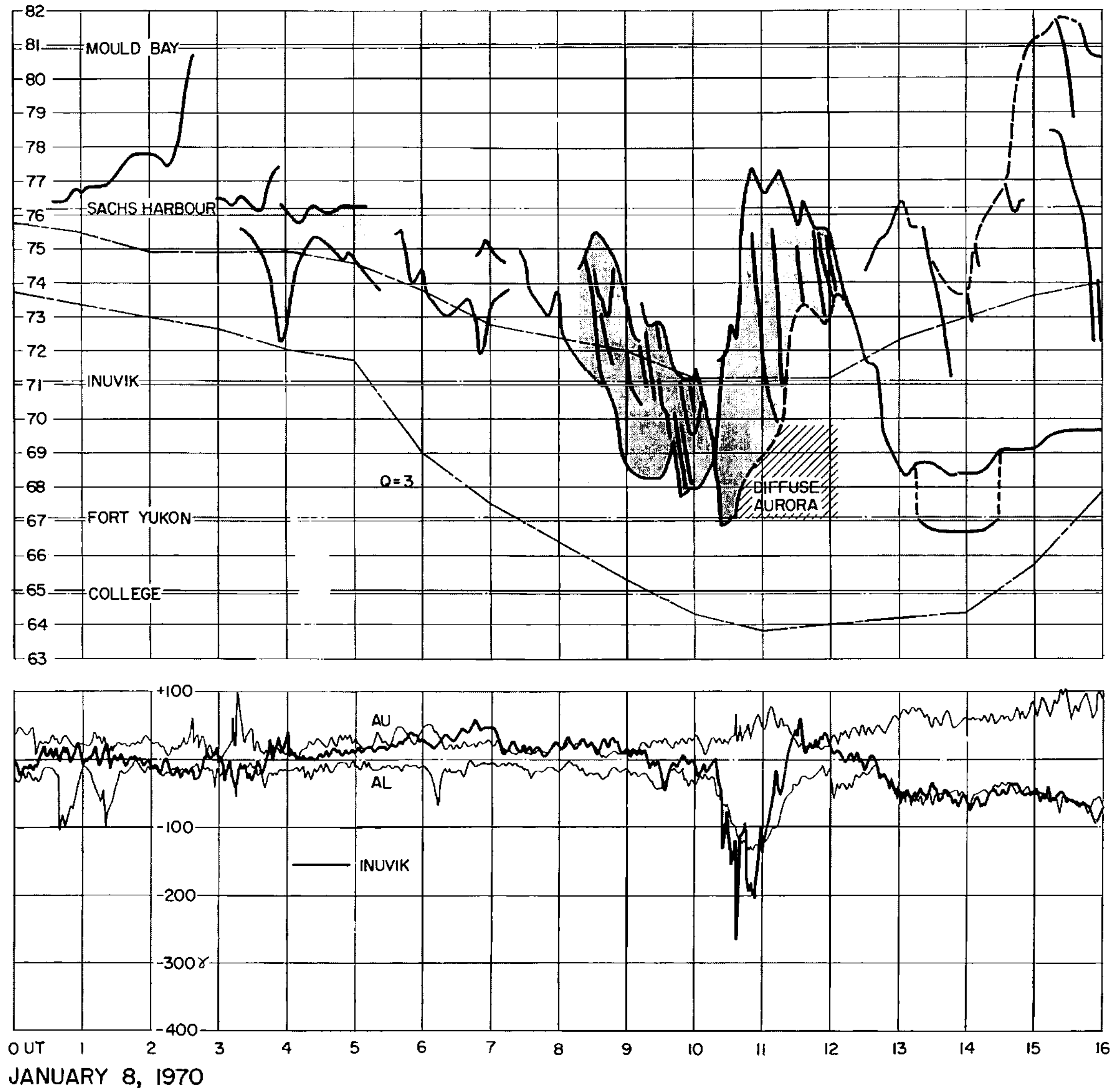


Figure 3. Upper: The auroral oval scanned by the Alaska meridian chain of all-sky cameras on January 8, 1970. Lower: The AE index.

Zmuda called me, saying that the tape recorder of his satellite TRIAD failed and asking me to install his satellite receiving station at the top of the Geophysical Institute building; it was installed at 50 below zero that winter.

We found that the location of the field-aligned currents agreed well with the oval. In particular, we found that regions of the upward field-aligned currents coincide with those of auroral arcs. *Iijima and Potemra [1976]* completed Zmuda's work by showing the distribution of field-aligned currents at the ionospheric level. Further, energetic solar electrons were found in the area bounded by the auroral oval. Therefore, the significance of the auroral oval began to be recognized toward the end of the 1960's. Actually, it was in the later part of the 1970s when the 'open' magnetosphere was fully recognized and when the auroral oval was found to delineate the boundary of the open region in the polar cap.

However, we had to wait for the full recognition of the auroral oval until 1971 when a scanning instrument, devised by Cliff Anger, aboard the ISIS-2 satellite, imaged the entire oval *[Lui et al., 1975]*. After this, even the concept of the auroral oval was accepted as if nothing had happened earlier.

In any modern monograph on the aurora, one can find a simple statement that auroral arcs lie along the auroral oval. It is interesting to recognize that such a simple fact had a long history; it was about a decade of struggle for acceptance by the scientific community. Personally, it started out with a naive question about the well known daily auroral behavior at that time. Recalling those days, I appreciate the foresight and courage of both Chapman and Elvey for taking the leadership of the all-sky camera project, in spite of the fact that auroral spectroscopists and auroral physicists in general paid little attention to it. Also, as far as I know, an all-sky camera was used for the aurora by Carl Gartlein and Stoffregen for the first time.

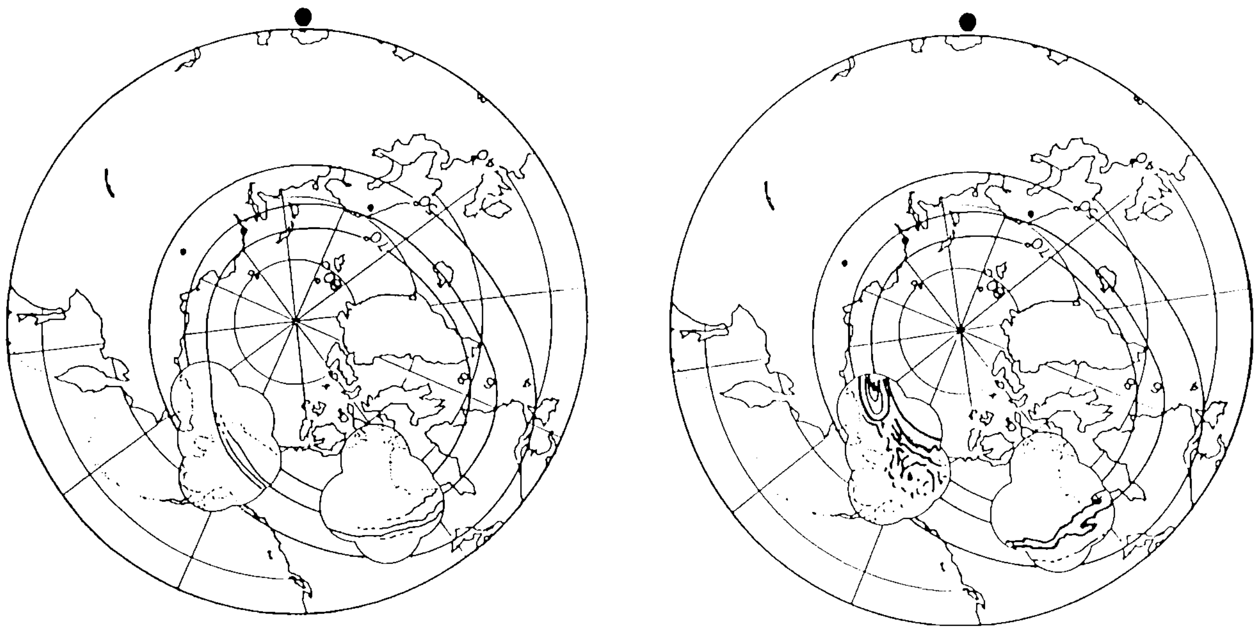
3.AURORAL SUBSTORMS: FIXED PATTERN TO SUBSTORM PATTERN

It had long been believed on the basis of study of the aurora by *Fuller [1935]* and *Heppner [1954]* that in the evening sky, auroral arcs always had a quiet and homogeneous form, that auroral arcs were always very active in midnight hours and became patchy in the morning sky. In this view, auroral activities are fixed with respect to the sun (and thus to MLT), and the earth and a single observer at a point on it rotates under such a pattern of activity once a day. That is, quiet forms, active forms and patchy forms are familiar features in the evening, midnight and morning skies, respectively.

At the beginning of the IGY, little was known how auroras behave *simultaneously* in Siberia (in evening hours) and Canada (in morning hours) when auroras became suddenly active over the Alaskan sky (in midnight hours). There had not been any simultaneous observations of auroras over a long local time span until that time. An early analysis of the IGY all-sky films also supported the fixed pattern concept (Davis, 1962).

However, all-sky films from even a single station gave me quite a different picture from the statistical concept mentioned in the above, since auroral arcs can be quiet during midnight hours. Further, all-sky films during a single night showed that auroral arcs can transform themselves from a quiet to active and back to a quiet form twice or three times during a single night. This fact suggested to me either that the fixed pattern concept was not correct or that the earth rotated two or three times in a single night!

The graduate student was obviously puzzled, but overwhelmed by the firm believers of the fixed pattern. I decided to examine *simultaneous* all-sky photographs from Siberia, Alaska and Canada, when Alaska was in the midnight sector. It was my finding that when an auroral arc is quiet in the Alaskan sky, it is also quiet over Siberia and Canada. When an auroral arc suddenly brightens and moves rapidly poleward over the Alaskan sky, this activity generates a large wavy or folding structure (the westward traveling surge) which propagates along the arc toward Siberia (toward the evening sky); Figure 4. This surge-like activity was recorded first at the Siberian station closest to Alaska several minutes after its



08 30 08 42

Figure 4. An auroral substorm recorded by Siberian, Alaskan and Canadian all-sky cameras, February 13, 1958.

formation over Alaska and subsequently at other earlier evening stations. This activity could propagate all the way to the dayside of the oval with a speed of a few kilometers per second. At the same time, auroras over Canada became active, often forming an inverted Ω-shaped form (called the omega band). To the south of the omega band, auroral arcs became folded in a very complicated way. Folded portions appear as shafts of lights or patchy forms, scattering all over the sky.

More importantly, when auroras over Alaska become quiet again, in about two to three hours, auroras over Siberia and Canada also become quiet. Further, such activity often repeated two or three times a night. Chapman coined the term 'auroral substorm' for this transient phenomenon *[see Akasofu, 1964]*. There was little mention of such auroral features in the then newest and most authoritative book published by *Chamberlain [1961]*. Therefore, I sent a paper to the *Journal of Geophysical Research*, reporting on the above findings. The paper was rejected, on the basis that there was nothing worth reporting. So I decided to analyze simultaneous all-sky films from a large number of stations and became more convinced of the validity of my findings. A new paper was then sent to the late Sir David Bates, the editor of Planetary and Space Science, who accepted it without review. I could assume this because I received his acceptance letter only about 10 days after sending the paper to him.

However, I found it very difficult to convince my colleagues of my findings. This was particularly the case for those who were experienced in observing the aurora. This was because a single observer standing at a point on the earth is carried by the earth's rotation with a speed of 15° (in longitude) per hour, so that he gets an impression *statistically* that the fixed pattern was correct. Elvey was a firm believer of the fixed pattern concept. Many auroral scientists who have little experience in observing the aurora simply followed the experienced ones. Thus, it was hard to convince any one about the validity of the concept of the auroral substorm. The only exception at that time was Feldstein who strongly supported my findings.

Therefore, I had to devise a scheme to prove the validity of the concept of the auroral substorm. The best way would have been to observe the aurora from a fixed point (with respect to the sun) well above the north polar region for many hours, as the Dynamic Explorer (DE) satellite did in the 1980s. In the middle of the 1960's, this was nothing but a dream. One method I conceived was to fly westward under the aurora in a jet plane. Along the latitude circle of 65° or so, the speed of a jet plane is approximately the same as that of the earth's rotation. Thus, a jet plane can stay in the midnight sector for about six hours by flying from the East coast at midnight to Alaska. Both NASA and Air Force jet planes contributed to the so-called 'constant local time (midnight)' flights.

On my way back from one such trip to Hanscom Air Force Base, I learned that Elvey, who had retired by then in Tucson, Arizona, was critically ill, and thus decided to visit him. Elvey was waiting for my results. We sat together at his bedside to scan the all-sky film obtained by one of the constant local time (midnight) flights, which clearly registered intermittent auroral activities in the midnight sector. We firmly shook hands. He said, "Syun, you did a good job". I believe that I had finally convinced him of the validity of the concept of the auroral substorm. I noticed that his arms were just skin and bones. I was told of his death about ten days later.

During the next decade, it was fortunate that many people realized that they can understand and interpret their observational results better in terms of the concept of the auroral substorm, rather than of the fixed pattern. However, I had to wait for the result of the Dynamic Explorer satellite. Indeed, it was the final test of the concept of the auroral substorm *(Frank and Craven [1988])*. After all, the auroral substorm seen from below and above must be the same. Nevertheless, It is important to learn that it takes a much longer time than one thinks to convince one's colleagues if one's finding is radically different from what has been believed for years.

4. NIKOLSKY'S SPIRALS

The concept of the auroral zone had greatly influenced the study of geomagnetic disturbances. The SD current system, obtained by *Chapman [1935]*, was an example of this influence. He suggested that the auroral currents consist of a pair of concentrated currents; the westward auroral electrojet in the morning sector, and an eastward electrojet in the afternoon sector, and their return currents in the polar cap and in lower latitudes (Figure 5). It was considered that a magnetic

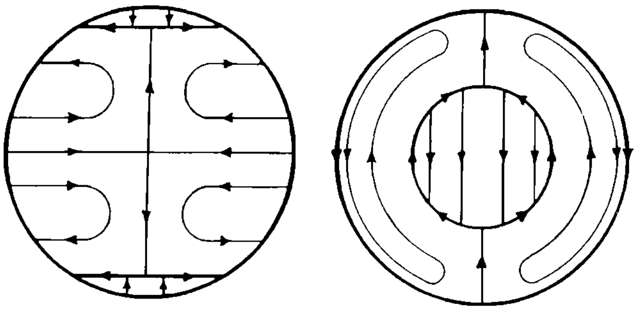


Figure 5. The SD current system (Chapman, 1935).

observatory rotates under such a fixed current system, registering the daily magnetic variations. Under the eastward electrojet (in the morning sector), there occurs positive (poleward) magnetic disturbances in the horizontal (H) component, while the westward electrojet produces negative (equatorward) magnetic disturbances.

The SD current system had become the standard model and a major paradigm for a few dacades. However, *Nikolsky [1947]* found an interesting phenomenon. He found that geomagnetic disturbances in high latitude stations have three peaks in a day. He denoted those peaks A (afternoon), N (night) and M (morning), respectively (Figure 6). Further, he found that the time of the peak tends to occur earlier at higher latitudes for the A and N peaks, later for the M peaks. Thus, in a polar plot, the peak occurrence time for A, N and M delineates three spiral curves, respectively, I recall I was fascinated by Nikolsky's results, but had no idea as to how to interpret them. However, one day, I recognized that the combination of the N and M spirals delineates the auroral oval; the A peak spiral indicates the eastward electrojet. The results suggested to me that the westward electrojet does not stop in the midnight sector and continues to flow westward with the westward traveling surge along the auroral oval in the evening sector. Thus,

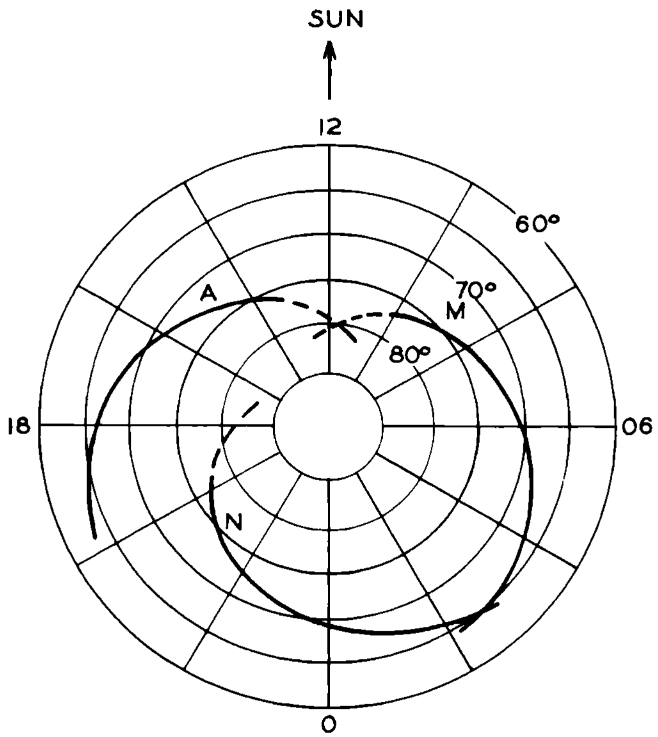


Figure 6. The gm. latitude-MLT (magnetic local time) relationship for the afternoon peak (A), the night time peak (N) and the morning peak (M), Nikolsky (1947).

the westward electrojet is located at latitudes higher than 65-70°, namely not along the auroral zone, in the evening sector. When I reported this result in Moscow in the 1960s, Nikolsky was very happy and gave me a typical Russian bear hug, saying that I was his son's age.

I thought that this was reasonable, because the auroral ionization takes place along the auroral oval, not along the auroral zone. Therefore, there is no reason that the westward electrojet stops in the midnight sector. The Alaska meridian chain of all-sky camera stations was also equipped with magnetometers. An examination of both all-sky photographs and magnetic records indicated clearly that the westward electrojet extends into the evening sector with the westward traveling surge. During the passage of westward traveling surges to the north, auroral zone stations register positive changes in the H component, while at stations of gm. lat. 70-75°, negative changes with greater magnitudes are observed. Chapman accepted my revision after examining those records together. Although these results were presented in several papers at that time, it was difficult to convince many of my colleagues of the results. I recall that there were even emotional objections to such a significant change of the configuration of the electrojets. Again, Feldstein was one of the strong supporters of my results.

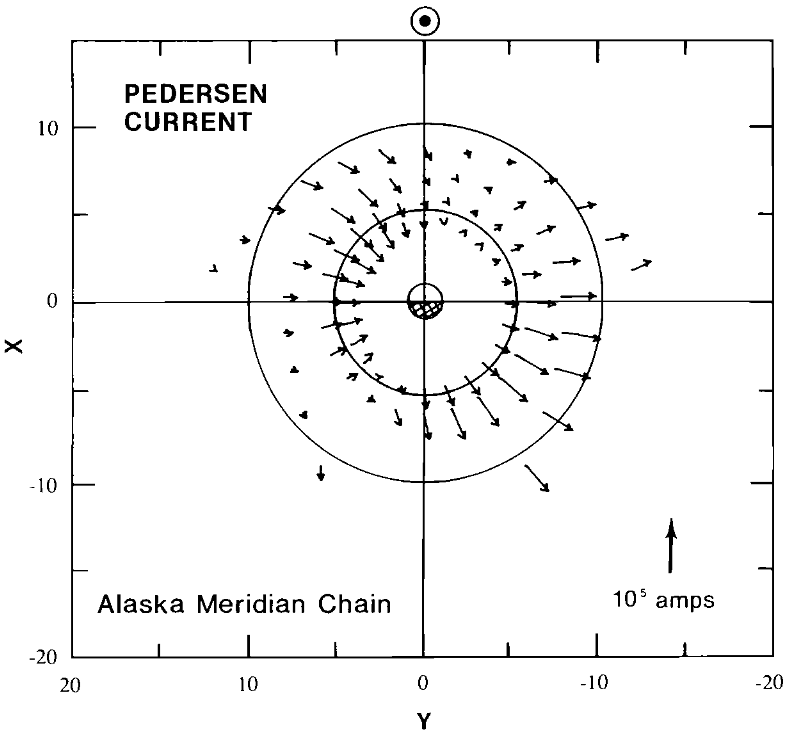
Chapman's SD current system is, strictly speaking, an equivalent current. That is, Chapman assumed that all the currents causing magnetic disturbances on the ground were flowing in a conductive shell which is concentric to the earth. Chapman told me that he thought that since there are an infinite number of possible current systems for a given distribution of magnetic disturbance fields on the ground, it does not make sense to choose just one arbitrarily. Instead, he thought that he could remain accurate, so long as he dealt with the equivalent (two-dimensional) current system. Although Chapman had many deep insights into physical processes, he tended to become an applied mathematician when he encountered mathematical uniqueness issues. Mathematical rigor was his life and it was part of the reason which caused some friction with Hannes Alfvén, who tended to be intuitive in interpreting physical phenomena.

Unfortunately, most researchers took Chapman's equivalent current system as the true current system for many decades. Ching Meng and I decided to examine whether or not the observed distribution of magnetic disturbance vectors can be reproduced by a model three-dimensional current system which was developed by Alfvén and modeled by *Kirkpatrick [1952]*. Kirkpatrick's three-dimensional model is not too much different from what is a currently accepted one. It was a great surprise that Kirkpatrick's model reproduced the observations very well. When I showed my results to Alfvén, I recall Alfvén was almost irritated. He said that I was too slow to recognize the validity of a three-dimensional current system. I could well understand his impatience.

Incidentally, I object to the use of the term 'Birkeland current' for the field-aligned currents in magnetospheric physics, because Birkeland's currents are far from what we define as the field-aligned currents today, which flow between the magnetosphere and the ionosphere. I also object to statements which imply that Chapman was wrong in rejecting Alfvén's paper on magnetic storms. Note that neither Birkeland nor Alfvén could envisage the magnetosphere. In their first paper on the formation of what we now call the magnetosphere, Chapman and Ferraro (1931) obtained an equation similar to the Debye length and concluded that the solar gas flow must be treated as what we now call 'plasma'. Note that Birkeland, Stormer and Alfvén treated the solar gas to be composed of single, independent particles. Alfvén's magnetosphere, if any, is quite different from what we know today, although his paper appeared later than the Chapman-Ferraro paper. I am not trying to criticize the monumental work by Birkeland and Alfvén. What I emphasize here is that we must be cautious in carelessly commenting on the works by our great pioneers. We must give credits accurately where they belong.

When the Alaska meridian chain of magnetometers became operational, I was very surprised that we could obtain a fairly systematic magnetic vector distribution over the entire polar region by averaging the data for only one month. Yosuke Kamide joined my group as a post doc from the University of Tokyo. I suggested to him that there may be a way to obtain the true current system, not the equivalent current system, using such a systematic data set. After moving to the National Center for Atmospheric Research, Boulder, Colorado, he completed a computer algorithm with Art Richmond and Sadami Matsushita for this purpose, which is now called the KRM.

Gordon Rostoker was also establishing a meridian chain of magnetometers in Canada. Thus, we agreed to operate Inuvik Station as a joint station. Our operation stimulated our colleagues to establish four other chains during the International Magnetosphere Study (IMS). Thus, six meridian chains of magnetometers, consisting of 71 magnetometer stations, became operational during the IMS. Applying the KRM method to the data thus obtained, a great wealth of knowledge on the ionospheric currents, electric fields, potential field-aligned currents, and the Joule heating rate was obtained. Thus, finally the dream of Birkeland and Chapman to obtain the



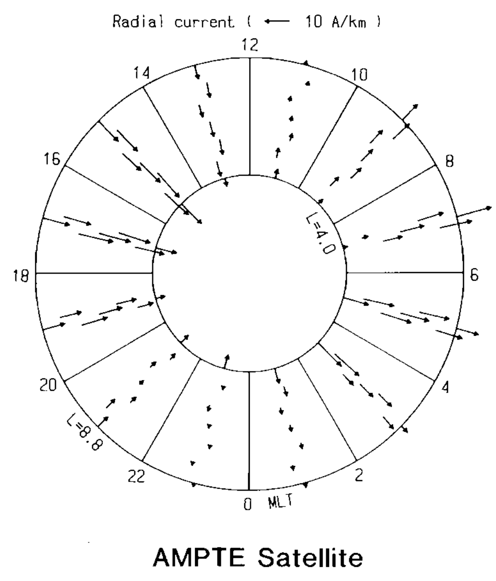


Figure 7. Left: The equatorial current associated with the Pedersen current. Right: The radial component of the equatorial current obtained by the AMPTE satellite (Iijima et al., 1990).

important to encourage creative young researchers to develop the field in the ways they believe in, rather than directing themselves towards more quantitative measurements of old parameters. It is also important to recognize that a new development is qualitative by definition. On the other hand, young researchers should be reminded that it is an extremely difficult task in many ways to nurture a new idea to the point that it establishes itself as an independent discipline against an established field.

Acknowledgments: I would like to thank my colleagues who have participated in auroral research for many years. In particular, I would like to thank Dr. Charles Deehr and Dr. Eugene Wescott for reading the manuscript and providing their comments. The paper presented here was supported in part by a grant from the National Science Foundation ATM 93-11747.

REFERENCES

Akasofu, S.-I., The development of the auroral substorm, *Planet. Space Sci., 12*, 273, 1964.

Boström, R.A., A model of the auroral electrojets, *J. Geophys. Res., 69*, 4983, 1964.

Chamberlain, J.W., *Physics of the Aurora and Airglow*, Academic Press, New York, 1961.

Chapman, S. and V.C.A. Ferraro, A new theory of magnetic storms, Part I. The initial phase, *Terr. Magn. Atmos. Elect., 36,* 77, 1931.

Chapman, S., The electric current-systems of magnetic storms, *Terr. Magn. Atmos. Elect., 40,* 349, 1935.

Davis, T.N., The morphology of the auroral displays of 1957-1958, 1. Statisical analysis of Alaska data, *J. Geophys, Res., 67*, 59, 1962.

Feldstein, Y.I. and E.K. Salomatina, Some problems of the geographic distribution of aurorae in the northern hemisphere, Results of IGY Researchers. *Aurora and Airglow No. 7, Sec. IV, IGY Program, 51,* Publishing House Acad. Sci. USSR, Moscow, 1961.

Feldstein, Y.I., Some problems concerning the morphology of auroras and magnetic disturbances at high latitudes, *Geomagnetizm i Aeronomiya, 3*, 183, 1963.

Frank, L.A. and J.D. Craven, Imaging results from Dynamic Explorer 1, *Rev. Geophys., 26*, 249, 1988.

Fritz, H., Das Polarlicht, 348 pp., Leipzig, 1881.

Fuller, V.R., A report of work on the aurora borealis for the years 1932, *Terr. Magn. Atmos. Elect., 40,* 269, 1935.

Heppner, J.P., Time sequences and spatial relations in auroral activity during magnetic bays at College, Alaska, *J. Geophys. Res., 59*, 329, 1954.

Hultzvist, B., Circular symmetry in the geomagnetic plane for auroral phenomena, *Planet. Space Sci., 8*, 142, 1961.

Iijima, T. and T.A. Potemra, The amplitude distribution of field-aligned currents at northern high latitudes deserved by TRIAD, *J. Geophys. Res., 81*, 1976.

Iijima, T., T.A. Potemra and L.J. Zanetti, Large-scale characteristics of magnetopheric equatorial currents, *J. Geophys. Res., 95*, 991, 1990.

Kirkpatrick, C.B., On current systems proposed for SD in the theory of magnetic storms, *J. Geophys. Res., 57*, 511, 1952.

Loomis, E., On the geographical distribution of auroras in the northern hemisphere, *Amer. J. Sci. Arts, 30*, 89, 1860.

Lui, A.T.Y., C.D. Anger, D. Venkatesan, W. Sawchuk and S.-I. Akasofu, A uniform belt of diffuse auroral emission seen by the ISIS-2 scanning auroral photometer, *J. Geophys. Res., 80*, 1795, 1975.

Nikolsky, A.P., Dual laws of the course of magnetic distrubances and the nature of magnetic disturbances and the nature of mean regular variations, *Terr. Magn. Atmos. Elect., 52*, 147, 1947.

Vestine, E.H., The geographic incidence of aurora and magnetic disturbance, northern hemisphere, *Terr. Magn. Atmos. Elect., 49*, 77, 1944.

Zmuda, A.J., J.H. Martin and F.T. Heuring, Transverse magnetic disturbances at 1100 km in the auroral region, *J. Geophys. Res., 71*, 5033, 1966.

S.-I. Akasofu, Geophysical Institute, University of Alaska Fairbanks, P.O. Box 757320, 903 Koyukuk Drive, Fairbanks, AK 99775-7320.

The Earth's Magnetosphere: Glimpses and Revelations

Kinsey A. Anderson

Space Sciences Laboratory and Department of Physics,
University of California, Berkeley, California

This article is a rather individualized account of research activities over a 20 year interval beginning in 1955. Balloon-borne detectors designed to measure galactic cosmic rays encountered several phenomena which later were seen to be workings of the magnetosphere. After 1958 solar energetic particles were employed to test the Dungey hypothesis. Initially solar ions of energy ~10 to 100 MeV were used and in 1966 fast solar flare electrons provided clear experimental support for the connection of the interplanetary magnetic field and the geomagnetic field. Data from small, low-altitude lunar orbiting spacecraft released from the service modules of Apollo 15 and 16 led to (1) direct determination of the cross tail electric field and (2) a new method to remotely measure small scale magnetic fields near the surface of planetary bodies was devised and applied to lunar magcons.

1. INTRODUCTION

In celebration of its 75th Anniversary, the American Geophysical Union published a collection of retrospective papers written by authors the Editor of the *Journal of Geophysical Research* called the "Pioneers of Geophysics." The several authors used a variety of approaches ranging from somewhat personalized accounts to abundantly referenced reviews. The revelation of the Earth's magnetosphere is also a tale that can be told in many ways depending on how individual scientists found their way to this rich and fundamentally significant scientific endeavor and which of the manifold aspects of the magnetosphere attracted them. My account will be rather personalized and will begin with the pre-discovery phase in order to indicate what backgrounds at least one of the participants in magnetospheric research came from, and to recount some of the early glimpses that scientists were getting of phenomena that later turned out to be appreciated as workings of the Earth's magnetosphere. My telling of the tale begins in 1955, four years before the term "magnetosphere" was applied by Tom Gold, and continues into the 1970s.

From the late 1940s in the Midwestern United States, a number of mostly state-supported universities developed research groups focussed on study of galactic cosmic radiation. A great impetus was given to this research by the discovery that atomic nuclei with $z \geq 2$ were arriving at Earth with great energies. Groups in Chicago, Iowa, and Minnesota developed strong experimental programs using balloons and rockets to carry instruments to high altitudes over a wide range of latitudes. Following the discoveries of "soft radiation" by the Iowa group [1955] and X-rays beneath visible aurorae by Winckler in Minnesota [1957] much of this expertise was transferred to magnetospheric studies. Following James Van Allen's demonstration of the geomagnetically trapped radiation in 1958 this process accelerated.

2. BENEATH NORTHERN SKIES, EARLY GLIMPSES OF THE MAGNETOPSHERE

In 1955, after receiving a Ph.D. degree from the University of Minnesota, I was asked by John Winckler to partici-

Discovery of the Magnetosphere
History of Geophysics Volume 7

pate in a survey of cosmic ray intensity over the north geomagnetic latitude range 51 to 65°. The objective of the expedition was to determine the latitude of the "knee" and compare its location to earlier measurements, especially those of Martin Pomerantz and G. W. McClure [1952]. Our payloads were designed to have very small mass so that a single latex balloon could carry each one to altitudes above 30 km. The highest geomagnetic latitude of 65° was reached at Flin Flon, Manitoba. We were now very close to the auroral oval although at that moment thoughts of auroral phenomena were not clearly in our minds. Near a lake we found an open grassy area suitable for balloon launches. As the balloon launched there on 26 August 1955 approached its highest altitude, the counting rate of the triple coincidence geiger tube telescope increased dramatically, quite unlike anything encountered on the earlier, more southerly flights. We knew that "soft radiation" had been detected at this magnetic latitude by the Iowa Rockoon group in the Northern summer of 1953, and that the Iowa scientists had established that the effects in their detectors were produced by X-rays. However, we did not understand how X-rays could produce such a high rate of triple coincidences in the counter telescope. This striking result gave both John Winckler and me a lingering feeling that wonderful mysteries were waiting to be revealed beneath northern skies.

Later in September 1955, I arrived in Iowa City having accepted an offer from James Van Allen to work as a Research Associate in the Physics Department at the State University of Iowa. During my first year there I followed up on my thesis work, but soon realized that cosmic ray research was changing rapidly and my research must move in a new direction. In late September, the SUI Rockoon group returned from their 1955 expedition bringing more results on the "soft radiation." When shown their data, I noticed that at times the soft radiation penetrated to altitudes sufficiently low to be detected by instruments carried on plastic film Skyhook balloons available at that time. These low cost "space vehicles" had been increasingly used for cosmic ray research and other applications since the late 1940s. They had proven to be capable of remaining at high altitude for 10 to 30 hours. Flown in the auroral zone, I thought they might provide synoptic and time variation studies of "soft radiation" and low energy cosmic rays not attainable from the brief flight times of sounding rockets. In thinking about how to go about doing such flights, I recalled a seminar in the Physics Department at the University of Minnesota late in 1954, in which Phyllis Freier and others had discussed what the cosmic ray group there would do during the International Geophysical Year (IGY), an 18-month interval beginning 1 July 1957. She described a program of regular balloon launches throughout this period. Each balloon would carry a standard set of cosmic ray particle detectors and keep them aloft for many hours. J. R. Winckler, Freier, and E. P. Ney developed this concept over the next 2 -1/2 years into one of the major projects carried out under the U.S. IGY program.

These several influences coalesced in my mind into a project that perhaps would meet my need to find a new research direction. I set to work developing the scientific objectives and the logistics for a series of balloon flights in the auroral zone and soon concluded that such a project was feasible to do if financial support could be obtained. By the end of November 1955, I had sent a proposal to the United States National Committee for the IGY. The proposal called for a large number of high altitude, long duration balloon flights to be carried out in 1957 at Fort Churchill, Canada, a site close to the auroral oval.

The scientific package I proposed for each flight would contain (1) a three-fold counter telescope using geiger-mueller tubes, (2) a single geiger-mueller tube, and (3) a scintillation counter consisting of a thallium activated sodium iodide crystal mounted on a photomultiplier tube. The scintillation counter would have high efficiency for detecting the "soft radiation" X-rays of energy above about 10 KeV. Its greater sensitivity might reveal many more "soft radiation" events and make it possible to infer something about the energy spectrum of the parent electrons. The budget totalled about $60,000. A serious problem then arose. The faculty and Administration policies at the State University of Iowa did not allow junior research persons the privilege of submitting research proposals to outside funding agencies, a policy common to many universities. James Van Allen must have removed this hurdle for me. I am enormously grateful to him for this essential support.

In late 1955, I was called to appear before the Technical Panel for Cosmic Rays of the United States National Committee for the International Geophysical Year. This group, headed by Scott E. Forbush, screened all proposals for cosmic ray projects to be conducted during that interval. Forbush explained that nearly all funds available for cosmic ray research had been allocated but he could find $15,000 if the Panel was convinced something useful could be done at that much lower level of funding. My response was that a reduced number of balloon flights could still produce interesting results on auroral zone phenomena and that I would send a revised proposal to the Panel. The Panel also stipulated that my payloads must fly two of the instruments that the University of Minnesota group was preparing for their IGY balloon flights. These instruments were a 10-inch diameter integrating ionization chamber of the Neher type and a single Geiger-Mueller counter and a stack of photographic emulsion. Since recovery of the payloads would be

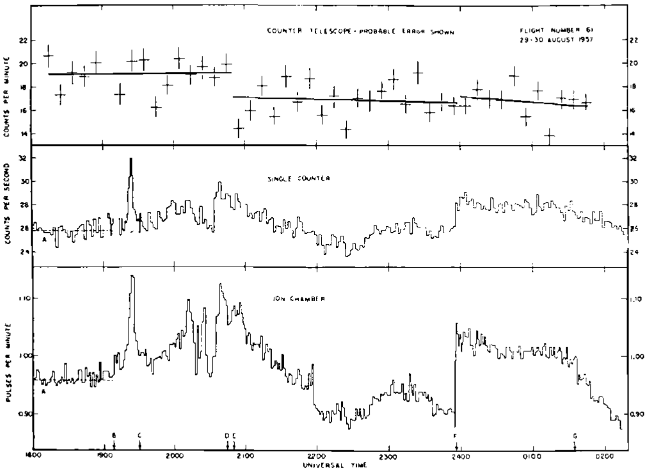


Figure 2. The responses of a single geiger counter, a coincidence telescope, and ionization chamber, during a balloon flight near the auroral zone. The short burst of X-rays near the beginning of a geomagnetic storm is, in hindsight, due to precipitation of electrons from the radiation belt caused by passage of a hydromagnetic wave through the magnetosphere. Here was another glimpse of the unknown magnetosphere. [Anderson, 1957]

(electrons) with a geomagnetic phenomenon, in this case the sudden commencement phase of a geomagnetic storm. Years later, the effect could be interpreted as the action of a hydromagnetic wave moving through the magnetosphere and causing the precipitation of electrons trapped in the Van Allen radiation zone. In 1953 Tom Gold [1955] attributed the sudden commencement effect to a "highly supersonic shock wave with the characteristic sharp wave front." In the summer of 1957 I met a co-worker of S. Fred Singer preparing a "wave detector" for flight on a high altitude balloon. I believe now that the objective was to detect the waves Gold had predicted.

(3) At other times during this flight, "soft radiation" did appear. Observers at Fort Churchill reported bright and active aurora at these times. This result confirmed and extended the 3 July 1957 observations by the Minnesota group.

(4) During three night flights when a balloon was at high altitude, quiet auroral arcs appeared in the sky but no X-rays were detected by the balloon-borne instruments. Quiet arcs evidently did not involve electrons of sufficiently high energy to produce X-rays that could penetrate to the atmospheric depth where the balloons were floating.

Some of these results were reported in a Letter to the Editor of the *Journal of Geophysical Research*. The manuscript was received on 25 November 1957 and was published in the December 1957 issue of that journal.

In early 1958 the physics laboratories at the State University of Iowa were in a high state of excitement and scientific activity. Explorer I had been launched in January and data from that spacecraft were being interpreted. In a hallway Carl McIlwain asked me if I could explain why the counting rate of a geiger counter would go to zero during a portion of the Explorer I orbit. I did not have the explanation. The correct answer as the satellite group soon realized, was that the spacecraft had encountered such high radiation levels that the operation of the G-M tube was completely blocked. When the discovery was made and carefully checked, Ernest Ray chortled loudly in the same hallway, "Space is Radioactive!"

Having had some success in demonstrating the scientific value of long-duration balloon flights in the auroral zone, I proposed a series of balloon flights for August and September of 1958, again at Fort Churchill. The proposal was written and submitted to the U.S. National Committee for the IGY in October 1957. It was accepted and Donald Enemark and I proceeded to make improvements in the electronic circuits during the winter of 1957-1958. The low temperature performance of germanium transistor amplifiers and scaling circuits was improved and the weight of the instrument package reduced by replacing the twin triode vacuum tube in the telemetry transmitter with silicon transistors just then coming on the market.

In Fort Churchill during August and September 1958, we made 10 balloon flights and accumulated 150 hours of high altitude data. Most of the launches took place in the lee of the large aircraft hangar. For certain wind directions we sought protection for the helium-filled bubble elsewhere. On one such occasion, the shelter was provided by the Fort Churchill elementary school. Normally, we would have launched the balloon at dawn, but strong winds delayed the launch until school had begun. Many small faces peered out at the strange activities on their playground.

We launched the first balloon on 10 August and obtained several hours of auroral zone X-ray activity in the single geiger counter and the ion chamber. An interval of windy, rainy weather ensued making balloon launches too risky to attempt. A brief interval of favorable weather beginning on 14 August allowed us to get a balloon off the ground a few minutes past midnight. The balloon reached its expected float altitude and the instruments provided 20 hours of data, but no effects above the normal cosmic ray background appeared. The wind and rain returned and balloon launches were again impossible. Several days went by without a launch, and we began to fall behind our plan to launch two

set of particle detectors for flight on the first Interplanetary Monitoring Platform (IMP-1). This spacecraft provided the comprehensive measurements of the outermost portions of the Earth's magnetosphere including the magnetotail to a radial distance of 32 Earth radii. In the tail we found many examples of energetic electron "island" fluxes which later investigations by other groups, particularly the Los Alamos group, showed to be related to dynamics of the plasma sheet. Another area of interest to beginning in 1964 was the phenomenon of electron acceleration at the bow shock of Earth, and the propagation of the electrons far upstream into the solar wind. Our work began on the IMP-1, -2, and -3 spacecraft and continued with the ISEE-1, -2, and -3 spacecraft well into the mid-1980s. A particularly significant finding was that the acceleration of electrons was most effective in a small region about the surface of contact between the interplanetary field lines and the bow shock surface. We also showed that at times these electrons remained gyrophase-bunched for large distances into the upstream region.

In Berkeley I was working with a group of gifted graduate students and I soon learned that spacecraft projects by themselves were not ideally suited for supplying Ph.D. thesis topics. The main problem was the very long times from experiment design to data plots. The plan developed at Berkeley was to have graduate students participate in some aspects of a spacecraft project (instrument testing, calibration, data analysis). They would then carry out an experiment in all its aspects using a balloon or rocket vehicle, most often in the auroral zone. For auroral zone investigations carried out from 1961 through 1966 support was obtained from the Office of Naval Research and the National Science Foundation and for rocket flights from Fort Churchill, the National Aeronautics and Space Agency. The Ph.D. research of Michael L. Lampton provides a good example of how graduate students participated in magnetospheric research. Balloons carrying scintillation counters to detect electron precipitation were launched at an auroral zone site. When X-ray microbursts appeared, a rocket carrying electron detectors was launched. The parent electron energy spectra was measured and motions of these electrons on the field lines inferred [*JGR*, 1967]. The microburst phenomenon had been discovered in 1963 [Anderson and Milton, 1964] and examples of this form of particle precipitation from the Earth's magnetosphere are shown in Figure 5. The auroral zone balloon and rocket project was largely taken over by graduate students in the years 1964-1966.

In the mid-1960s the auroral concept of Akasofu [1964] was generalized to become the magnetospheric substorm. Jelly and Brice [1967] noticed that auroral zone particle precipitation "is intimately associated with large scale processes that occupy a substantial part of the magnetosphere." Coroniti, Parks, McPherron, and Anderson [1968] in a series of papers showed that the microbursts, other types of particle precipitation, and certain geomagnetic micropulsations also fit into a definite local time pattern further generalizing the magnetospheric concept. They suggested removing the word elementary from the term introduced by Jelly and Brice.

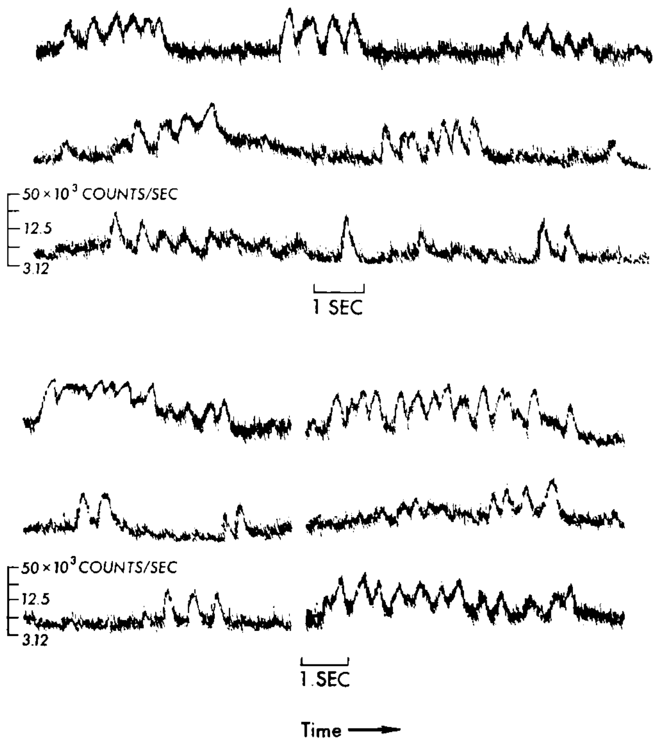


Figure 5. Scintillation detectors of high Z material for efficient detection of auroral zone X-rays revealed a remarkable form of electron precipitation. These ~ 1/4 second long microbursts arrive singly, in pairs or in groups of 4 to 6 or more. Groups of these elementary bursts may recur at periods of 5 to 15 seconds. The nature of this form of magnetospheric particle precipitation is not entirely known and is the subject of current studies. [Anderson and Milton, 1964]

3. EXPERIMENTAL TESTS OF DUNGEY'S HYPOTHESIS AND MEASUREMENTS OF ELECTRIC FIELD IN THE MAGNETOTAIL (1965-1975)

In 1965 James Van Allen and S. M. Krimigis published "three clear cut cases of ~40 keV electrons from the Sun." This was a highly significant discovery and the authors' belief that "Such electrons contribute a new tool for the study of the interplanetary magnetic field. . ." has been fully justified. In Berkeley, Robert Lin and I [1966] found these same events and nine more in the IMP-1 data. We also found that the electrons reaching the Earth were quite directional (non-isotropic) and not behaving at all as predicted by an isotropic diffusion model [Anderson and Lin, 1966]. Furthermore, most of the parent flares of the electron events were centered about W60° solar longitude about where the average Parker spiral field line reaching Earth would root into the Sun. This result led us to the idea that these particles, moving at about 1/3 the speed of light with small gyroradii (~ 100 km), could be used as magnetic field line tracers in the interplanetary and geomagnetic tail region.

At this time the issue of open versus closed models of the magnetosphere was being vigorously discussed. We thought that if solar-interplanetary field lines were actually connected to geomagnetic field lines as Dungey had suggested in 1961, that the fast solar electrons, following field lines like beads on a wire, would have easy access to the Earth's magnetic tail. We compared the flare electron travel times from Sun to a spacecraft outside the Earth's magnetosphere to the travel times when the spacecraft had moved out of the Earth's magnetosphere. From a statistical study of a considerable number of solar flare electron events, observed both inside and outside the magnetotail, we concluded (1) at least some of magnetotail field lines connect with interplanetary field lines behind the Earth. (2) The connection region (tail "length") must be $\gtrsim$ 0.2 AU (5000 R_E) distant from Earth. (3) Quoting from our 1966 paper [Lin and Anderson, 1966], "The upper limit [of 5000 R_E] we have arrived at can be greatly reduced by means of simultaneous observations on two satellites near Earth, one in the tail and one in interplanetary space."

In mid-1967 the Explorer 35 spacecraft, also known as lunar anchored IMP, was placed in orbit about the moon. Among the on-board instruments were energetic particle detectors built in Berkeley. Using data from these instruments, Robert P. Lin [1968] published several examples of "shadowing" of solar electron fluxes by the Moon in the solar wind and in the geotail. He showed that the shadowing of the fast electrons was complete when the Moon was in the Earth's magnetotail and that the full shadow appeared on the Earthward side of the Moon. From these results Lin concluded that ". . .these [solar] electrons gain direct entry into the tail at a distance greater than 64 R_E behind the Earth."

In their analysis of lunar shadowing of solar particles Van Allen and Ness [1969], published a large number of examples of full shadows for both the solar wind case and the geotail case. They also did a conceptual analysis of what happens when field lines carrying an isotropic distribution of fast solar particles of small gyroradius is transported across (a) a non-magnetized, non-conducting uniform sphere; and (b), a highly conducting sphere. Analysis of case (a) confirmed that a full shadow would develop Earthward of the Moon. Their analysis also predicted that a half shadow would appear on the opposite side of the Moon (Figure 18 of their paper). That they did not observe the predicted half-shadows in their Explorer-35 data was because the radiotelemetry signal from the satellite was usually occulted by the Moon when half-shadows would be expected. At Berkeley we found a few occasions when the radio signal was able to reach Earth and we were able to demonstrate the existence of the half-shadows [Anderson and Lin, 1969]. In their paper Van Allen and Ness set an upper limit to any electric field that might exist in the geomagnetic tail. The estimate was based on the sharpness of the shadow edges and came out to be

$$E_{\perp} \leq 5 \times 10^{-4} V/m \, .$$

A full theoretical treatment of particle shadowing by the Moon (or other non-magnetized, non-conducting spherical object) was later carried out by Robert McGuire [1972].

Continuing our attempts to test the Dungey model, Robert Lin and I looked in the particle shadow data to see if we could determine the location of the reconnection region in the geotail. If the spacecraft were on a field line that had just reconnected a uniquely different shadowing pattern would appear. The pattern would be two half-shadows each revolution about the Moon. We found one such example and concluded that reconnection at times may occur less than 60 R_E behind the Earth [Anderson and Lin, 1969].

By early 1970 Robert Lin and I had a large catalog of solar flare electron events, many of which had anisotropic, field aligned pitch angle distributions. In looking for other field line tracing applications for these particles I realized that a large-scale electric field with a component perpendicular to the magnetotail field lines would produce an orderly (non-diffusive) field drift motion of the fast electrons across the magnetotail field lines. Electrons entering the tail via merged field lines would pass the Moon and move toward Earth where they would encounter stronger fields and mirror, returning to the vicinity of the Moon. But due to the $\vec{E} \times \vec{B}$ drift their path would be shifted in the $\vec{E} \times \vec{B}$ direction. This motion would then displace them into the geometric particle shadow produced by the Moon. Such electrons would be distinguished by their motion away from the Earth. This situation is shown in Figure 1 of the paper "Method to determine sense and magnitude of electric field from lunar particle shadows" [Anderson, 1970]. At that time the particle measurements were not adequate to test the method and make quantitative measurements of the electric field. With the Apollo Subsatellite data James McCoy and others would obtain many quantitative vector determinations of the electric field in the magnetotail using this method.

4. THE APOLLO 15 AND 16 PARTICLES AND FIELDS SUBSATELLITE

By the spring of 1969 knowing that the lunar shadowing of the solar electrons could lead to conclusions about the interaction of interplanetary and geomagnetic field lines, we looked for ways to obtain additional and more refined measurements in the vicinity of the Moon during times it was crossing the geomagnetic tail. About that time we saw an Announcement of Opportunity (AFO) from NASA for scientific participation in the Apollo program. Perhaps we could use the Apollo program to continue our magnetic topology studies and also learn something about the energetic particle environment of the Moon. The measurements we needed to study such phenomena required months of observing time since the solar electron events occurred at the average rate of one a week. But the Apollo command module stayed in lunar orbit only a few days, and the Lunar Excursion Module was not able to return data to Earth.

A possible solution was hit upon late one Friday afternoon in March 1969 after work in a pizza and beer establishment named La Val's a half-block off the North Side of the Berkeley campus. This place was much frequented by University students and staff. Our research group often went there to celebrate a student's success in an examination or the award of a Ph.D. degree. On this particular Friday afternoon I recalled how Khrushchev ridiculed the United States Space Program in 1958 when the Vanguard project had tried to launch a small (about 12" diameter) scientific satellite and failed in the attempt. He referred to the small satellite as a "grapefruit." I wondered out loud if we could build several "grapefruit-size" satellites containing our particle detectors, then hand them over to the astronauts. In lunar orbit they could open a window, reach out and let go of these little spacecraft. They would then remain in low altitude lunar orbit taking data for many months after the astronauts returned to Earth.

There was some laughter but we soon began to think how this notion might be turned into scientific reality. Within a very few days Robert Lin, Lee Chase, Richard Paoli, and I had sketched out the instrument design and roughly estimated power, weight and telemetry requirements for a small spacecraft carrying the electron detectors. Our first proposal was dated March 1969, and a revised version of it was submitted to NASA Headquarters on 1 April 1969. Its title was *"Description of an experiment to explore the topology of the Earth's magnetosphere."* We soon realized we had not gone far enough in exploiting the idea of small scientific spacecraft in low lunar orbits launched from the Apollo spacecraft. A magnetometer would be required to do definitive studies of the magnetic field near the Moon. We contacted Professor Paul Coleman and his colleagues at the University of California, Los Angeles. They quickly agreed to provide a flux-gate magnetometer for the subsatellites. Later Dr. William Sjogren of NASA's Jet Propulsion Laboratory joined the experiment team. His interest was to do very precise tracking of the Fields and Particles subsatellite orbits using the S-band telemetry signal from the small spacecraft. Such information would enable him to locate and characterize lunar mass concentrations (Mascons) beneath the ground tracks of the subsatellites.

Our proposal moved quickly through NASA Headquarters and the Manned Space Flight Center between late April and early July. We were called in early July 1969 and asked to give an oral presentation to a review panel to be convened at the Los Alamos National Laboratory in New Mexico just two weeks later. Evidently, our presentation of the science objectives to the selection panel was successful -- only a few days later we had a go-ahead for our project from NASA Headquarters and soon we were working directly with a group at the Manned Space Flight Center. We added a scientist from the Manned Spacecraft Center to our team. James McCoy was very helpful in communication between Berkeley, the manned spaceflight center, the subcontractors and MSC. Later McCoy would be the lead author on two of the most important papers to come out of our Apollo work. MSC turned the implementation of the Particle & Fields Subsatellite project over to Thompson, Ramo, Wooldridge Corporation. Events moved at extraordinary speed from mid-July to November 1969. By November, TRW had essentially completed the spacecraft design. But the launch of Apollo 15 was only 19 months away and instruments and spacecraft subsystems were yet to be built and tested. One subsatellite would be flown on Apollo 15, a nearly identical one on Apollo 16. A third one was built and we hoped it would be put in orbit about the Moon by Apollo 17 but that did not happen. The third P & FS subsatellite now resides in the National Air and Space Museum in Washington D.C.

The Apollo 15 subsatellite was ejected from the service module by action of a coil spring on 4 August 1971. As the small spacecraft receded from the Apollo command module it was observed and photographed by the astronauts using a hand held Hasselblad camera and a 16 mm movie camera. The first P and FS made about 2000 revolutions about the

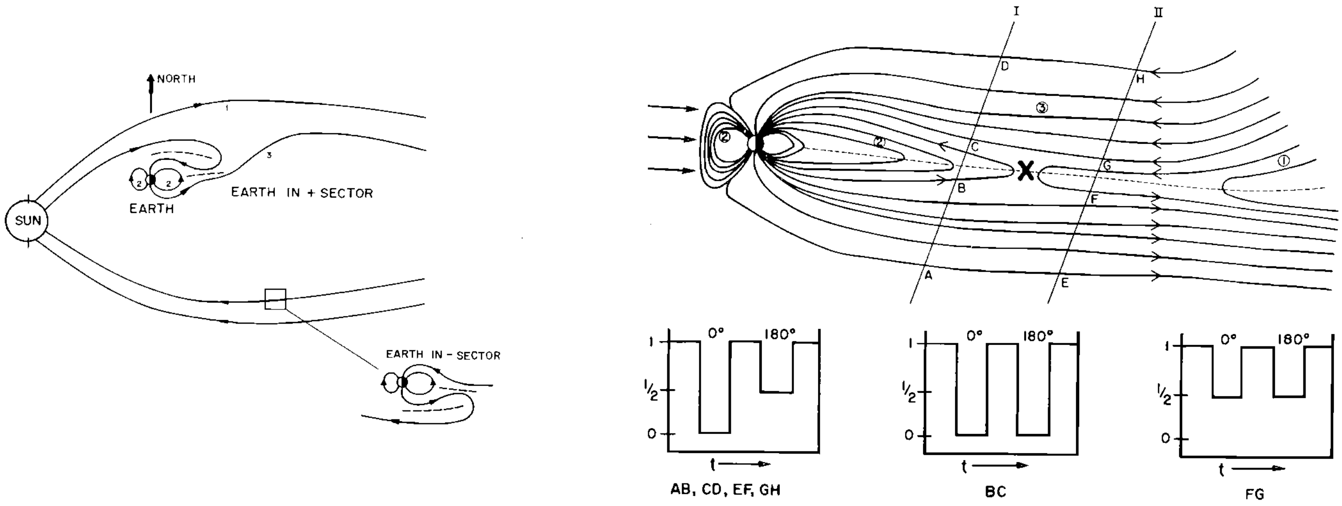


Figure 6a. The left hand sketch shows three types of field lines implied by the Dungey reconnection hypothesis. Two of the three neutral sheets required by this topology are indicated. The right hand sketch shows the three types of shadows predicted from the Dungey hypothesis. All three types have been observed on the Explorer-35 and Apollo 15 and 16 subsatellites. Study of these shadows revealed an instance when Dungey reconnection was occurring Earthward of the Moon (i.e., at a geocentric distance less than 60 Earth radii). [Anderson and Lin, 1969]

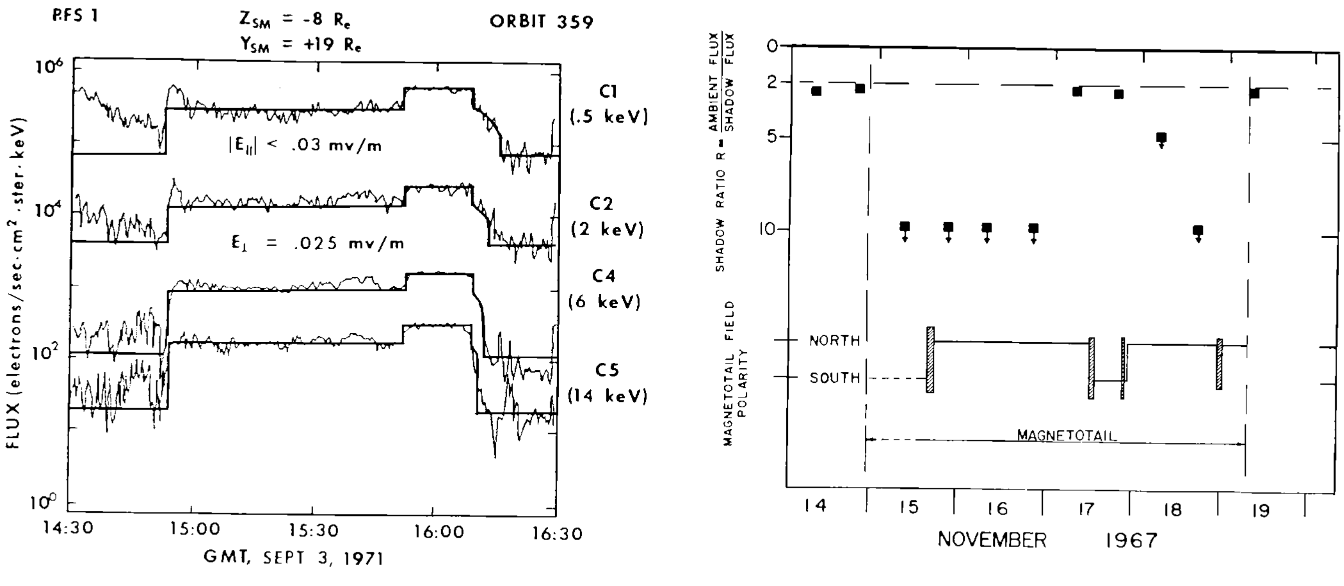


Figure 6b. The heavy line in the left hand plot is the theoretical fit to electron intensities for one revolution about the Moon by the Apollo 15 subsatellite. Electrons drifted into the Earthward facing full shadow are to the right of the "shoulder." In the absence of electric fields a symmetrically located shoulder would appear. The shoulders are sensitive indicators of the presence of an electric field. The right hand plot shows the shadow type encountered as the Moon moves from interplanetary space through the geotail and back into interplanetary space. The two double half shadows near the center of the geotail show Dungey reconnection is occurring Earthward of the Moon. [McCoy, 1975; Anderson and Lin, 1969]

Moon. Our first publication of P and FS results in the *Journal of Geophysical Research* [1972] included a color image of the spacecraft just after ejection from the science module. We reported on the lunar shadowing of the solar wind and particle energy spectra in and outside of the shadows. The second P and FS went into orbit on 24 April 1972. Because of a problem in the main engine gimbal control system, that engine could not be used to place the command module in an orbit that would give the subsatellite a long lifetime. Mascon perturbations caused it to crash on the far-side of the Moon after 425 revolutions during its 35 day lifetime. Several improvements had been made in the second P and FS so the data were of very high quality. The subsatellite made more than a full revolution about the Earth and crossed the geomagnetic tail once and spent many days in the magnetosheath and solar wind. We thus had many opportunities to study the Moon's shadowing of solar particles in considerable detail. And we would achieve a totally unexpected result which would carry our group into a completely new area of research.

Despite the degraded performance of PFS-1 and the short lifetime of P and FS-2, we were able to achieve the major objective we had set for the subsatellites in 1969: measurement of the large scale electric field in the Earth's magnetotail. The existence of a dawn-to-dusk directed electric field had been inferred but never directly measured far into the Earth's magnetotail. James McCoy and co-authors [1975] demonstrated that a dawn-to-dusk directed electric field was continuously present over the several day period while the spacecraft were in the magnetotail. The magnitude of the electric field ranged from 0.2 to 2 millivolts per meter (0.2 to 2 volts per kilometer). The typical or average field strength was about 0.15 v/km. Multiplying the average field value by the known width of the magnetosphere gave a total cross tail electric potential of about 40 kilovolts in good agreement with indirect ionospheric measurements (Refs). Our results also showed that the electric field strength varied with time, an effect attributed to variations in the convective motions of the magnetotail plasma.

By 1971 the Space Physics Group at the Space Sciences Laboratory on the Berkeley campus had participated in a many investigations on Earth-orbiting spacecraft, on sounding rockets and high-altitude balloons. The rocket and balloon programs had been conducted mainly in the Northern auroral zone. At the time of the Apollo 15 launch and the orbiting of its subsatellite about the Moon, the capability to mount rocket and balloon field operations was well in hand and could be mobilized in a matter of a few months. It was therefore natural for us to try to think of experiments with these lower-altitude vehicles that could extend the scientific value of the two Particles and Fields Subsatellites.

We linked the scientific goals for the subsatellites to our interests in auroral zone and polar cap research using balloons and rockets in the following way: The Moon carries the subsatellite across the geomagnetic tail once each revolution of the Moon about the Earth. The geomagnetic field lines at the times the moon is in the geotail are rooted in the polar caps of the Earth. We would launch rockets and balloons carrying a variety of detectors identical or similar to those on the subsatellite at times when the geotail field lines were more or less connecting the two instrument sets, a distance of about 60 Earth radii apart. This four day time interval is centered on the time of full moon. Charles Carlson and Lee Chase instrumented a Black Brant V rocket with particle instruments very similar to those carried on the subsatellite. Forrest Mozer and his group supplied a boom system to measure A.C. and D.C. electric fields. These instruments were quite similar to his electric field detectors carried to 100,000 feet altitude by balloons. We would then launch the rockets and balloons from a site in Earth's North polar cap when the subsatellite was in the magnetic tail of Earth and thus in the same bundle of field lines as the rocket and balloons. The site we chose was Resolute, a Canadian base on Cornwallis Island in the Arctic Ocean. Resolute is only about 150 km from the Earth's northern magnetic dip pole and thus well inside Earth's northern polar cap. A three person rocket launch crew from the Bristol Corporation, manufacturer of the rocket would carry out the launch operations and track the rocket. Mozer's team was ready to launch a balloon when the right moment came. That moment came on 5 September 1971. Balloons were already in the air above Resolute and Thule, Greenland, when the rocket was fired (Figure 7). All experiments worked well. We were able to show that the electric potential difference along the to magnetotail field lines between the polar cap and the Moon could not have exceeded 500 V showing that at this time the parallel component of electric field in the geotail was very small - less than 1.3×10^{-6} kv/km [McCoy et al., 1976].

In the summer of 1973 I was visiting the Centre d'Etudes Spatiales Rayonnements (CESR) a research laboratory connected with Université Paul Sabatier and supported by the French space science agency, CNES. While there, Robert Lin informed me of a remarkable development in the analysis of data from the two P and FS subsatellites. We had been puzzled by rather large fluxes of electrons unexpectedly entering the particle detectors from the direction of the Moon's surface and on many occasions the intensity of electrons coming up from the Moon's surface approximately equalled the intensity of the solar electrons headed toward the Moon. Herbert Howe and Robert Lin found that the large, upward-moving electron fluxes were correlated

the merging-reconnection hypothesis; a search for correlation between geomagnetic activity and the presence of a southward component in the interplanetary field. The search for such correlations has been pursued intensively by many groups with apparent success. Lin and I believed this was important to do but a more direct physical test was needed if only to provide an independent check. So in 1966 when Lin published the result of rapid access of fast and anisotropic distribution of electrons into the geotail we believed we had found a much needed direct experimental test. And we believe that the lunar shadowing of the fast, small gyroradius solar electrons did reveal the fundamental topology of the interplanetary magnetic field-geotail magnetic field system. The method also showed that the field line reconnection in the geotail required by the Dungey hypothesis on at least one occasion took place between the Earth and Moon.

Acknowledgments. The work referred to in this article was supported by the Office of Naval Research, the National Aeronautics and Space Administration, and the National Science Foundation. Many co-workers contributed to the success of the many experimental activities; undergraduates and graduate students, post-doctoral researchers, engineers, technicians, and administrative support persons.

REFERENCES

Akasofu, S.-I., *Planetary Space Sci.*, *12*, 273-282, 1964.

Anderson, K. A., Occurrence of soft radiation during the magnetic storm of 29 August 1957 (Letter to Editor), *J. Geophys. Res.*, *62*, *641*, 1957.

Anderson, K. A., Method to determine sense and magnitude of electric field from lunar particle shadows, *J. Geophys. Res*, *75*, 2591, 1970.

Anderson, K. A., and D. W. Milton, Balloon studies of X-rays in the auroral zone, 3. high time resolution studies, *J. Geophys. Res.*, *69*, 4457, 1964.

Anderson, K. A., and R. P. Lin, Observations on the propagation of solar-flare electrons in interplanetary space, *Phys. Rev. Lett.*, *16*, 1121, 1966.

Anderson, K. A., and R. P. Lin, Observation of interplanetary field lines in the magnetotail, *J. Geophys. Res.*, *74*, 3953, 1969.

Anderson, K. A., L. M. Chase, R. P. Lin, J. E. McCoy, and R. E. McGuire, Solar wind and interplanetary electron measurements in the Apollo 15 subsatellite, *J. Geophys. Res.*, *77*, 4611, 1972.

Anderson, K. A., R. P. Lin, and R. E. McGuire, Linear Magnetization feature associated with Rima Sirsalis, *Earth and Planet. Sci. Lett.*, *34*, 141, 1977.

Coroniti, F. V., R. L. McPherron, and G. K. Parks, *J. Geophys. Res.*, *73*, 1715-1722, 1968.

Gold, T., Gas dynamics of cosmic clouds, Symposium held at Cambridge, England, July 6-11, 1953, North Holland Publishing Company, Amsterdam, 1955.

Howe, H. C., R. P. Lin, R. E. McGuire, and K. A. Anderson, Energetic electron scattering from the lunar remanent magnetic field, *Geophys. Res. Lett.*, *1*, 101, 1974.

Jelly, D, and N. Brice, *J. Geophys. Res.*, *72*, 5919, 1967.

Lampton, M., Daytime observations of energetic auroral-zone electrons, *J. Geophys. Res.*, *72*, 5817, 1967.

Lin, R. P., Observations of lunar shadowing of energetic particles, *J. Geophys. Res.*, *73*, 3066, 1968.

Lin, R. P., and K. A. Anderson, Evidence for connection of geomagnetic tail lines to the interplanetary field, *J. Geophys. Res.*, *71*, 4213, 1966.

McCoy, J. E., R. P. Lin, R. E. McGuire, L. M. Chase, and K. A. Anderson, Magnetotail electric fields observed from lunar orbit, *J. Geophys. Res.*, *80*, 3217, 1975.

McCoy, J. E., R. P. Lin, F. S. Mozer, R. E. McGuire, L. M. Chase, and K. A. Anderson, Comparison of simultaneous magnetotail and polar ionospheric electric fields and energetic particles, *J. Geophys. Res.*, *81*, 2410, 1976.

McGuire, R. E., A theoretical treatment of lunar particle shadows, *Cosmic Electrodynamics*, *3*, 208, 1972.

Meredith, L. H., M. B. Gottlieb, and J. A. Van Allen, Direct detection of soft radiation above 50 km in the auroral zone, *Phys. Rev.*, *97*, 201, 1955.

Parks, G. K., F. V. Coroniti, R. L. McPherron, and K. A. Anderson, *J. Geophys. Res.*, *73*, 1685-1696, 1968.

Pomerantz, M. A., and G. W. McClure, *Phys. Rev.*, *86*, 536, 1952.

Van Allen, J. A., and S. M. Krimigis, Impulsive emission of ~40 - keV electrons from the Sun, *J. Geophys. Res.*, *70*, 5737, 1965.

Van Allen, J. A., and N. F. Ness, Particle shadowing by the Moon, *J. Geophys. Res.*, *74*, 71, 1969.

Winkler, J. R, and L. Peterson, Large auroral effect on Cosmic-Ray detectors observed at 8 g/cm^2 atmospheric depth, *Phys. Rev.*, *108*, 903, 1957.

Kinsey A. Anderson, Space Sciences Laboratory, University of California, Berkeley, CA 94720-7450

The Boundary and Other Magnetic Features of the Magnetosphere

Laurence J. Cahill, Jr.

School of Physics and Astronomy, University of Minnesota, Minneapolis, Minnesota

This is a personal account of the discovery of the boundary of the magnetosphere. The concept of a boundary and of related ideas about the magnetic field and plasma around the earth were gained at the University of Iowa. My first investigations of currents in the ionosphere were also done there. An account of the Iowa years, of participation in expeditions to launch rockets and of the findings from the earliest satellite magnetic observations in space is given first. In New Hampshire, the Explorer 12 preparation and launch follow, with discussion of the boundary discovery. Next are described the first observations of other magnetic features of the magnetosphere, obtained through data from Explorer 12 and later spacecraft. Finally, the discovery of the location and cause of the ring current is briefly mentioned.

INTRODUCTION

The author was directly involved in the discovery of the magnetosphere boundary and in the study by satellite magnetometer of other features of the magnetosphere, such as the location of the ring current. Before making discoveries one first must become familiar with the ideas and actions of other people. My ideas about the earth's magnetic field and about the environment in space about the earth began forming in the mid 1950s. By 1960 it was possible for me to manage the preparation and operation of an instrument, a three-axis, flux-gate magnetometer, used to observe the boundary of the magnetosphere. Through the 1960s we were able to observe in space, in rapid succession, other features and properties of the mag-netosphere that had been predicted earlier as the result of ground-based geomagnetic research.

In the first part of the report the years spent at the University of Iowa in the late 1950s will be covered, when the events of the earliest times of the satellite era occurred and where my ideas about the nature of the solar wind/magnetosphere interaction developed. The second part will be concerned with my direct responsibility for preparation and launch of several satellite magnetometers and for the interpretation of the scientific results from these instruments. This work took place at the University of New Hampshire in the 1960s. An earlier review article, limited to the boundary of the magnetosphere only, was prepared for a conference on the physics of the magnetopause [*Cahill,* 1995].

Discovery of the Magnetosphere
History of Geophysics Volume 7

AT THE UNIVERSITY OF IOWA

It was the summer of 1957 at the University of Iowa. The International Geophysical Year [IGY] was underway and each person in our small research group had a project. Carl McIlwain was preparing to study "soft radiation" using ground-launched rockets at Ft. Churchill [*McIlwain*, 1960]. The University of Iowa group had been studying this soft [low energy] radiation for several years with balloons and balloon-launched rockets [rockoons]. The radiation was found in the auroral zone, at balloon and small-rocket altitudes, apparently associated with auroral displays. Evidence was mounting that it was due to bremsstrahlung X-rays, produced by electrons causing the auroral displays.

Frank McDonald was planning balloon flights to continue study of cosmic rays. Kinsey Anderson had an elaborate set of balloon-lifted detectors to study cosmic rays and the soft radiation. Ernest Ray was working on a theoretical cosmic-ray project. Perhaps the most unusual project was assigned to George Ludwig. A Geiger-Mueller [GM] tube apparatus was intended for flight on one of the first U.S. artificial satellites, to be launched during the IGY by the Naval Research Laboratory. A new rocket, the Vanguard, was being designed and built to carry the scientific payloads into orbit. The Iowa experiment was to be carried in a near-earth orbit to study soft radiation, as well as cosmic rays, above the altitudes reached by previous Iowa flights. The satellite would also provide global coverage rather than the limited spatial coverage given by a balloon or rocket flight. In addition to the GM tube, George was building a miniature tape recorder to store the accumulated data [*Ludwig*, 1959]. The present world-wide network of receiving stations wasn't in full operation then and the data were to be stored for occasional readout when the satellite was near one of the few receiving stations. My sounding-rocket project was planned for study of electrical currents in ionosphere, particularly the equatorial electrojet, a sheet of electrical current flowing in the lower ionosphere near the magnetic equator. First, it was necessary to adapt the recently-invented, proton-precession magnetometer for rocket use.

George, Carl, Ernest and I were graduate students; Frank and Kinsey were post-doctoral research associates. We worked in close proximity in the basement of the old Physics building , discussed our work and problems, went to lunch and coffee breaks together and shared a sense of challenge and excitement as we prepared to go out and make measurements. Frank and Kinsey managed the lab and the students and were very accessible for advice. In overall charge of our enterprise was Professor James Van Allen, self-described as the "scoutmaster". He determined the direction of the research and found support. He also provided the graduate students with research projects. He was busy with teaching and adminstrative duties, as Department Head and director of the research lab, but was always available for advice on major problems and for long-term guidance. Of the greatest value for research training was his policy of giving each student as much responsiblility as the student could handle.

Of course, all of the graduate students were engaged in the standard graduate-level physics courses. In addition there were Physics colloquia and seminars with frequent presentations by visitors engaged in interesting research projects in cosmic rays, atmospheric physics and other subjects that would now be grouped together as "Space Physics". One semester we had a very distinguished visitor, Professor Sydney Chapman. A mathematician by training he had a long and productive career in applied mathematics and also in geophysics. He gave us a research seminar in geophysical topics, including his ideas on the causes of magnetic storms. Because I had been given a research project that involved detecting the magnetic effects caused by electrical currents in the ionosphere, it was of particular interest to me to study the two-volume treatise, "Geomagnetism", by *Chapman and Bartels* [1940]. My wife bought the books for my birthday and I worked slowly through them, reading as a change of pace from the graduate course study. Several figures in the 2nd volume caught my interest. They showed the effects of a neutral, but ionized, stream of particles from the sun, impacting on the earth's magnetic field. According to Chapman and Bartels, the charged particles of the stream would be deflected by the magnetic field but the field would be compressed as the stream was slowed. A boundary would be established between the ionized stream and the earth's magnetic field. The stream would flow around the earth's magnetic field forming a cavity in the flow. One figure from Chapman and Bartels is reproduced here as Figure 1.

The boundary between the earth's magnetic field and the ionized stream from the sun was first proposed by *Chapman and Ferraro* [1931, 1932] to explain some observed features of magnetic storms. Magnetic storms were known to be associated with solar flares and ionized streams from the flares were assumed to travel to the earth in a day or two. Compression of the earth's field by the initial impact of the stream was thought to produce the sudden world-wide increase in the field called the *sudden commencement*. Continued compression of the earth's field by the flowing stream was thought to correspond to the world-wide, positive, initial phase of the storm. The following, negative, main phase of the storm was thought due to a westward-flowing ring current, some how set up by the ionized stream. Chapman and Ferraro were not entirely convincing in their main phase explanation and the cause of the the negative phase was still under discussion in 1957.

The ionized stream from the sun was thought to flow only at times of solar flares and at the beginning of the related magnetic storms. At other times the earth's magnetic field would not be compressed and would assume its approximate dipole shape. *Parker* [1959] proposed a more or less continuous *solar wind* of ionized gas from the sun, an outflowing of the solar corona as an ionized plasma, composed of electrons and protons. A continous flow of ionized gas offered an always-present boundary. The earth's magnetic field would always be enclosed in a cavity, surrounded by the continuous, outward-flowing, solar wind. Fluctuations in the solar wind intensity would produce changes in the size of the cavity. Large outflows of solar wind, at times of magnetic storms, would squeeze the geomagnetic field into a smaller cavity and cause the observed, initial, magnetic storm effects.

One of the major research expeditions of our lab in 1957 was a rocket launching trip on the USS Glacier, a Navy icebreaker, from Boston, through the Panama Canal, across the Pacific and then down to McMurdo Base in the Antarctic. Professor Van Allen and I were to conduct the launchings with the assistance of Steven Wilson, a Navy Lieutenant, Junior Grade, assigned as our liaison with the Navy. He arranged for other Navy people to help us as needed. We had 20 Loki rockets, each 3" in diameter and about 9' long, including solid-propellant rocket and payload. The rockets were to be launched as rockoons; balloons carried the rockets to 70,000 feet in altitude, then the rocket fired and took the payload up to 120-130 km. I was to send rocket magnetometers up through the equatorial electrojet as we crossed the magnetic equator. With his rockoons, carrying GM tubes as the principal sensors, Professor Van Allen intended a latitude survey of soft radiation and the cosmic radiation as we proceeded from the equator to the Antarctic. Earlier in 1957 we had conducted similar flights while sailing north, through the Davis Straits, up to Thule in Greenland. The projects amounted to studies of radiation and electrical currents in the bottom side of the magnetosphere, but that term hadn't yet appeared.

We started launching rockoons from the heliocopter deck on the stern of the icebreaker as we approached Christmas Island, south of Hawaii. The first launch was not auspicious. We had filled the balloon with sufficient helium and were ready to release it. By that time the ship's direction and speed should have been adjusted to be traveling with the wind, so that there was zero wind over the launch deck. The balloon, tethered by a line to the rocket and payload, should be straight overhead, ready for release. Unfortunately, despite our instructions, the people on the bridge hadn't got it quite right. The line to the balloon was trailing to the stern, at an angle of 45° to the vertical. In past flights, off Greenland, Professor Van Allen had been on the bridge to counsel the Navy people on how to obtain zero wind. He was confined to his bunk for this launch, however, after gashing his shin in a fall while we were loading helium cylinders in Panama. The wound had got infected and the ship's doctor had required him to be completely immobile. Meanwhile, I was holding the rocket and attached firing mechanism box in my arms and was providing ballast to prevent the balloon from rising. I couldn't release the rocket since it would surely swing and hit the deck or some other part of the ship before the balloon rose high enough. Zero wind was essential so the balloon and payload could go

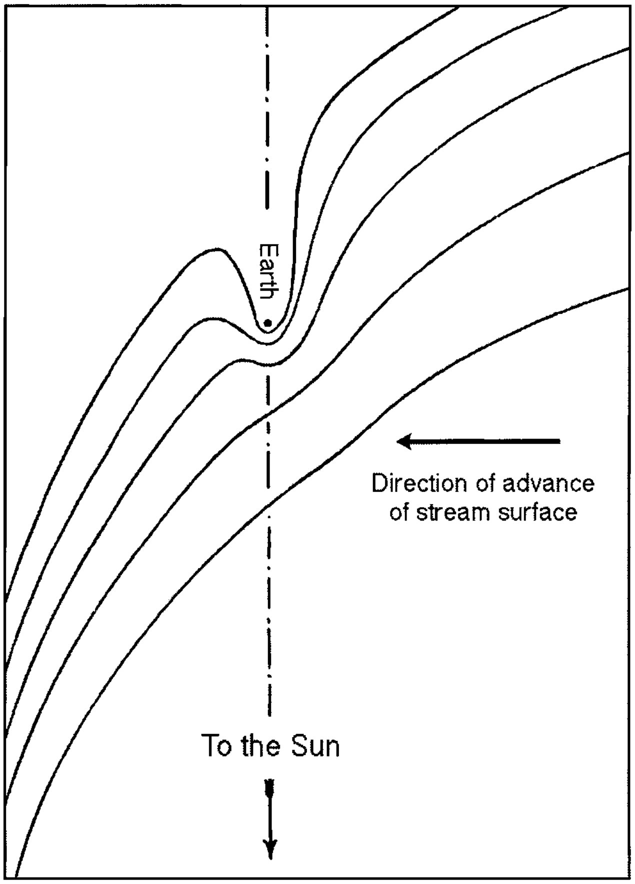


Figure 1. The lines show the advancing front of a stream of ionized, but neutral, gas from the sun. The earth's magnetic field is forming a cavity in the stream as it sweeps past the earth [p. 856, *Chapman and Bartels*, 1940].

straight up. Steve Wilson took off for the bridge and I hung on as the balloon pulled me toward the stern. As the helmsman tried to adjust course the situation got worse; the angle increased to 60°. The tug of the balloon was considerable and I moved slowly toward the edge of the flight deck. By this time the balloon had moved to the port side of the ship, but there seemed to be little progress toward zero wind. Standing at the edge of the flight deck and leaning backward to counteract the balloon tug, I felt somewhat uneasy. As several more minutes passed with no improvement in the angle, I began to consider releasing the rocket. From where I was standing the rocket would swing out under the balloon as the balloon rose. It would not hit any part of the ship but would the balloon rise fast enough so it wouldn't hit the water? After several more minutes with no improvement, I released the rocket. The rocket swung in an arc toward the water. The balloon rose sluggishly and the firing box hit the water first. Next, a wave caught the bottom of the rocket and the balloon stopped rising. Eventually the inflated bubble of the balloon sat on the water as we moved away. We conducted a thorough seminar on achieving zero wind before the next launch.

Some of my rocket magnetometers failed after launch but good data was received from several others, launched as we crossed the magnetic equator. We were surprised to hear, through the Navy communications network, of the launching of the first artificial satellite, the Russian Sputnik, on Oct. 4, 1957, as we cruised near the equator. We had assumed the U.S. Vanguard program would launch the first artificial satellites. Although the Russians had announced plans for a satellite program, it was thought we were comfortably ahead of them. There was a mild sense of shock while we listened to the Sputnik beeping overhead as it circled the earth. A series of radio messages traveled back and forth between the Glacier and the U.S. Because the upper atmosphere groups considered that we were far in the lead in rocket capability, there was no sense of urgency for the first satellite launch date. Vanguard was proceeding with due care and deliberation; the first launch was planned for some months in the future. Sputnik introduced a sense of urgency. There was another competent, large-rocket group in the U.S. that was also quite ambitious for the honor of launching a science payload into earth orbit, Von Braun's group at the U.S. Army's Huntsville base in Alabama.

Correspondence between the Glacier, the Jet Propulsion Laboratory, Huntsville, and Iowa was very active. George Ludwig was finishing up his M.S. thesis project, the GM tube with tape recorder designed for Vanguard launch. There was already a plan to include an Iowa GM package, as a later backup to the Vanguard project, on an Army rocket. A new idea emerged, to launch the Iowa science payload, originally intended for Vanguard, with an Army rocket. The Alabama group would be responsible for preparing the launch rocket; the Jet Propulsion Laboratory would provide a satellite structure with telemetry, power and other supporting functions. George Ludwig and Ernest Ray conducted some of the negotiations for Iowa, keeping Professor Van Allen informed by Navy radio. Van Allen approved the proposal with JPL, all the details were worked out and the Iowa package [but without the tape recorder] was prepared as the payload of Explorer 1.

Our rockoon program was completed in the Antarctic with reasonable success. We returned to New Zealand with the Glacier and flew home to Iowa in the fall of 1957. With all the IGY rocket launching activity in 1957, there wasn't much progress for me in completing required graduate courses. I returned to graduate study with gusto and also commenced analyzing the magnetic tapes with electrojet results. George Ludwig's GM tube was launched on Explorer 1 on February 1, 1958. A similar payload, this time including the tape recorder, followed on Explorer 3, launched on March 26, 1958. Data from the Explorers started arriving at the University and analysis revealed unexpected results. As these results appeared, Carl McIlwain and Ernest Ray joined George and Professor Van Allen in studying and interpreting the data. Intense radiation was observed, increasing rapidly with altitude, from above 700 km up to the highest altiudes reached, about 2000 km, and at all magnetic latitudes traversed by the satellite [*Van Allen et al.*, 1958]. The radiation appeared to be related to that observed in auroral latitudes by the earlier University of Iowa rocket flights. After some considerable deliberation, it was realized that the radiation must be due to charged particles, trapped on and circling in the earth's magnetic field. The particles were spiralling along the field lines, bouncing from one hemisphere to the other, and drifting around the earth. These particles constituted belts around the earth, trapped by the earth's magnetic field, the Van Allen belts.

Gold [1959] named this new region, filled with trapped particles and dominated by the earth's magnetic field. He called it the *magnetosphere*. Aided by the dignity of a title, the new region around the earth became the focus of intense study. As with any new field of research there were lots of things to measure. The particle experimenters were quick to come up with new instruments to determine the identity of the trapped particles as well as their energies and densities in all parts of the region. The boundary of the magnetosphere was also important and the relatively few people with experience in both magnetism and rocket instrumentation were busy planning how to determine the sunward extent of the magnetosphere, as well as its shape and extent behind the earth. There were predictions that the boundary between the solar wind and the geomagnetic field would be broad and unstable [*Parker*, 1958]. Providing the welcome stimulation of a mild controversy, others including *Dessler* [1961] suggested that it might be stable.

Towards the end of 1958 another series of spacecraft was started as an Air Force project, the Pioneer lunar probes. Pioneer 1 carried a search-coil magnetometer through the outer geomagnetic field toward the moon [*Sonett et al.*, 1959; 1960]. The search coil detected only the portion of the earth's magnetic field perpendicular to the Pioneer spin axis. Since the spin axis was essentially perpendicular to the local geomagnetic meridian plane, the search coil managed to measure most of the geomagnetic field in the early part of the flight. Lunar orbit was not achieved but magnetic measurements were available in two intervals of the flight. In the first interval from 3.7 to 7 R_E the measured field, perpendicular to the rocket spin axis, agreed with a dipole model within 10%. In the second interval from 12.3 to 14.8 R_E, near local noon, the magnetic field measured was at first 20 to 35 nT, considerably higher than the model, and was very irregular. There were large changes in magnitude and in the phase angle of the field, measured in vehicle spin. At times phase changes were large enough to constitute reversals of **B**. The magnetic field was then pointing opposite to the expected direction of the earth's dipole field. Of course there was no knowledge of the component of **B** parallel to the spin axis. Near the end of the interval the field decreased to a very low level, about 5 nT. The decrease in magnitude of the perpendicular component at 13.6 R_E was interpreted as the "cutoff" [boundary] of the geomagnetic field. The irregular field region was interpreted as evidence of the expected instability in the outer geomagnetic field, caused perhaps by hydromagnetic waves or by interplanetary gas imbedded in the outer field. The outermost low field measurement, 5 nT, was taken as the first observation of the interplanetary magnetic field. A similar magnetometer was flown on Pioneer 5, which traversed the outer magnetosphere in the afternoon hours, local time. The observations in the outer magnetosphere appeared to confirm those of Pioneer 1 with an irregular field beyond 9.4 R_E and a drop in the perpendicular component magnitude, the "termination" or boundary of the geomagnetic field, near 14 R_E [*Coleman et al.*, 1960].

After I finished the analysis of results from the equatorial electrojet flights and the Greenland flights, my Ph.D. thesis based on these results was completed in 1959 [*Cahill*, 1959a; 1959b]. A few months after graduation from the University of Iowa, I was offered a faculty position at the University of New Hampshire.

AT THE UNIVERSITY OF NEW HAMPSHIRE

I commenced teaching Physics at the University of New Hampshire [UNH] in 1959 and also started an ionosphere research program, with plans for the use of sounding rockets. Research support at the University was somewhat primitive in those days. Fortunately there was an excellent research machine shop in the Physics building. We built up other needed services slowly as the research activity increased. There were also a number of bright and enthusiastic undergraduate students who wanted to work on research projects. There was no Ph.D. program in Physics then.

An opportunity to prepare a magnetometer for flight on Explorer 12 soon appeared. The orbit was to be elliptical with apogee at 10 to 15 R_E, initially near noon, local time. A 3-axis, fluxgate magnetometer with 1000 nT full scale seemed appropriate. This would give full vector measurements from 3 R_E out to apogee. A higher apogee would have been better, since the Pioneer results suggested geomagnetic field termination at 14 R_E. We put together a fluxgate magnetometer package, with major support from Eric Schonstedt of Schonstedt Engineering in Silver Spring, MD. Schonstedt built the 3-axis sensor and the supporting electronics. My students at UNH and I mounted the magnetometer in suitable enclosures and accomplished the operational and environmental testing prescribed by the Explorer 12 program office at NASA/Goddard Space Flight Center. Magnetic testing and calibration were done by Eric and me at the Fredericksburg Magnetic Observatory. After my last class on many Fridays, in the spring of 1961, I drove to Boston, caught the plane to Washington, DC, and proceeded to Silver Spring. Early the next morning Eric and I drove to Fredericksburg and spent the day in the large, 3-axis, Helmholtz coil system there. The coils could be used to cancel the earth's magnetic field and produce any desired test field, precisely known in magnitude and direction. We did a long and careful operational test program [to be sure the 3 sensors were precisely perpendicular, for example] and then calibration. Later I participated, with other Explorer 12 experimeters, in integration of all the experiments into the spacecraft structure. The program office was even kind enough to bring the spacecraft to Fredericksburg for a couple of days, where we did an operations and calibration check of the installed mag-netometer through the satellite systems and telemetry.

Meanwhile, another exploration of the outer magnetosphere was accomplished with the Explorer 10, launched in March, 1961 [*Heppner et al.*, 1963]. This time the launch was toward the rear of the earth on the evening side. Two fluxgate magnetometers were used, mounted so that with spacecraft spin they were able to give, once each spin, a measurement of the full magnetic field vector. The results showed, below the equator and at distances between 8 and 22 R_E, a geomagnetic field directed away from the earth and the sun, the first direct evidence of the tail of the magnetosphere. Beyond 22 R_E and out to 42 R_E, where the battery power supply ran out of energy, there were 6 occasions when the satellite was in a magnetic field lower in magnitude and substantially different in direction from the tailward magnetosphere field. Apparently the tail boundary was flapping back and forth across the spacecraft trajectory. When the spacecraft was in the outside field regime, an anti-sunward streaming plasma was observed [*Bonetti et al.*, 1963]. Explorer 10 was then outside the geomagnetic cavity and immersed in the solar wind. The magnetic field observed outside the magnetosphere was, however, fluctuating and irregular in magnitude and direction similar to that seen by Pioneer 1 and 5 and then considered a broad, unstable, boundary region.

Launch of Explorer 12 was planned from Cape Canaveral. In early August of 1961. I was there checking everything connected with the magnetometer. There was some tension because this was the first satellite launch for me and I felt great responsibility for proper operation of the magnetic field experiment. I had insisted, late in the launch preparations, on an opportunity to make a last-minute systems check of the magnetometer. The launch personnel escorted me to the top of the rocket and payload assembly, with the 3-axis sensor on its boom at the very tip. The test was crude but reassuring; I waved a small bar magnet near the sensors and received confirmation from the telemetry readout below that the 3 sensors were responding. Another responsibility was also on my mind; my wife and I were expecting the arrival of our third child at the time of the satellite launch.

Explorer 12 was launched on 16 August, 1961. I had left Cape Canaveral, after the payload was buttoned up but before launch, to help my family back in Durham, NH. An excited voice on the phone from Canaveral, on the evening of 16 August, described the magnetometer post-launch record: two sensors with sinusoidal variations, one relatively steady. It was a relief to respond that this was just right for sensors spinning in the earth's magnetic field, with one sensor along the spin axis. Early the next morning, our third son was born in Exeter, NH.

The most important result of the Explorer 12 magnetometer experiment was clear determination of the location of the subsolar boundary of the magnetosphere [*Cahill and Amazeen*, 1963]. In Figure 2 one pass of Explorer 12 shows very clearly a boundary located at 8.2 R_E. Other passes showed the boundary at positions from 8 to 12 R_E, near noon. In all of the boundary crossings observed with Explorer 12 the most reliable indication of boundary crossing was a large, sudden, change in field direction. In the crossing shown in Figure 2 the magnitude also has a sudden drop but in some crossings the magnitude stays about the same. However, there are always larger fluctuations in magnitude outside, as well as larger fluctuations in direction. The expected compression of the outer, subsolar, magnetic field is seen [Fig. 2] in the value of the measured magnitude [about double the dipole model] just prior to boundary penetration. Another important consideration was the nature of the boundary. Some had predicted that the boundary would be broad and unstable; the Pioneer observations appeared to confirm that. The Explorer 12 results, however, usually showed rapid passage of the spacecraft through a single, relatively-thin, boundary, with no multiple penetrations that would be expected if the boundary were oscillating rapidly or if it had a thick structure of multiple layers.

For a few years the Explorer 12 magnetic results were the only substantial body of data about the magnetic state of the outer magnetosphere and the region beyond, then called the transition region. A number of colleagues came to UNH to work with me and with the copious amounts of magnetic data. In 1962 a Ph.D. program in Physics was established at the University of New Hampshire; V. L. Patel was the first candidate to enter the program. He was the first colleague to work with me on Explorer 12 data. As one result, we published a more complete boundary study, including all 3 months of the Explorer 12 boundary measurements [*Cahill and Patel*, 1967]. For his thesis project, Vithal Patel focused on a search for evidence of magnetic pulsations, long observed on the earth's surface and thought to be generated and to propagate in the magnetosphere. He found magnetic pulsations of small amplitude, 6 to 8 nT, after smoothing relatively noisy data by taking 1 m sliding averages. The pulsations were perpendicular to the geomagnetic field lines, polarized clockwise and the period was about 2 to 3 m [*Patel and Cahill*, 1964]. The pulsation observations were made close to the magnetic equator and near 8 R_E; the boundary was crossed at 8.6 R_E shortly after the pulsations were observed. Similar pulsations were also observed at College, Alaska, at the same time and near the same longitude as the Explorer 12 pulsations.

Sudden impulses in the geomagnetic field at the earth's surface had been attributed to compressions of the magnetosphere due to changes in the solar wind. Atsuhiro Nishida had taken up study of sudden impulses and came to Durham to look for them in the Explorer 12 data. He found ample evidence of sudden impulses throughout the magnetosphere whenever they were observed on the ground [*Nishida and Cahill*, 1964]. The propagation time between the boundary and ground level was in good agreement with hydromagnetic wave theory.

Richard Kaufmann joined the faculty of the University of New Hampshire Physics Department in the early 1960s and became a research collaborator in studying the Explorer 12 data. He joined with Andrei Konradi of Goddard Space Flight Center [associated with the GSFC electron instrument] in studying evidence for boundary motion. With magnetic measurements alone it wasn't possible to tell whether the boundary was moving when a single crossing occurred. Of course, if there were multiple, in-and-out traversals associated with a boundary crossing, one might infer that the boundary was flapping in and out as the satellite passed through; there were some of these cases. *Konradi and Kaufmann* [1965] took a different tack. By considering the gyro-radii of the trapped particles they were

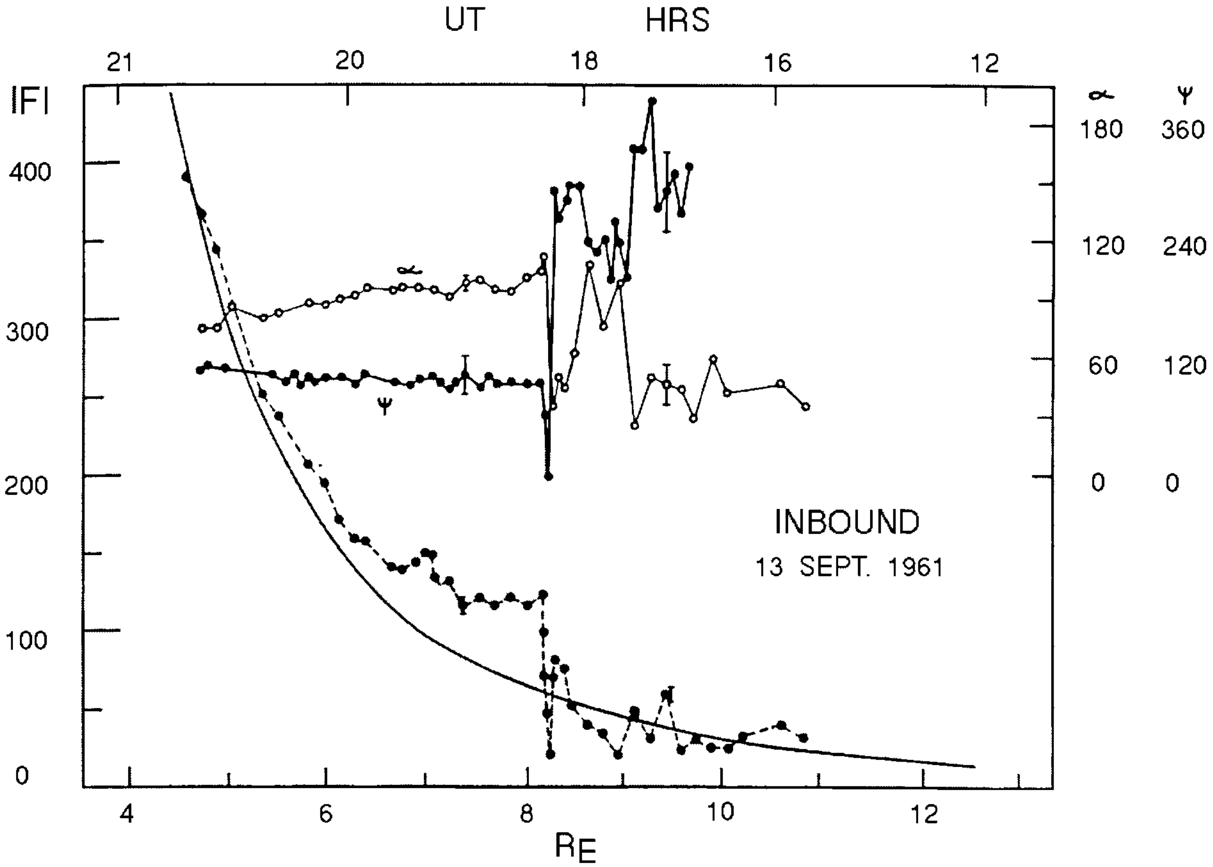


Figure 2. An inbound boundary crossing, by Explorer 12 on 13 September, 1961, is seen at 8.2 Re as sudden, large changes in the two magnetic field direction angles, alpha and psi, in spacecraft coordinates. The field magnitude scale, in nT, is shown at left; the angle scales, in degrees, at right. Radial distance from the center of the earth, in earth radii, R_E, is at bottom and the Universal Time in hours at top [*Cahill and Amazeen*, 1963].

able to show the boundary was moving in some crossings and to determine the speed of boundary motion.

The finding of an irregular magnetic field beyond 10 R_E and a very low field beyond 14 R_E in the Pioneer results as well as the observations just outside the boundary with Explorer 10 and with Explorer 12 suggested another interface in the solar wind/magnetosphere interaction. A bow shock was predicted by *Axford* [1962] and by *Kellogg* [1962], in analogy to the shock wave preceding a supersonic airplane, where in the present case the magnetosphere is moving relative to the solar wind at a velocity higher than the appropriate wave velocity [Alfven waves through a magnetized plasma]. Later *Kaufmann* [1967] was able to find several cases when Explorer 12 had penetrated the bow shock.

Dungey [1958] had proposed that sometimes the interplanetary magnetic field was able to connect with the geomagnetic field at the frontside boundary. The connection was facilitated when the interplanetary field was pointing south, opposite to the earth's field when they became adjacent at the boundary. When this happened the connected field was swept along with the solar wind into the magnetic tail. This so-called reconnection process also was thought to cause polar magnetic disturbances. A graduate student at Pennsylvania State University, Donald Fairfield, whose work was directed by Dungey, came to UNH to work as a research assistant for several months. A comparison study of transistion region magnetic field observations from the Explorer 12 data and polar ground-level magnetic observations revealed that when the field outside the magnetosphere had a strong, southward [negative] component there were polar magnetic disturbances [*Fairfield and Cahill*, 1966].

Soon after the initial discovery of the magnetosphere several groups were engaged in modelling the interaction between the earth's magnetic field and the solar wind

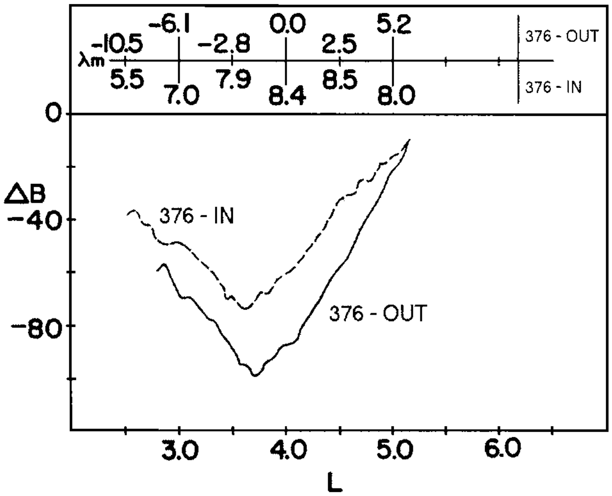


Figure 3. Profiles of the distortion in magnetic field magnitude by the ring current are shown for the outbound and inbound passes of Orbit 376 of Explorer 26. These passes occurred on 19 April, 1965, during the recovery phase of a magnetic storm that began on 17 April. During this storm the maximum depression of the Dst storm index was -150 nT on 18 April. The scale for field magnitude depression [observed magnitude minus dipole magnitude] is shown at left; the L value scale is at bottom [the L value gives the geocentric distance in R_E at the equatorial crossing of the field line]. The magnetic latitudes of the passes are shown in degrees at the top [*Cahill*, 1966].

[*Beard*, 1960; *Beard and Mead*, 1964; *Mead*, 1964]. Mead made several visits to UNH in 1966 and 1967 to compare his model predictions with Explorer 12 measurements of the field [*Mead and Cahill*, 1967]. He found that in the outer magnetosphere, 7 R_E out to the boundary, the observed field agreed quite well with his disturbed field model.

Another approach to study of the reconnection model was attempted by Bengt Sonnerup, then with Gold at Cornell. Bengt spent a couple of summers at UNH looking very closely in the detailed data [three vector measurements per second] to find evidence of reconnection. He found in several crossings during magnetic storms that there was evidence of a component of the field normal to the boundary surface; this suggested that the interplanetary and geomagnetic fields were indeed connected [*Sonnerup and Cahill*, 1967; 1968].

M. J. Laird of King's College in London spent the summer of 1967 with us at UNH. He was interested in the neutral sheet separating the lobes of the magnetotail behind the earth. Explorer 14, launched late in 1962, had a 3-axis magnetometer similar to that of Explorer 12 but an orbit with slightly higher apogee, 16 rather than 13 R_E. The orbit was initially moving into the tail in the early morning hours, local time, and provided some views of the near-earth tail. In a few orbits early in 1963 we were able to find passages through the southern lobe of the near-earth tail and observed an anti-sunward direction for the lobe field, confirming the earlier results of Explorer 10 [*Cahill*, 1964, 1965]. Laird was able to study several cases of neutral sheet passage near midnight local time and found, during a magnetic storm on February 10, 1963, that the tail lobes were connected across the neutral sheet by a substantial southward field component. He concluded that at this time there was field reconnection occurring in a region that was earthward of the satellite observations [*Laird*, 1969]. Before Laird's work *Ness* [1965] had completed a very thorough study of the magnetotail with data from the IMP 1 satellite; this spacecraft had a highly elliptical orbit with apogee at 31.4 R_E, very suitable for magnetotail studies.

One major puzzle of the magnetosphere, the origin of magnetic field depression in the main phase of a magnetic storm, had been addressed, but unconvincingly, by Chapman and Bartels in the 1940 volumes. The source was understood to be an electrical current circling the earth, flowing west, that caused the world-wide depression, sometimes as large as several hundred nT, but what was the location and the nature of the current?

The magnetometer experiment on Explorer 26 provided the location [*Cahll*, 1966]. Figure 3 shows a radial profile of the effect of the ring current on the magnitude of the earth's magnetic field. The distortion of the ring current has been made more apparent in this figure by subtracting a model, mostly dipole, field magnitude from the magnitude measured along the trajectory of this particular pass. The ring current was apparently due to trapped charged particles, protons drifting west and electrons drifting east. Even in quiet times there was a ring current, producing a much smaller, but measurable, effect than the one shown in the figure.

By this time the physics of the ring current was well understood, the trapped charged particles drifted and produced an electrical current that constituted the ring current. There were also other effects, including a local field decrease in the magnetosphere due to the diamagnetism of the new, storm-time charged particles that constitute the ring current. Actual measurement of the protons and electrons, that cause the current and the diamagnetism, and correlation of these particle measurements with the magnetic effects of the current were still lacking. The Explorer 26 magnetic field results were used to derive the ring current particle distributions that could have caused the magnetic perturbations [*Hoffman and Cahill*, 1968]. Finally, in the next decade, proton observations from 1 to 872 keV and magnetic observations were made on the same satellite [Explorer 45] and the observed proton distributions were demonstrated to produce the major part of the magnetic field

Konradi, A., and R. L. Kaufmann, Evidence for rapid motion of the outer boundary of the magnetosphere, *J. Geophys. Res., 70,* 1627, 1965.

Laird, M. J., Structure of the neutral sheet in the geomagnetic tail, *J. Geophys. Res., 74,* 133, 1969.

Ludwig, G. H., Cosmic ray insrumentation in the first U. S. earth satellite, *Rev. Sci. Instr., 30,* 223, 1959.

Mead, G. D., Deformation of the geomagnetic field by the solar wind, *J. Geophys. Res., 69,* 1181, 1964.

Mead, G. D., and L. J. Cahill, Jr., Explorer 12 measurements of the distortion of the geomagnetic field by the solar wind, *J. Geophys. Res., 72,* 2737, 1967.

McIlwain, C. E., Direct measurement of particles producing visible auroras, *J. Geophys. Res., 65,* 2727, 1960.

Patel, V. L., and L. J. Cahill, Jr., Evidence of hydromagnetic waves in the earth's magnetosphere and of their propagation to the earth's surface, *Phys. Rev. Letters, 12,* 213, 1964.

Ness, N. F., The earth's magnetic tail, *J. Geophys. Res., 70,* 2989, 1965.

Nishida, A., and L. J. Cahill, Jr., Sudden impulses in the magnetosphere observed by Explorer 12, *J. Geophys. Res., 69,* 2243, 1964.

Parker, E. N., Interaction of the solar wind with the geomagnetic field, *Phys. Fluids, 1,* 171, 1958.

Parker, E. N., Extension of the solar corona into interplanetary space, *J. Geophys. Res., 64,* 1675, 1959.

Sonett, C. P., D. L. Judge, and J. M. Kelso, Evidence concerning instabilities of the distant geomagnetic field: Pioneer 1, *J. Geophys. Res., 64,* 941, 1959.

Sonett, C. P., D. L. Judge, A. R. Sims, and J. M. Kelso, A radial rocket survey of the distant geomagnetic field, *J. Geophys. Res., 65,* 55, 1960.

Sonnerup, B. U. O., and L. J. Cahill, Jr., Magnetosphere structure and attitude from Explorer 12 observations, *J. Geophys. Res., 72,* 171, 1967.

Sonnerup, B. U. O., and L. J. Cahill, Jr., Explorer 12 observations of the magnetopause current layer, *J. Geophys. Res., 73,* 1757, 1968.

Van Allen, J. A., G. H. Ludwig, E. C. Ray, and C. E. McIlwain, Observation of high intensity radiation by Satellites 1958 Alpha and Gamma, *IGY Bull., Trans. Am. Geophys. Union,* 39, 767-769, 1958.

L. J. Cahill, Jr., School of Physics and Astronomy, University of Minnesota, Minneapolis, MN 55455.

Lightning Whistlers Reveal The Plasmapause, an Unexpected Boundary in Space

D. L. Carpenter

Space, Telecommunications and Radioscience Laboratory, Stanford University, Stanford, California

In the 1950s my efforts to obtain government work in International Affairs were frustrated by "McCarthyism." I then became a graduate student of electrical engineering at Stanford, where I soon obtained part time work as a data aide to Bob Helliwell. Bob's research group were investigating very low frequency (VLF) radio phenomena such as the impulsive signals from lightning that propagate in space plasmas as "whistlers." Before long I was the project expert on the dynamic spectra of whistlers as they appeared on frequency versus time records or sonagrams. As the whistler probing technique became more robust through the later 1950s, I began to study temporal variations in the magnetospheric equatorial electron density profiles inferred from the IGY and Byrd, Antarctica whistler data. Large, factor of >10, decreases in electron density were found and soon recognized as developing beyond a steep drop or "knee" in the equatorial density profile. In 1963, data from Eights Station, Antarctica, a true "window on space," revealed the knee to be a global phenomenon with a duskside bulge. In 1966 I introduced the terms "plasmapause" and "plasmasphere"; both were quickly adopted by the space science community. The reality of the knee effect was initially challenged and for several years debated by a group of GSFC particle experimenters. A method of tracking the cross-L motions of geomagnetic field aligned whistler paths was developed; it revealed the unsteady, substorm-associated penetration of the plasmasphere by high latitude convection electric fields. Since the early work based on whistlers and spacecraft, the experimental methods applied to the plasmasphere have not been matched to that region's huge size and complexity. For example, the phenomena of plasmasphere erosion and plasmapause formation have yet to be directly observed.

1. PROLOGUE

As a young person I decided that my future was not going to be in science or engineering. How could it be? I was reasonably good at high school algebra and geometry, but could never seem to understand what my engineer-father was saying when he tried to help me with homework problems. Even after a two year hitch in the Navy as an electronics technician, during which I enjoyed the challenge of trouble shooting radar systems, I proceeded to forget essentially all of the circuit fundamentals that I had been taught. I was convinced that my future lay elsewhere.

At Willamette University in Salem, Oregon I developed a passion for languages, studying French and Russian while pursuing undergraduate majors in political science and philosophy. Then in 1951, with an eye on the developing

Discovery of the Magnetosphere
History of Geophysics Volume 7

to inspect carefully a large number of records. One question concerned identifying the occurrence time of the lightning flash giving rise to a whistler. This time, or an estimate thereof, was needed in order for the travel time, or dispersive, properties of the whistler to be known and for information on plasma density along the magnetospheric propagation path to be extracted. One would like to find a way to recognize on the record a vertical line or impulse marking the arrival at the receiver of the causative "sferic," the energy that had propagated at essentially the speed of light in the Earth-ionosphere waveguide from the location of the whistler-producing lightning flash. Fortuitously, causative sferics turned out to be relatively easy to spot on some of the records of the Whistlers West stations, as illustrated by the example in Figure 1 (arrow in the lower margin). As I began to hear about whistler work elsewhere, I realized just how fortunate we were; in Europe, for example, causative sferics tended to be lost in the noise backgrounds produced by other, non-whistler-associated impulses from lightning.

In the course of further work I found that most lightning flashes excited multiple paths and thus multiple L values, a factor that was to become crucial to virtually all our later work on the equatorial density profile. Furthermore, I found that the f-t multipath fine structure of individual whistlers did not measurably change during a given short interval, say of 2 min.

2.2. *My First Knee Whistler*

Most of the whistler spectra that we studied during that early period seemed to be part of a relatively simple but self consistent picture. The first major anomaly that I can remember appeared during comparisons of simultaneous data from Seattle and Stanford. We had been making such comparisons so as to estimate the ground distance within which a whistler or VLF noise event emerging from the ionosphere could be detected. In most cases whistler components recorded at Stanford or Seattle were found to be independent of one another. The two stations were separated by about 1500 km, more than twice the typical ~500 km radius within which detection was eventually found to be common [*Helliwell and Carpenter*, 1961].

In data from January 13, 1958 I noted that whistler components from a lightning flash had arrived at Seattle and Stanford at approximately the same time, although the field line path to Seattle was much longer. This was my first view of what I came to call a "knee whistler," a type used later to support the concept of the knee phenomenon, or abrupt density decrease in the equatorial density profile. Other anomalies were seen; John Katsufrakis, a fellow student and later manager of Stanford's highly successful Antarctic field programs, found that the propagation times of whistlers received at Seattle following the great magnetic storm of February 11, 1958 were a factor of 3 to 4 shorter than usual.

2.3. *A Thesis Topic*

As 1959 approached I found myself becoming more attracted to the idea of the PhD and a career in research. The seeds of a thesis topic were planted as the result of a visit to our lab in August, 1959 by Alex Dessler, who was working at the time for Lockheed in Palo Alto. Alex suggested that we look at the effects on whistlers of a large magnetic storm that had begun on August 16th. Responding to this suggestion, I noted a pronounced decrease, by a factor of 3, in whistler dispersion in the aftermath of the storm, implying a decrease in magnetospheric electron density by about an order of magnitude! Other storm periods were then investigated and evidence of density decreases by varying amounts was obtained. I was impressed by these indications of magnetospheric unsteadiness.

Figure 2 illustrates the magnetic storm effect; the upper panel shows two closely spaced multicomponent whistler recorded at Seattle during a magnetically quiet period in June, 1959. Arrows in the lower margin mark the two causative sferics. The lower panel shows a whistler recorded at Seattle on August 18, 1959, only a day after the geomagnetic Kp index had reached 8+ for the second time within 24 hours. The propagation times of its components were roughly a factor of 3 shorter than those of the components in the upper event that propagated at similar L values, i.e. those that extended to or above about 12 kHz.

A number of research developments between 1956 and 1960 were of critical importance for my efforts to extract information from whistlers. In 1956, [*Helliwell et al.*] had shown that, in principle, every whistler component had a frequency of minimum time delay or "nose," as illustrated by the elegant example in the upper panel of Figure 2. This effect was attributed to a singularity in the expression for the whistler-mode refractive index at the electron gyrofrequency. Thus each whistler component carries with it information on the path it has followed through the geomagnetic field. Conveniently, the nose frequency turns out to be roughly proportional to the minimum electron gyrofrequency along the whistler path.

In those early days one might have expected the randomness in lightning source location to introduce a comparable randomness in the dispersive properties of the whistlers recorded within a short time interval, say of several minutes. In fact, it was found that successive whistlers propagate as if on a fixed set of discrete, field aligned paths [e.g. *Smith*, 1961]. Thus the nose frequencies of the multiple components of a whistler can in principle be used to identify a discrete set of magnetic shells or L values along which they have propagated.

Another important step was the development of a curve fitting technique for estimation of the path L values of whistlers such as the one illustrated in Figure 2, lower panel, for which the nose frequencies were not detectable on the records. At Stanford and at other mid- to low-latitude stations, where many of the early whistler recordings were

examples of the effect were not easy to find; the knee had frequently been at L<3 during the IGY period of very high solar activity, and the associated datasets were mostly from the mid-latitude Northern hemisphere, where whistler activity after magnetic storms was often sporadic and usually required application of the extrapolation method. However, in the fall of 1962 I decided that enough material had been accumulated and submitted an article on the knee to the *Journal of Geophysical Research*; it was published in early 1963 [*Carpenter*, 1963].

3.2. *Gringauz and the Lunik Rocket Data*

In 1962 I heard about evidence of a knee-like drop in thermal ions obtained by *Gringauz et al.* [1960] from ion traps on the LUNIK 1 and 2 rockets. Being quite unfamiliar with particle instruments, I discussed the LUNIK results with Bob Mlodnosky, who was then managing an ambitious Stanford experimental program in the upcoming OGO satellite series. Mlodnosky found the LUNIK results credible from the point of view of the detectors employed. In turn, I found them reasonable in terms of what I then knew about the density profile.

Having once been a student at the Russian Institute of Columbia University and thus a potential "cold warrior," I found it ironic that my new career had at its outset intersected the career of a prominent Russian researcher. This intersection became all the more real when Bob Helliwell, then chairman of Commission IV of URSI, invited me to review my work on the knee effect at the 1963 URSI Assembly in Tokyo. I prepared a plot, here reproduced as Figure 4 [*Carpenter*, 1965], comparing the LUNIK results with an idealized whistler-based equatorial profile with a knee. The comparison was complicated by the fact that the LUNIK trajectory had begun at ~60° invariant latitude (upper row of numbers in Figure 4), moving with altitude first toward lower and then toward higher L values. At low altitudes the LUNIK numbers were much smaller than the corresponding whistler values, probably because the rocket had initially been poleward of the knee or plasmapause, as the geomagnetic-field aligned boundary later came to be called. A density falloff was observed as the rocket, having apparently penetrated the plasmasphere, continued outward toward its ultimate collision with the moon. I later came to realize that had the magnetic activity at the time been more intense, the plasmasphere might have been missed altogether! In any case, I was pleased to meet K. Gringauz in Tokyo and to discuss our common interests. Around this time I heard that there had been strong disagreements among Soviet scientists about the LUNIK results, and was pleased that our whistler work at Stanford had played a role in supporting Gringauz's position.

The excitement of the Tokyo meeting was almost more than I could stand. Our Japanese sojourn began with a small lunch hosted by T. Obayashi, who is revered for his pioneering role in Japanese space science. At the meeting I not only had a chance to speak, but also got to see, hear, and occasionally meet major players in ionospheric radio research. In the area of VLF I met R. Rivault and Y. Corcuff of the Poitiers group, R. Gendrin from Paris, M. Morgan of the Whistlers East group, I. Kimura of the University of Tokyo, and G. Mck. Allcock, a whistler researcher from New Zealand. I was able to discuss the pathology of whistler data with Yvonne Corcuff, and, with her indulgence, to converse at least part of the time in French. I was introduced to a key member of the URSI secretariat, Jela Stefanovitch, and was delighted to find myself speaking her native Serbian, a language I had learned some 10 years before. Upon arriving home after three weeks of this kind of stimulation, I slept almost continuously for 48 hours!

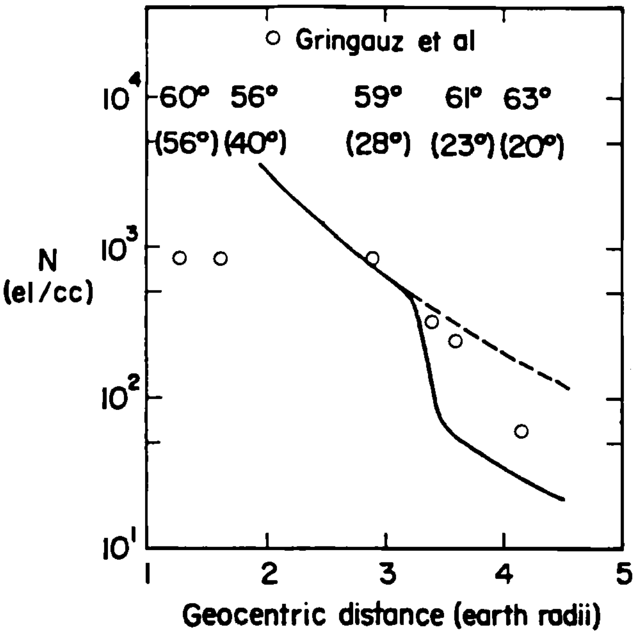


Figure 4. Plot of plasma density versus geocentric distance comparing an idealized equatorial density profile obtained from whistlers with measurements of ion density reported by *Gringauz et al.* [1960] from the Lunik 1 rocket (open circles), launched on January 19, 1959. The upper row of numbers shows the invariant latitude of the rocket, the lower row the geomagnetic latitude. From *Carpenter* [1965].

3.3. *The Plasmapause as a Global Phenomenon*

I thought my 1963 paper was convincing on the subject of the knee as a permanent feature of the magnetosphere and on the negative correlation of its radius with magnetic activity. There was clearly much more to be learned, however, and the existing whistler stations and data were far from optimum for probing the region near L=4 where the knee was most likely to appear. Fortunately, much better data on the knee phenomenon were being acquired at that very time. In the austral summer of 1961-62, Neil Brice, a graduate student at Stanford, had carried a portable receiver on a traverse into an area of Antarctica called Sky-Hi, at L~4 near the Ellsworth mountains. His findings were surprising, even for those of us at Stanford who had anticipated good results. In a season when whistler activity originating in Northern-hemisphere lightning would be expected to be minimal, Brice had recorded elegant one-hop whistlers that covered a wide range of L values, from ~3 to 6.

In 1963, Eights Station, the first U. S. Antarctic winter-over activity whose site was chosen purely on scientific grounds, began year-round operations. Located in the vicinity of Sky-Hi and in the region magnetically conjugate to the eastern U.S. and Canada, and hence to one of the highest concentrations of middle latitude lightning activity on Earth, it produced whistler data that to this day defy description in terms of the wealth of information provided. Data from the Whistlers-West stations and Byrd had revealed the knee in the density profile, but in a true sense it was Eights that allowed us to see the knee in three dimensions, in its global manifestation as the plasmapause, bounding a region that would be called the plasmasphere. Eights was a true "window upon space."

1963 was a year of low solar activity. Conditions for whistler propagation were ideal. Magnetic storms that occurred tended to be weak, such that storm-related interruptions in whistler activity, well known from the great events of the IGY period, were minimal. Post-storm recovery periods, known to be favored intervals for whistler activity, were long in duration. When the 1963 data arrived in the Spring of 1964, I soon realized that it was going to be possible to track the position of the knee for many hours each day and for days in succession.

Case studies promptly showed, essentially in real time, the kind of inward shift in the knee location during magnetic storms that only crude statistics had previously been able to suggest. After the initial large inward displacement of the knee, a distinctive diurnal variation in its location appeared and was observed for several days in succession as relatively steady substorm activity continued into the recovery period.

This diurnal variation became a crucial element in the interpretations of the plasmapause phenomenon that were soon to be advanced by *Nishida* [1966], *Brice* [1967], and *Dungey* [1967]. The nightside was characterized by inward displacement from a relatively large radius; on the dayside the radius increased slightly, on average, to a minor peak at noon, and then in the late afternoon shifted abruptly outward by order of one Earth radius to form a bulge-like extension. From these recovery-phase data it was possible to distill the crude model of equatorial plasma density shown in Figure 5.

1964 contained enough excitement for a lifetime. In November I was privileged to watch the launch of OGO 1 at Cape Kennedy with a Stanford receiver on board, and while immersed in the remarkable 1963 whistler data from Eights I came across a powerful new whistler method for detecting plasma drift motions in the magnetosphere.

4. TRACKING CROSS-L MOTIONS OF WHISTLER PATHS

In studying a set of Eights whistlers with long trains of echoes, for which the one hop propagation time could be very accurately measured, I noticed a steady decrease in the echo period over several nighttime hours. Having earlier read the work of *Axford and Hines* [1961] on the subject of magnetospheric convection, I concluded that the whistler path, in the form of a tube of slightly enhanced ionization, was participating in the bulk $\mathbf{E}\mathbf{x}\mathbf{B}/B^2$ drift motions of the surrounding plasma, and that what I had seen represented an inward component of drift in the cross-L direction.

To confirm the existence of this effect, I searched the 3-h continuous recordings that observer Mike Trimpi had made at Eights and found the example shown in Figure 6. A component with a well defined nose frequency near 4 kHz, corresponding to propagation at L~4.5, appeared in a succession of whistlers recorded between 0500 and 0820 UT (~0000 MLT to 0320 MLT). From top to bottom at the left, the entire whistler, including several lower L components, is shown at three local times, 0000 MLT, 0150 MLT, and 0310 MLT. Extending to the right along the top panel are spectrogram segments of the component of interest, recorded at intervals over the ~3-h period. Arrows at the left and right ends of the panel show the beginning and ending levels of the component's nose frequency, while at the lower right is a panel showing in f-t space a tracing of the beginning and ending configuration of the component. The overall increase in nose frequency implied an inward displacement of the path by about 0.3 R_E, while the reduction in propagation time was consistent with what would have been expected from an essentially electron-content preserving motion of the flux tube. The decrease in path L value, occurring as it did in the outer plasmasphere, was consistent with what I had been finding about the nighttime inward displacements of the plasmapause. I imagined that there could be no better example of the concept of equipotential field lines: if significant variations in potential had existed along the duct between conjugate topside ionospheres, the duct could not have retained its field-aligned form and hence its properties as a waveguide.

5. ANGERAMI'S WORK ON THE EQUATORIAL PROFILE

In 1964 I was joined in studies of the knee phenomenon by a graduate student from Brazil, Jacyntho Angerami. While I pursued issues involving plasmapause location, Jacyntho worked on the equatorial density profile as well as the distribution of electrons along the geomagnetic field

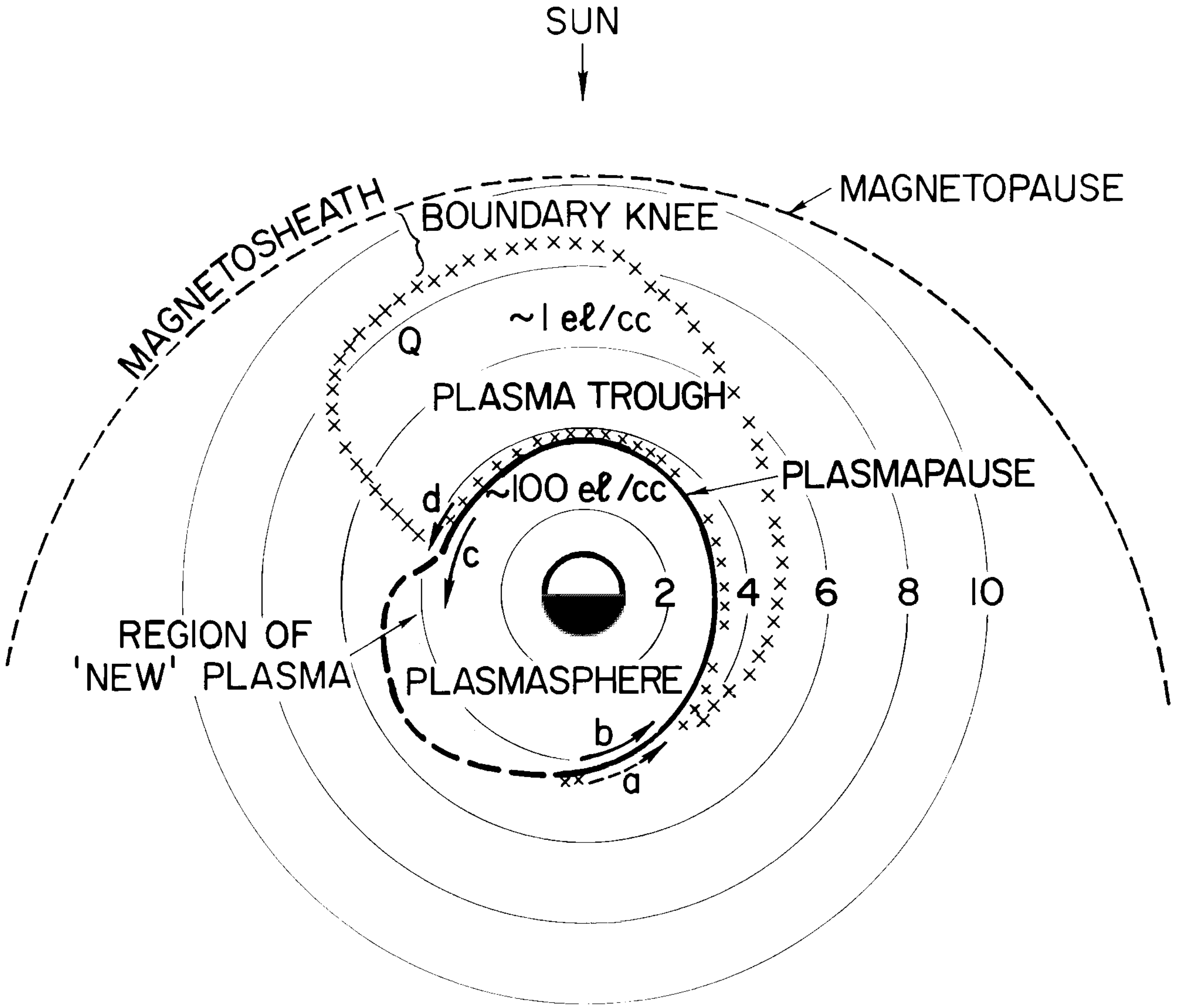


Figure 5. Model of equatorial plasma density for periods of steady, moderate geomagnetic agitation based upon 1963 data from Eights, Antarctica. The part of the trough outlined with x's indicates the region of most frequently observed ducted whistler activity exterior to the plasmapause. Dashes along the plasmasphere boundary near dusk show where the outer limits of the plasmasphere tended to be poorly defined for averaging purposes. The "boundary knee" refers to the expectation of a density increase from trough levels to magnetosheath levels. The bulge region did not appear to develop as the result of outward flow from interior regions, but instead appeared to have some other origin. Adapted from *Carpenter* [1966].

lines. Displaying remarkable skill in both data handling and interpretation, he showed that a diffusive equilibrium type model was consistent with whistler and topside sounding data inside the knee, while a model varying much more rapidly with radial distance was required in the region beyond [*Angerami*, 1966]. Through his work on the electron content in tubes of ionization above the regular ionosphere, he clarified a point that had been left uncertain in my 1963 paper, namely that the density depletion beyond the knee could be assumed to extend essentially all along the field lines between the conjugate hemispheres.

6. CONVINCING THE COMMUNITY

The work on the 1963 Eights data that began in 1964 culminated in a presentation at an AGU-URSI symposium on solar-terrestrial relations in 1965. If any presentation I have helped to produce can be called a *tour de force*, that

geomagnetic conditions. At first I gathered crude statistics on the apparent direction and velocity of cross-L flow vs time. The general trends emerged quickly, and not surprisingly, seemed to agree with the known temporal changes in plasmasphere radius. It was actually possible to summarize these trends at the 1966 symposium in Belgrade. Incidentally, my return to Belgrade after 13 years was a joyful experience. I was speaking Serbian and reunited with a dear friend from my earlier visit, now a local newspaperman. I was duly photographed and then thoroughly and humorously misquoted in a popular press article.

Having read in the work of *Akasofu* [1964] about what were then called auroral or polar substorms, I also sought detailed evidence of drift effects in Eights whistlers during periods of enhanced magnetic activity at the higher latitude Byrd Station and at its conjugate Great Whale River, Canada. A remarkable case was found, in which it was possible to track the motions of several whistler paths distributed over a range of ~0.6L in the outer plasmasphere [*Carpenter and Stone*, 1967]. The several path L values, carefully measured by a graduate student Keppler Stone, changed with time as if driven by a moderately large scale electric field. Figure 8 shows how the equatorial radii of the whistler paths changed with time in relation to riometer absorption and geomagnetic field changes at Byrd and Great Whale.

We were excited about having what we believed to be the first direct evidence of convection-associated inward plasma work at Cambridge, Ratcliffe had been skeptical of the

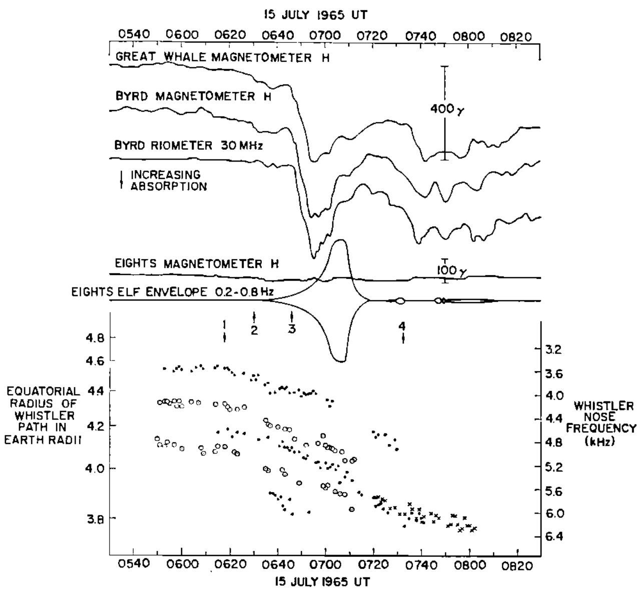


Figure 8. Data showing cross-L inward drifts of Eights whistler paths during a magnetospheric substorm. Above are magnetometer, riometer, and ULF signatures from the higher-latitude Byrd station and Great Whale, Byrd's geomagnetic conjugate. From *Carpenter and Stone* [1967].

validity of Storey's results. A measure of his eventual acceptance of the whistler probing technique was his decision to feature it on the cover of his popular book.

10. GEOPHYSICAL CONSEQUENCES

As awareness of the plasmapause phenomenon increased, so did the numberof questions concerning its relations to other geophysical phenomena such as ionospheric troughs, drifts in the nightside magnetosphere during a substorm. The inferred peak velocity of cross-L flow was ~0.4 R_E/h, nearly 25% of the speed of corotation at L=4.

10. SPREADING THE WORD ABOUT THE WHISTLER METHOD

During the 1960s I experienced satisfaction in being able tocontribute to the growing body of knowledge about the magnetosphere, but also a sense of isolation due to the arcane nature of the experimental technique we were using. At a typical meeting, most of the attendees would be familiar with the capabilities of devices such as particle detectors, all-sky cameras, incoherent scatter radars, and topside and bottomside radio sounders. Almost no one, however, would know much about whistler experiments, so I had almost no one to talk to about data. The other whistler groups around the world were few in number and were not yet engaged in whistler collection and analysis programs ona scale comparable to ours.

In the 1970s word did begin to spread. A diagram on the dustcover of a popular introductory monograph on the ionosphere and magnetosphere, published in 1972 by *J. Ratcliffe*, became a bright spot in the progress of community awareness of the whistler technique. The diagram showed a set of geomagnetic field aligned whistler paths, distributed on both sides of a plasmapause. In the early 1950s, as supervisor of *Storey*'s [1953] pioneering whistler SAR arcs, magnetic pulsations, substorm particle injections, etc. It was humbling to encounter the difficulty inherent in going beyond one's narrow area of expertise. In 1966 I had naively thought that in a short time we in whistler work would be telling ionospheric workers all they ever needed to know about the then unknown conditions in the topside ionosphere ionosphere that are imposed by the structure and motions of the overlying dense plasmasphere. Seven years were to pass before Chung Park and I would summon the nerve to write a review paper entitled "On what ionospheric workers should know about the plasmapause-plasmasphere" [*Carpenter and Park*, 1973]. Even then it was only a beginning.

11. EPILOGUE

Since its discovery over 30 years ago the plasmasphere has been largely hidden from view behind the shield of its

Smith, R. L., and D. L. Carpenter, Extension of nose whistler analysis, *J. Geophys. Res.*, *66*, 2582, 1961.
Storey, L. R. O., An investigation of whistling atmospherics, *Phil. Trans. Royal Soc. (London) A*, *246*, 1953.
Taylor, H. A., Jr., H. C. Brinton, and C. R. Smith, Positive ion composition in the magnetosphere obtained from OGO-A satellite, *J. Geophys. Res.*, *70*, 5769, 1965.

D. L. Carpenter, Space, Telecommunications, and Radioscience Laboratory, Stanford University, Stanford, CA 94305.

REPRINTED FROM JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 99, NO. A10, PAGES 19,189–19,197, OCTOBER 1, 1994.

Memories, Maxims, and Motives

J. W. Dungey

Blackett Laboratory, Imperial College of Science and Technology, London, United Kingdom

This account displays continuing development flowing from my Ph.D. project. MHD was sufficient for the discovery of reconnection and observations on spacecraft provided supporting evidence. If the rate of reconnection is to be predicted, more sophisticated plasma theory is needed. Extensive analysis of wave-particle interactions is described. They were thought to be important in the neutral sheet, but another mechanism, electron viscosity, is described and no conclusion is reached.

1. INTRODUCTION

It is a big honor to contribute to the celebration of the senility of the American Geophysical Union (AGU). In my ignorance I hope to be able to read the history of AGU and the *Journal of Geophysical Research* (JGR). How has JGR reached volume 99? I dimly remember references to a journal called *Terrestrial Magnetism and Atmospheric Electricity* and wonder if this was involved. In my earliest research years I also remember a loneliness, because there was so little literature on cosmic magnetism; this was followed by an explosion to excessive literature in the 1960s, but I recognize my luck in working in that golden age. Recently, the loneliness has returned, because my interest is extremely narrow, but I enjoy not needing to read and would only wish for a small club in neutral sheets.

I remember being immediately struck by the complexity of plasma physics and note that "complexity" seems to be a current trendy buzzword. Without wishing to upset any applecart, I feel bound to say that complexity is not new. "Seeing the wood for the trees" is an old cliche, though which meaning of "wood" is intended is obscure. "Waterfall physics" was a common metaphor long ago, while turbulence is almost synonymous with complexity and supplied the useful distinction between large and small eddies. I soon acquired some pictures of aurorae (paintings as there were no photographs), which confirmed the complexity described in dynamic accounts. Later in this paper I will mention complexity in velocity distributions and in the dynamic spectra of waves. The last paper I published in JGR [*Dungey*, 1985] was concerned with fine structure in the velocity distribution and its repercussions for simulations. My experience has been that complexity appears daunting, but sometimes a route through the jungle is found and with hindsight it is seen that the apparent difficulty was deceptive. Consequently, theorists should be willing to gamble, with crude approximations say, since very little in the way of resources is at stake. Ethically, however, one should not stake a young person's career on too wild a gamble.

Motives are required for any proposal to be approved by a committee, but more general motives are less often discussed, rather puzzling and perhaps somewhat subjective. The generation of physicists previous to mine gave strong encouragement to plasma physics and the space race gave financial encouragement. Subsequently, the subject drifted gradually into middle age, and the community became more set in its ways. I drifted into middle age about the same time, but my value judgements diverged from the consensus: now my Holy Grail is someone else's red herring.

Discovery of the Magnetosphere
History of Geophysics Volume 7

2. MAGNETOHYDRODYNAMICS

By the 1940s, Alfvén had laid the foundations of both MHD and the particle approach to plasma theory. Alfvén waves had his name on, and he also discovered the magnetic moment invariant of the guiding-center approximation. MHD is obviously hydrodynamics with extras and familiarity with conventional fluid mechanics is a prerequisite. Some mathematicians tackled MHD, but it was also used in a handwaving way in semiquantitative and often speculative applications. Handwaving is facilitated by certain attractive features, which are sometimes criticized and so will now be defended.

The most attractive feature of MHD is the familiar "frozen field" theorem. To any mathematician struggling with partial differential equations, this must seem like finding a priceless jewel, and it must be even more priceless to those who prefer pictures to equations.

A pretty and simple example is the Parker spiral model of the interplanetary magnetic field, which furthermore is useful in interpreting observations. However, the perfect conductivity equation is an approximation, and the result may need to be modified as in the case of reconnection, a major topic to be discussed later. It is possible to define a velocity for magnetic field lines even when the electric field is completely unrestricted, but this will not be pursued here. All my publications are short except one, a book [*Dungey*, 1958], in which a recipe for field line motion is presented, though I might now need to add a reply to rather hairsplitting criticism involving noise in the fields. Another stumbling block, the induced electric field, is dealt with in my book in a section on causal relationships; this is important and will be recapitulated in outline. Causal relationships are discussed by considering numerical simulations, a very young industry at the time, leading to the conclusion that causal relationships depend on the approximations made and so are not really very fundamental. In MHD it is found that curl **E** should be regarded as causing $\partial\mathbf{B}/\partial t$, and I have been surprised that several intelligent people cannot swallow this, presumably because they were brainwashed by unusually charismatic teachers in their youth. Some, however, will admit that it is silly to add an induced field when curl **E** already matches $\partial\mathbf{B}/\partial t$. The other feature of MHD that is conducive to handwaving is the magnetic stress, which can be expressed as the sum of an isotropic pressure and a tension parallel to B. Consequently, magnetic field lines can be viewed as elastic strings, whose tension "tries" to shorten them, and this is useful in understanding MHD behavior, most obviously in the primitive laboratory experiment known as the "pinch."

Waves also involve the magnetic stress and knowledge of waves is essential to MHD theory. Probably the best observed waves are associated with geomagnetic pulsations, which are observed on the ground and are notably narrowband. They were discovered long before the ionosphere but not explained before my investigation in 1954. Since the Earth is a conductor and the free space wavelength is many Earth radii, I believed that the ionosphere would have little effect and investigated MHD waves in the magnetosphere. It was of course tempting to relate the pulsation period to the wave travel time along a field line, but this depends on the mass density of the plasma. Art Waynick drew my attention to Owen Storey's whistler story, which was quite recent and not understood. Partly to fit the pulsation periods, I favored the idea that Storey's plasma was atomic hydrogen rather than oxygen and, with Marcel Nicolet's advice, came up with the charge exchange mechanism to switch ions from oxygen to hydrogen. This aspect of the picture was established early in the space era and became so familiar that later discoveries of oxygen ions caused excitement.

My neglect of the ionosphere is an example of crudity yielding progress. It bothered me for years, but I was not aware that anyone else worried. Eventually, I suggested the problem to Jeff Hughes as part of his Ph.D.; he smashed the problem by computing and, with David Southwood's help, discovered a simple and elegant approximation: "the Hughes rotation." According to Hughes and Southwood the magnetic perturbation just above the ionosphere must be rotated through a right angle about the vertical to obtain the perturbation on the ground, which exchanges N-S and E-W polarizations while leaving circular polarization unchanged. This cleared some puzzles and was timely because of the observational programs starting at that time.

I speculated on the mechanisms exciting pulsations in the MHD picture. Given the existence of a magnetopause, as in the much earlier model of Chapman and Ferraro, the Kelvin-Helmholtz instability could be anticipated, and later the MHD formulation was very elegantly laid out and analyzed by David Southwood. Unfortunately, observational testing tends to be frustrated by the complexity of the magnetopause. Another enjoyable handwaving speculation suggests that Pi 2 are the pizzicato form of pulsations, probably twanged by impulsive reconnection. Observations may be able to test this hypothesis, but it is currently controversial. One other mechanism for exciting pulsations will be discussed in the section on wave-particle interactions. It is an example of MHD with an additional collision-free effect, which is common in the subject. MHD underpins much of plasma theory and is itself underpinned by conservation laws, which of course should usually be obeyed.

3. RECONNECTION

Reconnection has been the thread running through my career, particularly at both ends. The idea started with Giovanelli's suggestion of the importance of neutral points, based on his observations of solar flares. Fred Hoyle pursued the idea and gave me the task of developing the theory and applying it to the aurora as a Ph.D. project. In those days there was a financial penalty for submitting a Ph.D. thesis more than three years after the beginning of one's grant and halfway through my third year I had the horrifying thought that Lenz' law would prevent the buildup of electric current, which Hoyle had asserted would be large at a neutral point. As I bicycled nervously to tell Hoyle this worry, it occurred to me that he might question whether Lenz' law applies in a fluid, and I made a U-turn and went home for a rethink. I made both algebraic and pictorial tests and was surprised to find that the induction effects actually caused the current density to increase. Then, when I had obtained the same results three days running, I happily told Hoyle and wrote a heading "dissertation" on a clean sheet or paper. Later some people have told me that Lenz' law does not say what I thought it did, but that is what I thought, believe it or not.

The buildup of current density is not the whole story of reconnection, though qualitatively it is crucial. It must be reemphasized that reconnection is not a consequence of the frozen field approximation, but of its breakdown, which is caused by high current density. The picture may need to be recapitulated, and the simplest case is of course clearest. Consider a two dimensional magnetic field, so that the field lines can be drawn on the page (Figure 1) and let there be just one neutral point N, with "limiting" field lines, two going into N and two coming out. Let the electric field be everywhere perpendicular to the page and define a velocity for the field lines as $\mathbf{W}$, defined by $\mathbf{E} + \mathrm{W} \times \mathrm{B} = 0$, so that the definition does not involve the plasma directly, but only through the dependence of E on the plasma. Now suppose the limiting field lines are moving in the directions indicated in Figure 1. According to the frozen field approximation they remain field lines, and they still go in to or out from N, so that they remain limiting field lines. This is the kind of motion that tends to increase the current density at N, but it is not reconnection.

Reconnection occurs when there is resistivity, anomalous resistivity, or something equivalent, as has now been

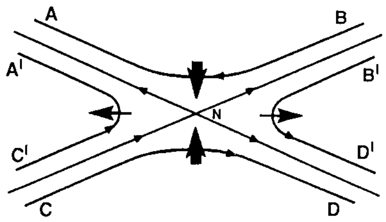


Figure 1. The pictorial explanation of reconnection (see text).

demonstrated by numerous simulations. The problem in the collision-free neutral sheet will be discussed later. Now, if $\mathbf{E}$ is finite at N, $\mathbf{W}$ has a singularity there, and it helps to make use of another approach, which may be described as using a one-component vector potential A, the vector being perpendicular to the page. This A behaves as a field line label, and it is consistent to say that field lines retain their A values, while moving with velocity $\mathbf{W}$; the flux between two field lines is the difference between their A values. Field lines flow toward N from two sides and away from N on the other two. At some instant a pair of inflowing field lines actually become the limiting lines (which all have the same A value) and immediately after that they form an outflowing pair. They are found to have "swapped partners" (in Figure 1, A partners B, but after reconnection A' partners C'). This process is described as cutting the field lines as they pass through N and "reconnecting" them. The rate at which flux is reconnected can be obtained from the vector potential A and the definition of $\mathbf{W}$ and, as it should, this yields Faraday's law.

My first postdoctoral post was in Sydney, and I much enjoyed interacting with Ron Giovanelli, the originator of the neutral point mechanism. He did a conscientious job as devil's advocate and submitted my paper on reconnection to *Monthly Notices*. It was rejected mainly on the grounds of my neglect of plasma pressure, but, since this has no connection with Lenz' law, my discovery was not demolished. I reacted with a study of MHD equilibrium in a class of two-dimensional fields and found that equilibrium required infinite current density at the neutral point, so pressure was not a worry. With this section added and on the helpful advice of Sydney Chapman, I submitted the paper to *Philosophical Magazine*, and it was accepted. Of course, *Philosophical Magazine* was not widely read, but my paper was occasionally cited, notably in the famous original paper on the tearing mode [*Furth et al.*, 1963].

There seems to be a psychological problem concerning the pictorial explanation of reconnection. I felt that the pictorial explanation was easy to assimilate and was therefore the best sales tactic, but over the years I have been told that many people find it repulsive. This seems to be an emotional reaction, and I must have been naive to expect other scientists to be as objective as myself, but it is not for me to discuss any aesthetic aspect. Some further criticism, following up the question of pressure, came in the "Parker- Sweet" approach. They started with a neutral sheet rather than a neutral line, and I still think it was a red herring. Their difficulty was, however, cleared up by *Petschek* [1964], who included Alfvén waves propagating along the outflowing field lines, which I accepted when he told me. His field lines have corners instead of sharp bends, but, more importantly, he specified the divergence of the outflow, whereas my flow was at the handwaving level. More importantly still, his story sold well. I have not kept up with MHD and do not know how simulations have compared with his model. I am not aware of any observational confirmation but would hardly expect any because of the limitations of MHD.

My quest for relevant observations followed from Hoyle's proposal of the aurora as an application. Once the reconnection picture was sketched, the open model of the magnetosphere followed easily, provided the interplanetary magnetic field was chosen to be southward, and the open model is in my Ph.D. thesis with diagrams of the noon-midnight meridian plane. My thesis was short of material, but I did not publish the open model until 1961, and it is barely mentioned in my book. I must have been influenced by the poor reception of reconnection but was frustrated by my inability to relate the open model to observations, not only because of the exhortations of senior physicists about confronting theory and experiment. I was lonely in my subject and fortunate to be given fellowships in ionospheric groups, so was quite well up on the ionosphere and felt that it should provide a test of the open model. It was like failing to recognize the significance of a clue in a detective story, and with this analogy, another suspect must now be introduced. Possibly being greedy, I had the idea of "viscous interaction" well before Axford and Hines. The only published evidence is a sentence in my book about "spray" blown off the magnetopause. The idea in a vague form occupied my mind quite frequently; however, I have a clear memory of discussing it with Jo Pawsey in Sydney; I left Sydney in 1953, and I doubt if I ever saw Pawsey again, as he died young. For the rest of the 1950s I wanted an observational clue to choose between the open model and viscous interaction. Though I expected both mechanisms to occur, it could well be that one would dominate. I know that my decision came in a flash while I was sitting at a sidewalk cafe in Montparnasse, because, when I got home, I immediately wrote the paper. What I saw was the open model in three dimensions with the boundary between closed and open field lines in the auroral zones and particularly the relation of the electric field in the magnetosphere to the currents in the ionosphere. Viscous interaction actually produces similar electric fields and currents, but I was impressed by the sharpness of the boundary between eastward and westward currents. This convinced me that it was a topological boundary. It may be noted that I had been familiar with this observational data for years. It was not from the International Geophysical Year (IGY) and satellites were not involved, although a few years later, they would provide strong support for the openness of the magnetosphere.

Since 1960 I have been particularly interested in data from magnetometers on spacecraft and can even remember the analogue data from Chuck Sonett's rotating coil on Pioneer 1. On Explorer XII, launched in 1961, Larry Cahill's digitization was designed for an exploratory approach, with the consequence that complex data reduction was required to obtain a B vector (assuming this was steady enough over many spins, which turned out to be usually acceptable). This instrument discovered the magnetopause and obtained valuable data just outside the magnetopause, but the satellite died when its apogee was near dawn.

For IMP I, launched in 1963, with the benefit of previous exploratory results, Norman Ness had his software ready before launch and was able to announce, almost in real time, his discoveries of the bow shock and the tail neutral sheet. Around this time I managed to persuade Cahill to allow Don Fairfield to make a correlation study between his Explorer XII data and data from magnetic observations on the ground. During this period of the early space race, the United States had a shortage of university faculty in some relevant areas and Art Waynick at Penn State obtained a U.S. Air Force contract to employ alien consultants to supervise postgraduates mainly by mail. Fairfield was the first of my students under this scheme and for his master of science was well grounded and ground into ground magnetic data from the IGY. Knowing the time intervals when Cahill's data existed, just outside the magnetopause, Fairfield selected the most conspicuous magnetic "bays," (which by then were known to relate to auroral substorms). He also sorted Cahill's hourly means into "clearly northward," "clearly southward" and "indeterminate" and found that overall, there were roughly equal numbers of hourly means in each category. Then he found that each of his 10 big bays had southward Cahill field, a result I viewed as like a student obtaining full marks on a multiple choice test of 10 questions with three choices for each. With hindsight, the fit was better than average, but it gave me the high of a successful prediction and great confidence and attracted attention, though I have several times heard gossip about unworthy emotions being aroused. It seemed to me that the result could be popularized because of its similarity to weather lore, but when I described it to science correspondents, they barely stirred from their after lunch torpor.

This is a convenient turning point to break off from the reconnection story and describe wave-particle interactions, which attracted me partly because of their possible role in the anomalous resistivity required for reconnection.

4. WAVE-PARTICLE INTERACTIONS

The analysis of wave-particle interactions is a very important and large part of plasma theory. I had some original ideas independently but will not claim any priority; most work in this area was associated with fusion, and much of it originated in the Soviet Union and was published first in Russian. I followed and used the subject, but a chronological order is inconvenient here. I will now progress through the linear, quasi-linear, and nonlinear branches.

In linear theory I deviated from the mainstream by using the technique that calculates the perturbation of the trajectory of a test particle as a past history integral and afterward uses this to obtain the distribution function. For years I gave a postgraduate course to a very small class always containing two or three postgraduates from the fusion group at Imperial College, and I remember sometimes one of them would object that my method was different from that in their books. Tom Stix, who very early wrote an excellent book on waves, told me he had been familiar with the test particle method for a long time. Eventually, I suggested to Colette Robertson that she should perform a rather general calculation, which would demonstrate the method as well as providing an assessment of "finite Larmor radius stabilization," which at the time was fashionable in the fusion group.

Another facet of linear theory, which occupied my attention and that of one of my first graduate students at Imperial College, Michael Houghton, was the behavior of wave packets. The distinction between convective and nonconvective instabilities was topical at the time, and it was necessary to discover how this depended on the size of the envelope of a wave packet, which is less constrained in space than in a laboratory experiment. Convective instabilities are important in space and will be mentioned again shortly.

Wave instabilities are synonymous with negative Landau damping and Landau and many books derived positive damping by analytical continuation from positive growth. In terms of past history integrals it is easy to see that the integrals converge well for growth but not for damping, but full satisfaction requires some physical confirmation of the mathematical magic and this is provided by quasi-linear theory. This uses zero damping or growth, but it can be argued that strongly damped waves are uninteresting, while the plasma would not exist in an unperturbed state in which the growth rate was high. Quasi-linear theory also uses the random phase approximation, and the validity of this can be addressed by nonlinear theory. The outcome of quasi-linear theory is described as diffusion in phase space, from which the rate of change of the kinetic energy in the plasma can be obtained. Balancing this with the rate of change of wave energy provides the formula for Landau damping, but for electrostatic waves the calculation of wave energy can involve some subtlety. The concept of Landau damping is much more general and for electromagnetic waves the expressions for wave energy are usually familiar. Landau damping always involves some kind of resonance and, when a magnetic field is imposed, the original resonance of particles traveling at the phase speed of the wave becomes just the-zero harmonic of gyroresonance. Remembering the connection with Landau damping, quasi-linear theory will be the next topic.

The concept of quasi-linear diffusion is related to earlier qualitative ideas like eddy diffusion and eddy viscosity. The concept occurs for instance in Fermi acceleration sometimes later known as "stochastic acceleration." The concept is also that of the Fokker-Planck formula for diffusion due to a random walk with small steps, which is almost obvious, but of course important. Not long before I retired, I discovered how to obtain the diffusion coefficient from past history integrals over unperturbed trajectories and used this in a way to be described in the next section. Later I met Alex Galeev, and he told me his group had also discovered the method.

The magnetosphere provides an excellent opportunity to study quasi-linear behavior in plasmas and conversely quasi-linear theory can explain a great deal about the behavior of energetic particles in the magnetosphere. The motion of a charged particle can have three periodicities: gyration in the magnetic field, bounce between magnetic mirrors and drift round the Earth. Corresponding to each periodicity there is an adiabatic invariant, which may be approximately conserved, and it will be convenient to treat them as conserved when the approximation is adequate in the context. Thus bouncing between mirrors depends on conservation of the first invariant (corresponding to gyration), while drift round the Earth depends on conservation of the "longitudinal invariant" (corresponding to bounce) as well. Quasi-linear diffusion always involves a resonance relating to one of the periodicities. It may also be necessary to point out that, since the phase space for the particles has six dimensions, the diffusion is constrained to a curve in phase space and the determination of this curve will be illustrated for each resonance and source of noise. It may also be noted that rough orders of magnitude are 1 ms for an electron gyroperiod, 1 s for a proton gyroperiod or electron bounce period, 1 min for a proton bounce period, and 1 hour for a drift of either species. It is suggested that the past history integrals be viewed as similar to Fourier integrals.

Diffusion may be best understood in the case of drift resonance. The noise is dominated by impulses, mostly associated with substorms, which are easily strong enough to show up individually. The observations are interpreted in terms of a dawn-to-dusk electric field causing particles to move inward and gain energy. Sharp increases of particle flux are observed on satellites and called "injections." Recurrent weaker increases generally agree with the drift period (proportional to energy) and observations at different local times and distances generally fit with acceptable electric field models. Thus, in this case, partly because of the long timescale, a microscopic view of the fine structure is easily feasible. Nevertheless, with a broad brush picture, the outcome can be described by diffusion, as was discovered by early observations that were of course rougher [*Nakada et al*, 1965]. Plotting the distribution function for fixed values of the first two adiabatic invariants showed a flat region far out and within a critical distance a rapid decrease in intensity toward the Earth. The latter indicates a rapid loss of particles, but the flat region shows that the dominant mechanisms far out can be described by diffusion, which conserves the first two invariants. So it can be claimed that, while substorms are far from understood, their effect on the trapped particle population can be comprehended both microscopically and macroscopically.

It is convenient next to discuss gyroresonance, which is a big subject. I shall restrict myself to resonance between electrons and the whistler mode, which still comprises a large literature, and is a topic in which I maintained an interest for many years. The whistler mode is an electromagnetic mode at frequencies below the electron gyrofrequency, and it will be assumed that the phase speed is slow compared to the speed of light. In the magnetosphere the plasma typically consists of a "thermal" component somewhat hotter than the ionosphere and a trapped energetic particle component. If the thermal component dominates the density, this determines the phase speed of the whistler mode, while gyroresonance requires energetic particles and these control the Landau damping, which is usually negative. In the real magnetosphere there are complications due to the nonuniformity of the magnetic field and nonuniformity of the plasma density, which is believed to provide ducts. Theory, however, begins with the simple case of uniform field and plasma and the unperturbed trajectory of a test particle is described by a constant parallel velocity and a rotating perpendicular velocity. A Fourier component of any noise spectrum is a plane wave and a general plane wave is considered, though the dispersion equation is tricky when the phase fronts are nearly parallel to B. The fundamental gyroresonance requires an electron to "see" the wave as having the gyrofrequency so a blue shift is needed since the wave frequency is less than the gyrofrequency in the plasma frame. Therefore the electron and wave are going in opposite directions. Now I like to use a theoretical techinique that will have two further applications in this paper; this is the use of a frame in which the wave has zero frequency. In this application the frame moves parallel to B relative to the plasma, to avoid other complications; its speed is modest provided the wave fronts are not nearly parallel to B, and it may be called "the wave frame." Because the wave magnetic field is constant in time, the electric field has a potential and the particle energy, including this potential, is conserved. For quasi-linear theory, however, only resonant effects are sought, which increase proportionally to time, whereas the potential is limited by the amplitude of the wave and can only oscillate in time, so the kinetic energy can be taken to be conserved. When the past history integral for a test particle is examined, it is seen to be just like a Fourier integral picking out the circular polarization, which is dominant in the whistler mode, and it pleasingly agrees with the physics. Now the point about the wave frame is that it is in this frame that energy is conserved, and this determines the surface in velocity space to which the diffusion is constrained. Either using a sketch or calculating algebraically and remembering that the electron is traveling in the opposite direction to the wave, it is easy to see that in the plasma frame an increase of energy must be associated with an increase of pitch angle. It may be noted that the terminology "pitch angle scattering" is not strictly accurate.

The simplified theory relates well to trapped energetic electrons in the magnetosphere, a useful realization of a mirror machine. If one looks for energetic electrons coming up just above the ionosphere, of course, there are very few, and this is the loss cone. If one follows the trajectories of the few along a field line, the number will gradually increase as a result of pitch angle scattering from the more populous trapped electrons with mirror points above the ionosphere. If enough are scattered over the full length of the field line, there may be a substantial flux, now downward, into the other ionosphere. This is confidently believed to be the origin of the diffuse aurora. It is also a loss mechanism for the trapped electrons and demonstrates the importance of waves for the radiation belts. It is fortunate that both particles and waves are observed so that their effects on each other can be analyzed.

The trapped electrons amplify the whistler mode whenever there is a loss cone. Because the energy of a test particle changes in the same sense as the pitch angle, diffusion to smaller pitch angles implies diffusion to lower energies, and the energy released goes into the waves, driving at least a convective instability. It has been found, using particle data, that the amplification integrated along a field line can be large, and this accounts for the strength of whistlers. Whistlers had been noticed in audio systems long before Owen Storey deduced their origin from lightning and propagation along field lines. A great triumph for quasi-linear theory was Kennel and Petschek's trapping limit. They considered a wave packet making a round trip, propagating along a field line, being reflected at each ionosphere and ending where it started. If the amplification should outweigh the losses at ionospheric reflection, the waves would grow and catastrophe would loom. However, pitch angle scattering causes a loss of trapped electrons and consequently reduces the amplification, so that the trapping limit corresponds to the condition of zero round trip amplification. Observations show much complexity in the whistler mode noise, which will be mentioned in relation to nonlinear theory.

Resonance involving the bounce period remains to be discussed. This has received little attention but has been of considerable interest to David Southwood and me. In the magnetosphere the waves involved are assumed to be MHD waves, which have already been introduced as geomagnetic pulsations, so the wave frequency is determined by MHD field line resonance. Electron bounce frequencies are high, and the interesting case concerns energetic protons with a Doppler shift due to their westward drift, which is important, because Landau growth can easily occur. Even in this case, the bounce frequency is substantially higher than the wave frequency and a large Doppler shift is necessary, so that the east-west wavelength of the wave must be short. The resonance condition can then be used to determine the east-west wavelength, which turns out to be comparable to the Larmor radius of a resonant particle, so that the MHD approximation is barely valid.

It is fruitful to consider a wave whose longitude dependence is simply $\exp(im\phi)$, and the interesting case requires m to be large. The frame transformation trick can be used, if the wave frame is allowed to be a rotating frame and the angular velocity is small enough to neglect the complications arising from the use of rotating frames in general. The important feature in the rotating frame is in the undisturbed electric field; this has an added component in the meridian plane, which can be described by a potential that is constant on a field line. Conservation of particle energy in the wave frame then relates the change of a particle's energy to its displacement across the magnetic shell. Consequently, if the first invariant is valid, it is necessary only to compute the change in either energy or displacement. As far as I know, accurate computations have yet to be performed, but *Southwood et al.* [1969] used guiding-center theory to estimate the displacement and even this was not trivial. We found that one mechanism dominates. While a wave may cause a drift across the field due to either field strength gradient or field line curvature, the important effect is the change in field direction, which of course diverts the parallel particle motion. Observation on geostationary spacecraft shows field perturbations in the radial direction, which results in inward or outward particle motion. For the fundamental bounce resonance the tilt of the field line reverses between northward and southward particle transits, so that the particles displacement is in the same sense for both. We used this picture to obtain an order of magnitude expression for the diffusion across shells and then, using the constraint on the change of energy, the growth rate can be similarly expressed. Amplification of the wave can be caused either by a spatial gradient or by population inversion in the particle energy spectrum. I expected that the amplification by a gradient would worry the fusion community but they seemed to be unperturbed. In the magnetosphere, either cause is possible; particle injections followed by drift cause population inversions that have been observed.

Again, a crude theoretical picture led to a notable feat of data analysis depending on some good fortune. A geostationary satellite, ATS 6, had as its principal mission the relaying of television in India, but initially the system was tested on some of the inhabitants of the Rocky Mountains, so it had to be moved from that time zone to the Indian longitude. Furthermore, ATS 6 carried a magnetometer, as did the GOES satellites and in particular one which ATS 6 would pass in the course of its move. Jeff Hughes undertook to compare the pulsations seen on ATS 6 and GOES. Before the move, when the spacecraft were hours apart in local time, he already found high coherence for pulsations in the morning, an interesting result. However, his good fortune was that the local time when ATS 6 actually passed GOES was in the afternoon. What he found was no coherence throughout the overtake [*Hughes et al.*, 1978], and this rigorously shows that the wavelength in a plane perpendicular to B is short. It strongly suggests that the waves are excited by energetic protons, which are most intense in the evening sector, but also more intense in the afternoon than the morning. So the outcome showed a successful prediction with a bonus, and it is possible that this class of pulsation will be surveyed by Cluster.

Turning now to nonlinear features of wave-particle interactions, these may be negligible for electromagnetic modes when the magnetic amplitude is much less than the undisturbed field, though this argument is sometimes too simplistic. Because of the complexity of nonlinear effects, however, it is easiest to start with a monochromatic electrostatic wave in one dimension without any magnetic field. In the frame of the wave near-resonant particles are trapped between potential maxima; in phase space, which is two-dimensional, their trajectories can be drawn and, for waves of constant amplitude, they are closed curves forming eddies in phase space. Next apply Liouville's theorem to see how the wave changes the distribution function f by viewing contours of constant f . It is seen that the eddies wind up the contours into ever tightening spirals and handwaving suggests eddy diffusion, a result qualitatively similar to quasi-linear diffusion. There is a difference, however, for a monochromatic wave. In the analogous case of stirring, paint say, the spiral is eventually smeared out by real diffusion and, in the plasma case, one smearing mechanism is represented by the Fokker-Planck formulation of collisions between particles, which resembles a diffusion in velocity. After smearing, f is nearly uniform throughout the eddies, but little changed elsewhere. I described the resultant f (v) as having a "ledge" centered on the resonant velocity [*Dungey*, 1961], but the fusion jargon is "plateau," which suggests their wave is louder.

There is a substantial literature on the nonlinear features of electrostatic waves, but my interest switched to the whistler mode. Having adapted the stirring picture for the whistler mode [*Dungey*, 1963], I was struck by the triggering of emissions by Morse pulses [*Helliwell*, 1965], which is necessarily nonlinear, because the frequency changes. Fortunately and fortuitously, emissions were seen or heard only from dashes and not from dots, so that a critical value was hit on. The criterion can be seen as the converse of the validity of the quasilinear approximation: the change in a particle's trajectory due to waves changes its resonant frequency and quasi-linear theory loses its validity when this frequency is removed out of the bandwidth of the waves. The criterion therefore involves the bandwidth as well as the amplitude. A basic nonlinear mechanism was described *by Das* [1968] using the "ledge" picture. The picture suggests that the flattening of the ledge should cause steepening of the distribution on either side. Das pursued this idea for the whistler mode and showed that the amplification was much enhanced at frequencies on either side of the triggering frequency. This became "the sideband instability" in the jargon, but I suspect that Das received less than the credit due to him. Not only must the electron distribution function contain much fine structure resulting from such interactions, but observation shows that the emissions are complex and furthermore, the naturally occurring noise consists of numerous features of relatively narrow band. The best known is "dawn chorus": on an audio system many rising tones are heard, reminiscent of birds, and a dynamic spectrum can be used to record these for analysis. The whole phenomenon is an illustration of complexity. The trajectory integrals for electron perturbations are dominated by the sections surrounding resonances, leading to fine structure in f, but if the resonant contributions are added as a random walk, the result is a smooth diffusion, compatible with quasi-linear theory using a time-averaged spectrum for the emissions. The fine structure may be detectable in the aurora; perhaps it has been seen.

5. THE NEUTRAL SHEET

The pursuit of reconnection becomes the study of neutral sheets and really there is only one, in the geotail, because it is so much better observed than any other. Appropriate coordinates have z perpendicular to the sheet and y in the direction of the current. Before resuming the main thread, it is convenient to mention two related developments in 1969. Stan Cowley, who had only just started graduate study, pointed out the likely need for an electric field in the z direction in the neutral sheet, and *Cloutier et al.* [1970] discovered a field-aligned current in an auroral arc, which to me was surprisingly intense. Still believing discrete auroras to connect magnetically to the neutral sheet, I expected the Cloutier current to be driven by the Cowley field. If the Cowley field changed the current should propagate as an Alfvén wave to the aurora. When the Cowley field is steady, there should be a standing Alfvén wave, on the assumption that there is also a steady flow corresponding to a dawn-to-dusk electric held. The crucial question seemed to be the thickness of the current layer, and I eventually tackled this [*Dungey*, 1982] rather crudely, finding the layer to be thin.

The main thread of the neutral sheet originates from my second Penn State student, Ted Speiser, looking at particle trajectories. It was well known that these oscillate across the central plane, but Speiser added a weak z component of magnetic field and discovered that this slowly turns the trajectories around the z direction: possibly obvious but an important result as these trajectories are typical. A y component of electric field can be removed locally by transformation to a frame moving in the x direction, allowing calculation of a change of energy in the original frame. This may be the mechanism of proton injections.

The crucial problem in the neutral sheet is the nature of anomalous resistivity. The importance of noise was emphasized by *Axford* [1967]. For many years it seemed the best candidate, but any theory would be beset by difficulties. Noise must be composed of waves. Clearly, one hopes to use the experience of plasma waves in the vast literature, but this application meets two obstructions at least. The magnetic field is as far from uniformity as possible, and the current associated with the wave at any point does not depend only on the wave field at that point. Strictly, the perturbation involves past history integrals over Speiser trajectories and the z dependence of the wave field for a normal mode is unknown. Some of the difficulties were recognized by *Dungey and Speiser* [1969], and we also pointed out the importance of radiation out of the sides of the sheet; with such a thin layer supplying the energy it seems likely that the noise will consist of waves that are ducted by the sheet.

During the next decade, several postdocs worked with me on the noise problem, but it remained obstinate. The last was Peter Smith, who, while a graduate student, made good progress with the linear problem. After that, and partly because I had decided to take early retirement, we bulldozed the quasi-linear problem with crude approximations to obtain a formula for the anomalous resistivity [*Dungey and Smith*, 1984]. We assumed that the waves had a low phase velocity and selected the resonance in which the electron sees the frequency of its own oscillation across the sheet. With further approximations we wrote a trajectory integral for the perturbation in the y component of velocity and used the newly discovered technique for relating the mean square perturbation to the power spectrum. Then using the wave frame the vector perturbation was described by an angle and the result was the diffusion coefficient for this. Finally, with a simple but plausible form for the velocity distribution, we obtained the rate of change of current density due to the noise and hence the resistivity. Putting in numbers expected for the tail neutral sheet showed a required noise level, which I felt was excessive. Perhaps subconsciously I started to seek an alternative.

Having seen that this problem is quantitative, it is as well to mention the evidence for the reconnection electric field. Flows in the ionosphere indicate an order of magnitude of 1 mV/m, and several other indirect measurements agree. Probably the most convincing evidence comes from a rather elaborate analysis of data from the energetic particle anisotropy spectrometer on ISEE 3 during its excursion down the tail [*Richardson and Cowley*, 1985]. Their value was 0.4 mV/m and was associated with plasmoids, which are believed to be produced by reconnection, so perhaps 1 mV/m is an unusually high value. Nevertheless, this electric field is strong: the acceleration of electrons can be thought of as the speed of light in a few seconds, so that the effective collision frequency needs to be many collisions per second.

An alternative crystallized three years after I retired. Following the idea of Speiser trajectories, suppose the z component of the magnetic field is proportional to x, βx say. From β and the electric field a timescale is obtained, which may be an effective collision time; if so, the resistivity increases when the current density increases. The value of β is not known and may fluctuate wildly. A large value is needed, however, to yield the required electric field and the problem is still open. I pursued the idea by looking at simplified velocity distributions and found that they were extremely anisotropic near the neutral line. This implies anisotropy of the stress tensor, which is effectively viscosity and I suggested that resistivity is replaced by "electron viscosity." I proposed a set of equations for computations with anisotropic stress [*Dungey*, 1988]. Later, I wondered whether electron viscosity in the inflow could be more important than that in the outflow and hoped some simulation expert would take up the problem. No one has yet, but last year Stan Cowley gave me a computer and now I am a DIY (Do It Yourself) simulator. This is my view of the future.

Looking into the future, I still have a copy of the suggestion I submitted for the "Second Large ESRO Project." Entitled "Tetrahedral observatory probe system," it is just like Cluster, and the scientific objectives specified in it still make sense. If I live to be 75, with luck I may see data from Cluster and perhaps some progress in modeling the neutral sheet.

6. THE FUTURE HAS STARTED

Comments on Simulations of Magnetic Reconnection

I am hopeful that recent activity [*Cai et al.*, 1994; *Mandt et al.*, 1994] represents the onset of rapid progress. The contrasts between these two papers may start a debate about which physical features are important:

C, D & L emphasize the pressure tensor
M, D & D neglect the pressure tensor
M, D & D emphasize electrons
C, D & L use only a background of electrons
M, D & D emphasize $\mathbf{B}_y$
C, D & L omit $\mathbf{B}_y$

A feature which is yet to appear occurs where the electrons are magnetized but the ions not, the magnetic field strength being a fraction of its value at the boundary. Then the electron perpendicular pressure should be reduced.

The Role of Satellite Measurements in the Development of Magnetospheric Physics: A Personal Perspective

D. H. Fairfield

Laboratory for Extraterrestrial Physics, NASA Goddard Space Flight Center, Greenbelt, Maryland

Included here are the author's personal recollections of the early days of magnetospheric research. Emphasis is on his role in the confirmation of Dungey's theory of the open magnetosphere and on the early definition of the magnetosphere magnetic field configuration.

INTRODUCTION

Contemporary students may find it difficult to realize that prior to the 1960's the magnetosphere was unknown. Even in the early 1960's the field of plasma physics was in its infancy with no textbooks, virtually no university courses, and only a very few monographs on ionized gases. Only a few individuals such as Sydney Chapman in the 1930's had even considered the question of what might happen when an ionized gas encountered a dipole magnetic field. For many years these and subsequent early speculations [*Stern, 1989*] were of interest to just a small community of scientists. Only with the advent of spacecraft measurements in the early 1960's did the field expand into the vital field of geophysical and plasma physics interest that we know it to be today. In the following I offer my personal perspective of what it was to be educated and participate in those early years of magnetospheric physics.

EDUCATION OF A SPACE PHYSICIST

In the fall of 1957, I was a sophomore at Beloit College taking my first college physics course when Sputnik was launched. In retrospect, my involvement with space physics began the following day when we were assigned a mechanics problem involving orbiting bodies. I had no particular interest in space research before completing my undergraduate degree in 1960 although I do remember giving a senior seminar report on the Van Allen radiation belts. My professor had no knowledge of the subject and could offer little help in understanding the physics of trapped particles.

Discovery of the Magnetosphere
History of Geophysics Volume 7

Published in 1997 by the American Geophysical Union

I accepted a research assistantship in the Ionosphere Research Laboratory at the Pennsylvania State University, where I commenced study in the fall of 1960. The next spring I was offered the opportunity of becoming the student of Jim Dungey, a consultant in the Ionosphere Research Laboratory. Jim was permanently located in England and only visited Penn State for a week or so once or twice a year. I accepted this unusual arrangement and we subsequently interacted via the occasional visits and frequent letters. I always felt these concentrated weeks with Jim were more valuable than most student's more frequent contacts with their advisors. I had Jim constantly available during his visits because the ionospheric physicists in the laboratory had little interest in the newly discovered "magnetosphere" above the ionosphere. They also had little appreciation of how the magnetosphere affects the ionosphere. Contact with Jim was particularly valuable since he was one of a few individuals with a deep knowledge of plasma physics and magnetohydrodynamics. The Penn State physics department did not teach a plasma physics course prior to my graduation in 1965 and my only formal introduction to the subject was via sitting in on a course taught by a young professor in the nuclear engineering department.

For my masters thesis I was given the assignment of analyzing high latitude magnetograms of the International Geophysical Year in order to determine the ionospheric current patterns that Jim believed to be the result of magnetospheric convection. He further believed that these currents were driven by the solar wind electric field connected to low altitudes via the open field lines in the revolutionary new model that he was just publishing [*Dungey; 1961*]. To accomplish my thesis I collected magnetograms and produced equivalent current patterns in high latitudes. Help in doing this was initially provided by a room of women working at mechanical calculators, although I later became the first person in the Ionosphere Research Lab to walk across campus and submit card decks to the new IBM digital computer – other laboratory computations were done on an analogue computer built by the electrical engineering department. My thesis results [*Fairfield; 1963*] generally supported Jim's ideas on convection and gave me further insights into the nature of ionospheric currents and what would later be called substorms.

Although Jim's famous paper on this subject had been published in early 1961, it was far from being accepted. The early thinking about the magnetosphere was based on the Chapman-Ferarro model, where a solar wind plasma enveloped the Earth and was completely excluded from the magnetospheric cavity. Without Dungey's interconnected field lines, it was difficult to see how the solar wind could do more than confine the Earth's field. Time variations could be due only to waves caused by a buffeting of the cavity by a time varying solar wind.

My first professional trip occurred early in my Penn State tenure when Jim suggested I visit Ottawa to talk to Colin Hines and Ian Axford, who had just published their famous paper [*Axford and Hines; 1961*] on magnetospheric convection, in which they suggested that the convection might be driven by a viscous interaction at the magnetopause. They received me very graciously, and I remember trying to explain Jim's new recently-published reconnection theory to them, which they had not yet read. As a new graduate student, I am sure I did a very poor job of explaining this new view. In any case, Axford particularly became one of the early converts to reconnection theory. Both these magnetospheric theories evolved out of an understanding of magnetosphere convection and differed mainly in the driving mechanism at the magnetopause. Despite their basically similar understanding of the magnetosphere, Dungey versus Axford/Hines has been portrayed as oppositional viewpoints ever since.

An example of how the closed magnetosphere paradigm dominated thinking of the time followed from work I did under Jim's guidance while we both visited NASA's Ames research center during the summer of 1963. Jim had me investigate the behavior of the mirror points of particles trapped in the magnetosphere as they drifted around the Earth in a distorted dipole magnetic field. By assuming conservation of the first two adiabatic invariants, it was possible to calculate particle mirror points in a magnetic field model in the noon meridian and then find the comparable location in an appropriate model at midnight; it was not necessary to have a complete field model and inquire how they moved from day to night. For the two model fields, we simply added a uniform northward field to a dipole on the dayside to produce a compressed field and a uniform southward field on the nightside to produce an extended taillike field. This work occurred before the discovery of the magnetotail and the paper was initially rejected by the referee who stated, "The author assumes, incorrectly, that the effect of the solar wind on the geomagnetic field is to push field lines in on the front side and push field lines away from the Earth on the night time side. However, it was realized some time ago [references given] that pushing in on the front face of the geomagnetic field also pushes field lines in on the night time side, thus the southward uniform field on the night time side of the diagram as assumed by the author is of the wrong sign." Fortunately, very early observations of the depressed fields of the magnetotail current sheet were just becoming available, and I was able to get the paper published [*Fairfield 1964*]. This example demonstrates how theory dominated the field prior to spacecraft measurements and how wrong theory can be without measurements to guide it.

Another highlight of that summer was a trip to a meeting in Los Angeles where, at Jim's suggestion, I talked to a young graduate student of Sydney Chapman, Syun Akasofu, who was just finishing his degree at the University of Alaska. Syun told me about his work on the morphology of an auroral substorm which I, along with many other people at that time, regarded as a great accomplishment. This new perspective helped me considerably in understanding the auroral zone magnetograms I had been looking at for two years. In the present era of global auroral pictures and stacked plots of magnetograms from chains of stations it is difficult to appreciate how difficult it was to separate time and space variations with the limited ground data of that time.

Since the Dungey theory needed a southward interplanetary field to open the magnetosphere and drive ionospheric currents, Jim gave me the task of investigating the relation between the north/south polarity of the interplanetary magnetic field (IMF) and geomagnetic activity. The Explorer 12 spacecraft launched in August 1961 was the first long lived (4 months) Earth orbiting spacecraft to provide measurements outside the magnetosphere. Jim arranged for me to spend the summer of 1964 at the University of New Hampshire with Larry Cahill who had the magnetometer experiment on this spacecraft. Explorer 12 was designed to study Earth's magnetic field and with an apogee of 13 Re was the first spacecraft to map out the magnetopause position. Since it was designed to measure strong fields, however, the experiment had a large dynamic range and a large digitization error (±12 nT). Launch successes were always problematical in those days, computers were new, and Larry's experience had previously been on rocket flights of a few minutes duration. Probably for these reasons, minimal preparations had been made to automate the data processing from an experiment that produced data at the rate of 3 vector measurements per second. When I arrived in New Hampshire the halls of the Physics building were lined with shelves containing printouts of this high resolution data; undergraduate students were busy at mechanical calculators computing average fields for immediate use while others were striving to automate the process on a modern (for the time) digital computer.

I had a very enjoyable and productive summer comparing ground magnetograms with the magnetic field outside the magnetosphere. With paper magnetograms and a limited amount of magnetosheath data, it was difficult to do anything more than prepare a set of examples. Even so, I was impressed with how well the southward field seemed to correlate with geomagnetic activity [*Fairfield and Cahill, 1966*]. I wondered if research would always be this easy, but having so little experience, I was in no position to judge. Little did I know that this correlation would remain one of the clearest results in sun/Earth relationships over the years.

I also remember being impressed with the potential importance of this work. It confirmed the major prediction of Dungey's new theory and it clearly offered the possibility of predicting magnetic activity and related phenomena such as the high latitude ionospheric disturbances that could result in communications blackouts so important during the cold war. It also appeared to be the answer to another longstanding question of why some sudden commencements are followed by storms (those with southward IMF) and others are not (those with northward IMF). In those days, however, the common view was that science was done for its own sake and it was up to others to apply the results if they wished; space weather had to wait until the 1990's. At least Jim Dungey was pleased with my results and I received my degree.

EARLY YEARS AT GODDARD

I was familiar with the work of Norman Ness and particularly remember attending a one day symposium on results from the IMP 1 (Interplanetary Monitoring Platform) spacecraft at NASA's Goddard Space Flight center in March 1994 when he first presented his mapping of the position of the Earth's bow shock. This was a very clear and exciting result and it left no doubt that collisonless shocks were a reality in space plasmas.

Norman's IMP magnetic field data were the best interplanetary measurements available at this time so I applied for, and later accepted, a postdoctoral position at Goddard in order to pursue the relationship between the IMF and geomagnetic activity. Norman had recently published his very clear evidence for a magnetotail [*Ness, 1965*] when I arrived at Goddard in the fall of 1965. This was of special interest to me since I had earlier puzzled with Jim Dungey over early unpublished Explorer 10 results after Jim had visited Goddard on the way to Penn State. This spacecraft had skimmed the then unknown magnetotail in 1961 on its one outbound pass, moving back and forth between two plasma/field domains that recurred out to 42 Re. At that early time their interpretation was not at all clear, although they were eventually published with a tentative association with an extended magnetospheric cavity popular at that time [*Heppner et al., 1963*]. These early Explorer 10 results, along with Jim Dungey's theory, all fell into place in the light of Norman's tail observations.

On arriving at Goddard, Norman turned the IMP 2 analysis over to me. IMP 2 was designated a failure by NASA headquarters because a rocket malfunction left it in a near equatorial orbit with an apogee of only 16 Re – only about half that planned for monitoring the interplanetary medium. Interestingly, this is an orbit that most magnetospheric physicists would kill for today. The IMP 2 experiments worked well and the spacecraft had a very productive 6 month lifetime [*Fairfield and Ness, 1967*].

At Goddard I was able to pursue correlations with geomagnetic activity with plenty of high quality IMP interplanetary magnetic field data, but the problem was still how to quantify the ground activity on a time scale better than the three hour Kp index. The geomagnetic AE index had just been defined by Davis and Sugiura [*Davis and Sugiura, 1966*] and digitized ground magnetograms were becoming more available. Given this situation I decided to produce the AE index on the 2.5 minute time scale of the digitized data. I was able to enlist the help of the National Space Science Data Center at Goddard and the index was produced, initially with six auroral zone stations. The NSSDC continued production of this index under my direction until the effort was transferred to the World Data Center in Boulder in 1968. With the AE index I was able to demonstrate the IMF – geomagnetic activity relationship in a more quantitative manner and I presented these results at the 1967 COSPAR meeting in London. I published the results in the meeting proceedings [*Fairfield, 1967a*], and in the process learned a lesson on the dissemination of scientific results.

Having grown up in the early days of space research where one gleaned the limited literature for any article of possible relevance to one's research, I had developed the attitude that it didn't much matter where one published one's results. My paper in the COSPAR proceedings provided confirming evidence for what became the prevailing theory of the solar wind/magnetosphere interaction, but it also became what is undoubtedly the least referenced paper I have ever written. This IMF/AE confirmation of Dungey's theory was known and accepted by a few knowledgeable people in the field but this result was probably not widely known and accepted until Roger Arnoldy's clear quantitative exposition in a JGR paper in 1971 [*Arnoldy et al., 1971*], which did not even reference my COSPAR paper.

Other results from the IMP 2 spacecraft included a paper which I titled "The ordered magnetic field of the magnetosheath" [*Fairfield, 1967b*] in an attempt to counteract the prevailing impression, based on the enhanced magnetic field fluctuation level in the magnetosheath, that this region was very turbulent. This paper also used a very limited amount of IMP I data taken simultaneously with IMP II. These data demonstrated how frozen-in magnetic fields are indeed convected from one spacecraft to another in what to my knowledge was the first demonstration of this well known property of fully ionized plasmas. Lacking a formal plasma physics education, I always considered myself fortunate in being able to learn plasma physics from such real life examples. These IMP 2 magnetic field data also demonstrated how the interplanetary field was convected into the magnetosheath and draped over the magnetopause.

The orbit of IMP 2 was also ideal for investigating the magnetospheric configuration and particularly the transition from the dipolar region to the tail [*Fairfield, 1968*]. The sweeping back of field lines was demonstrated graphically in this work and contours of constant B in the equatorial plane were determined. These data were used in obtaining the first quantitative determination of where low altitude field lines map to in the outer magnetosphere. Although static and graphical, it remained the only data-based magnetosphere model until later quantitative computer fits to experimental data were carried out. Since equatorial particles conserving their first adiabatic invariant drift around the Earth in a constant magnetic field, the constant B contours gave an early quantitative idea of how particles circled the Earth.

The launch of the IMP 4 spacecraft in 1967 carried Norman Ness's sensitive triaxial magnetic field experiment which produced a vector magnetic field every 2.6 seconds, which was a significant improvement over the earlier IMP's. Prior to this time there were a few observations suggesting that Earth-associated particles were seen upstream of the bow shock, but the accepted view was that the nature of a shock was incompatible with waves propagating into the upstream region. IMP 4 data greatly extended an earlier two-component search coil magnetometer measurement of apparent upstream waves and clearly demonstrated that there was a permanent presence of waves in an upstream region, later termed the foreshock by Gene Greenstadt [*Fairfield, 1969*]. These high quality IMP 4 data also allowed determination of the typical wave frequencies (.01–.05 Hz) and polarizations (left handed) and even indicated the speed of propagation into the upstream region. I was always grateful to Derek Tidman who suggested to me that generating the waves by protons streaming away from the shock was the way to get around the inability of waves to propagate into the upstream region. This foreshock work was particularly satisfying in that it revealed a new region that remains to this day one of the most fascinating plasma physical laboratories available for experimental study.

SUMMARY

I will remember the early years of space research as a golden age that can never be repeated. Little was known but the rapid pace of spacecraft launches rapidly produced a wealth of revelations and insights. At virtually every scientific meeting, data from a new spacecraft would yield exciting new results. I personally felt particularly privileged to work on magnetic field data from high apogee Earth-orbiting satellites that traversed all regions of the magnetosphere and interplanetary space. Magnetic field data is vital in any plasma environment and these eccentric orbits allowed me to work on problems in various regions of space. With so much less known in the early days, it was much easier to keep informed about different regions compared to today's burgeoning literature and greater specialization. In retrospect, it seemed relatively easy to achieve significant new results in those days, but one must remember how much more difficult it was to process and manipulate data without today's high speed interactive workstations, mathematical toolkits, and efficient plotters.

Without spacecraft data in the early days, theory led the field, but in some cases led in the wrong direction. Advances in magnetospheric physics are accomplished by individuals with different backgrounds and different scientific styles. I have realized that my own personal style is contrary to the classical view of science where one starts with a hypothesis and then compares it with data. I tend to view the magnetosphere as a giant jigsaw puzzle. I first look at the data to see what they seem to be saying and then look for an explanation and how it fits into the bigger picture. It is often rewarding to solve a little bit of the puzzle and one doesn't have to keep working on the same intractable piece.

Early Times in the Understanding of the Earth's Magnetosphere

T. Gold

Cornell University, Ithaca, New York

The recognition by Alfvén that magnetic fields in conducting gases are convected with the fluid, led to a better understanding of the relationship between solar phenomena and magnetic field disturbances at the Earth. From the observed abrupt beginnings of magnetic storms, I concluded that a type of shock-wave must exist, in which the magnetic field interactions and not particle collisions were the dominant effect. Solar flares, explosions at the Sun that often caused magnetic storms, were attributed to the sudden destruction of solar magnetic fields (in a paper by F. Hoyle and myself). Highly energetic particle emission at the Sun and their reception at the Earth allowed one to make statements about the magnetic fields in solar system space, even before direct spacecraft exploration was possible. The mobility of tubes of force of the terrestrial magnetic field, allowed by the insulating atmosphere, resolved a puzzling aspect of Van Allen's radiation belts. In the paper giving this result, I also introduced the word "Magnetosphere", to describe the region around the Earth in which the magnetic field had the dominant effect on gas motions.

My interest in the magnetic field of the Earth was aroused for two reasons. One was the great puzzle why that field existed at all: what in the Earth could give rise to such a field?

The other was the immense amount of observational evidence that had been gathered over many years, about relations between solar events and a variety of types of disturbance observed in the Earth's external magnetic field, as well as in long-distance radio propagation effects. How were all these to be explained?

The explanations in vogue at the time in both of these areas seemed unsatisfactory to me, or at least lacking in some essential components. Could I perhaps puzzle out some of the answers?

My first step was simply to read the major textbook on the subject: two thick volumes by Chapman and Bartels, entitled "Geomagnetism." They taught me more than I really wanted to know, but then I could not leave out anything since it might contain a vital clue. Although the books were well written and presented the knowledge of the day, they confirmed my conviction that major components were missing, and not much imagination had gone into the search for them.

Discovery of the Magnetosphere
History of Geophysics Volume 7

The enormous puzzle of the Earth's field itself was presented on just one page of the two books, and it explained nothing. There was not even the emphasis I would have expected on the difficulty of finding an explanation. (I am still doubtful that the present day viewpoint is the correct solution.) The other, the numerous external and mostly solar influences were given in the greatest detail, but also, I felt, without giving the sense of urgency to find an explanation for them.

Alfvén's influence

In the early 1950's I had much contact with Hannes Alfvén whom I had met when he visited Cambridge and I had in turn visited him several times in Stockholm. I was most impressed with his simplifications of that otherwise quite impenetrable subject, the combination of hydrodynamics, already a very difficult subject by itself, and the effects on this of electromagnetic fields: magnetohydrodynamics or MHD. He had seen that one approximation made it possible to think about it in simple terms, but only in conditions when the scale was large and the electrical conductivity was high: conditions that were met in many or most astronomical applications. This was simply the approximation that in such circumstances the magnetic fields would be convected around with the fluid. Faraday's useful fiction of the lines of force of a magnetic field assumed a new meaning: the particles of the fluid that were once on one line of force would continue to be on that common line, to whatever place and shape it came to be convected. Of course the magnetic forces would affect the fluid motions, and that combination still left us problems of great complexity; but at least these were much simpler than a step-by-step, self-consistent calculation of the flow, under the interacting hydrodynamic and electromagnetic forces would have been.

I understood that these were no more than helpful approximations, but I was prepared to use them wherever they were applicable. Many others, especially those with a puritanical mindset like Sidney Chapman, rejected these out of hand. I had many discussions with Chapman on this subject, and he insisted that nothing short of a complete, self-consistent calculation could ever be used to solve a problem in magnetohydrodynamics. "You might be misled terribly by such approximations" was his response, and he would then point out errors that Alfvén had (unquestionably) made in his writings.

Alfvén himself did not always apply his own approximations to problems he was trying to solve. On one of my visits to Stockholm he demonstrated to me a large vacuum system in which an electron beam was directed at a magnetized sphere. This was to represent the Earth under the influence of a beam of electrons from the Sun. I remember questioning him why he thought of an electron beam rather than of magnetized clouds emanating from the Sun. At that time, in the early 50's, he still resisted this viewpoint.

MHD shocks

In 1953 there was a symposium, organized by the International Astronomical Union, entitled "Gas Dynamics of Cosmic Clouds" [*I.A.U. Symposium number 2, July 1953, proceedings published by North Holland publishing Company, 1955*]. It was a splendid occasion, and really the first meeting in which MHD in an astronomical setting was taken seriously; but nevertheless plain gas dynamics of low density gases seemed to have been the mainstream of the discussions.

It was at that meeting that I made the acquaintance of many of the people with whom I had much contact in later years. There was the German contingent consisting of L. Biermann, A. Schlüter, R. Lüst, and also several persons from the U.S., including Arthur Kantrowitz.

Kantrowitz gave a major talk on shock waves, both on the experimental results and the theory. All this was pretty new stuff at the time, much of it spurred on by the U.S. rocket development program. The maintenance of sharp shock fronts with a thickness of only a few mean free paths in the medium was of particular interest to me, because I had puzzled a lot about one phenomenon in the solar-terrestrial relationship: the phenomenon of the abrupt beginning of magnetic disturbances, disturbances that then continued for many hours or days. These "sudden commencements", as they were called, represented a rise in the strength of the Earth's field from the preceding undisturbed state, taking as little as two minutes, and then followed by the "main phase" of the storm, which represented erratic fluctuations but mainly a diminution of the surface field. In many cases the event at the Sun that would give rise to such a magnetic storm was well recognized and so the time delay to the onset of the storm was known to be generally between one and three days. How could something travel here from the Sun in 1,500 minutes and have an abrupt beginning in two minutes? One could not imagine that the surrounds of the Earth experienced some critical phenomenon so that a gradual rise in the intensity of an arriving stream would have no effect, and then, suddenly, at a certain intensity, have a large effect. One therefore had to ascribe this to a property of the stream, that it had maintained a sharply collimated front whose thickness was no more than roughly one part in 1,500 of the Earth-Sun distance. Could this be due to a shock wave phenomenon similar to those Kantrowitz had described? Obviously not, because any estimate of the gas density in the solar system space would make the mean free path equal to or larger than the entire Earth-Sun distance, and we would need 5 mean free paths for the collimation of a shock wave.

But of course this consideration was in the regime of low density gas dynamics, not in the MHD regime. I concluded that this must be making the enormous difference required to make for a collimation of a shock wave to a thickness roughly 7,500 times narrower than the plain gas dynamical shock would have been. Could the presence of magnetic fields in space really change the phenomenon by a factor of 7,500? I concluded that was the only acceptable explanation for the sudden commencements. (The previous literature had described these but had neither offered an explanation nor discussed that this represented a major difficulty.) I presented this explanation at the conference, and stated, in opposition to criticism: "In considering the interaction between the stream and the residual gas one must not restrict oneself to the collision cross-section of neutral particles, but one has to consider the much stronger electromagnetic interactions that may occur between the two ionized gases" [*Gold*, 1955, Discussion Section p.104]. While this was not accepted very readily by many of the participants, Arthur Kantrowitz was much taken with the notion of an MHD shock wave, and he organized later a series of experiments to demonstrate the existence of such waves in the laboratory. Now of course MHD shock waves are regarded not only as the explanation of the problem I was addressing, but they enter into every phase of cosmical gas dynamics.

Solar Flares

The origin of gigantic explosions in the solar atmosphere, known as "solar flares" was a major puzzle. What was the origin of the sudden large energy releases that appeared to come from a level called the chromosphere, a level at which the thermal energy content was far too small to account for them? Suggestions had been made that perhaps some nuclear processes were responsible, but on closer analysis one could see that the density was far too small to get into any such regime. What then could be the explanation?

It was known that solar flares occurred in the surroundings of strongly magnetized regions, recognized as Sunspots. The sudden event of a flare could be identified by a brightening up in a Sunspot region of the H α band of hydrogen, and both the total energy and the peak temperature could be estimated from such observations. No doubt a much larger amount of energy had suddenly become available than had previously been identified to reside there. The only possibility seemed to be magnetic energy, and this suggestion would also account for the proximity of flares to the strongly magnetized Sunspots.

Hoyle and I published a paper based on these considerations [*Gold and Hoyle,* 1960] in which we selected the simplest configuration in which such an energy release from the magnetic fields could arise. The essential feature was that the conducting gas of the atmosphere could carry currents that would increase the energy content there beyond the value given by the "potential field" based on the flux emanating from the photosphere below. (The potential field is the field of minimum energy consistent with the boundary conditions below). This "non-potential" component of the field would derive its energy from the forceful motions in the photosphere or below, which will incessantly tangle and distend the fieldlines in the spaces above where the density and the hydrodynamic forces are too low to resist this. This implies that currents will be generated there, that can be annihilated or diminished, reducing the energy content of the field possibly even down to the potential field value. In particular sudden instabilities of the field could cause such annihilation, and release the energy as heat or as electric fields that would in turn cause particles to be accelerated to high energies. While we restricted ourselves to one case of particular geometrical simplicity, we believed that many more complex configurations would have similar results.

In recent times some doubts have been expressed that solar flares are the major cause of magnetic storms on the Earth [*Gosling,* 1993]. There are indeed other phenomena on the Sun that can also have similar results at the Earth, and possibly these are of a similar nature but occur at a higher level, and at the lower density they then emit much less light and are not seen in Hα [*Gold,* 1962; *Gold* 1959 d]. But there can be no question that many magnetic storms are indeed caused by the identified flares, as the temporal relationship has often been clearly observed. (In the early '50's I was at the Royal Observatory, and my duties included the supervision of the solar and the magnetic observatories. I observed many flares, in some cases followed within less than one hour by the arrival of high energy particles, and then followed within one to three days by the sudden commencement and the magnetic storm.)

The Heating of the Corona

The high temperature of the tenuous outer atmosphere of the Sun, the solar Corona, and the resulting constant evaporation [*Parker,* 1958] that we call the solar wind, has in my view a similar explanation. [*Gold,* 1964; *Gold,* 1965] But here we are dealing in most areas with much weaker fields, and we are not seeing large energy releases in major instabilities. Instead there appears to be a persistent energy source that provides the heating to the million degree temperature and the energetic evaporation of this atmosphere [*Biermann*, 1952. *Parker*, 1958]. I attributed this also to the destruction of non-potential components of the magnetic fields, but now not just of the strong Sunspot fields, but of the weaker general field of the Sun, constantly and everywhere moved by the violent convection cells that we see as the granulation of the photosphere. These photospheric motions must generate currents in the fieldlines above, and thus provide a constant supply of non-potential field energy there. "Mini flares", too small to see, will be the result, since it is not conceivable that the magnetic energy content of the atmosphere would increase indefinitely with time, nor that the motions below would contrive to disentangle the fields they had previously entangled.

The Van Allen Belts

The entire discussion of the magnetosphere and its properties was changed in 1958 with the earliest spacecraft discovery by Van Allen and his skillful team in the U.S. [*Van Allen,* 1959]:and later confirmed by Vernov and Chudakov in the USSR [*Vernov and Chudakov,* 1960], of the belts of high energy particles that were stored in orbits around the Earth. Van Allen realized readily that the particles would be on captive orbits in the magnetic field, and that such orbits would spiral around the Earth's field lines, oscillate rapidly in latitude between reflection points in the converging fields, and precess slowly around the Earth in longitude. Such orbits were soon calculated in detail and were found to have long term stability. In fact I found it surprising that no one had made such calculations before the discovery, and predicted that this phenomenon was possible. The well known diminution of the field at the surface, during the main phase of a magnetic storm required a current ring around the Earth somewhere above the surface, and that had been amply discussed; but neither the origin nor the stability of such a ring ever had a thorough discussion. That ring would have to have a much larger flux of particles than that observed now in the newly discovered Radiation Belts. But one may well consider that the dominant contribution may have come from a lower particle energy range than was detectable by the instruments that had been flown. I published a paper in Nature describing these orbits and giving a few deductions one could make about them. [*Gold, T.* , 1959a.]

If these orbits were so stable, then this immediately posed the question how they could ever be supplied with high energy particles. If particles cannot get out of the storage orbits, they also cannot get in. The mathematicians who had made the detailed calculations, seemed to think that the only possible entry would be by uncharged energetic particles such as neutrons from the Sun, some of which would decay into protons and electrons in the region of the storage orbits. This explanation was widely accepted at first, but I did not consider it adequate for quantitative reasons. I had been concerned with researches on neutron fluxes from the Sun, and when it became clear that the radiation belt particle flux suffered temporal variations, the solar energetic neutron flux was insufficient to account for these by a large factor.

Motions in the Magnetosphere of the Earth

I came to realize that the mathematical stability calculations had been based on an erroneous assumption. It was implied that the lines of force were anchored in the Earth, and therefore that they could distort, but not move their attachment points (their "feet") relative to the solid Earth.

Alfvén's simplification of MHD stated that in suitable and defined circumstances, his approximation meant that low energy particles that were on a common line of force would stay on that same line and in fact be convected around with the field. What had not been discussed in the Alfvén approximations, was the effect of a layer of insulating material running across the magnetic field lines. I believe that some people thought that if such a layer were sufficiently thin, it would have no effect. Thinking more about it, I realized that this was an error, and if the insulating layer interrupted currents that were flowing along the field lines, the whole situation was completely changed. The atmosphere is such an insulating layer, and one now had to puzzle out what effects this would have.

Within the Alfvén approximations of field lines moving with the gas, one could ascribe an "identity" to each field line since one can ascribe an identity to the particle fluxes that were on that field line. But when an insulating layer cuts across the field lines, these lines lose their identity. The field lines still have to be continuous from below to above the insulating layer, but the connections can now be freely changed so that the field emanating from the Earth cannot hold any particular tube of force in the same location. Thus, the tube of force and the particle fluxes that it contains can move around as they may be subjected to other forces, but they must merely obey the condition that in each area the total flux must be the same above and below that layer.

This consideration then implied that the stability calculations for the radiation belt particles were not applicable. Any one tube of force with the particles in it may be moved from one place to another and thus any one flux tube could be convected together with its particle content into a position which the individual charged particles could not have reached. The entire force field and its particle content was subject to convection that could bring outer tubes emanating at high latitudes into the position previously occupied by lower tubes and vice versa. Particles that could enter the outermost tubes of force might then later be convected into regions to which they would not have had access individually.

The forces that will cause such convection can be discussed, and they may arise either from hydrodynamic forces in the outer ionosphere or from pressure differences between neighboring tubes that have been distended to a different degree by the hot gas or energetic particle loading they may contain.

I published this in a paper entitled, "Motions in the Magnetosphere of the Earth" [*Gold,* 1959 c.] and it seemed to me that this largely resolved the stability problem and did not require now that only decay products of neutral particles would inhabit these regions. The magnetic storm onslaughts on the Earth's field could now be a source of resupply of the radiations belts.

The title of this paper introduced, as a new addition to the language, the word "magnetosphere". There was much opposition to this in the first place, but I insisted that we must have a name for the region in which the Earth's magnetic field is the controlling factor. It did not take long before this word was part of the English language and indeed translations of it became a part of many other languages.

The sequence of phenomena observed in a magnetic storm and in Van Allen's radiation belts now became much clearer [*Gold,* 1962]. A magnetized high speed cloud from the Sun would arrive here with a shock front, as it plowed its way through any gas that was in the Earth-Sun space and relative to which the cloud was moving at supersonic speed. The width of this shock front determined by the magnetic interactions was narrow enough to flow over the Earth's magnetosphere in a matter of one or two minutes. At that stage, it implied a compression of the external field and therefore an increase of the fieldstrength at the Earth. Following that, the outer field lines would become loaded with energetic particles and that would imply a distention of the field and therefore a decrease of the strength of the field during the main phase of the magnetic storm. Much interchange of tubes of force would take place as they were loaded differently with energetic particles. All this became much clearer still in the picture devised by Axford and Hines *[Axford and Hines, 1962]* who then considered all this together with the rotation of the Earth, and provided a clear explanation of the process of "folding in" of the hot and fast moving solar gas into the storage regions around the Earth.

The Solar Cosmic Rays

I had also directed my attention to another aspect of the solar outbursts, namely the high energy particles, called solar cosmic rays, that sometimes arrive at the Earth within minutes or hours after a solar flare has been observed, and long before the arrival of a storm cloud, often but not always associated with such particle bursts. The particles are mostly protons in an energy range of 10 Mev to 1 Gev, and would be strongly guided by the magnetic fields in the Sun-Earth region. Observations of these particles at the Earth allowed one to make deductions about the configurations of magnetic fields in the Earth-Sun space and beyond. For these energetic particles to arrive so quickly, one has to suppose that these fields had been drawn out previously from the Sun, to encompass the domain of the Earth. It often happens that the same event at the Sun that caused the high energy burst also causes a gas outburst, showing up 1 to 3 days later as a magnetic storm.

The most remarkable high energy particle burst occurred on February 23, 1956, following a large solar flare, with a delay of only approximately 20 minutes. Nothing of this magnitude or short time delay had been observed in the six prior years of observation of such phenomena. Working at the Royal Observatory in England, and being interested in solar-terrestrial relations of this nature, I had arranged for the construction of a monitoring instrument specifically for the observation of such solar outbursts. A small house had been constructed for it, so that the instrument would be free from any effects that could change the counting rate, since the solar events are only superposed, so we thought, as a small effect on the galactic cosmic radiation. Our instrument was the most sensitive and best protected from interfering effects that existed in the world at the time; it had been in operation for three months only when this giant event occurred. Optical observations of the Sun, various radio observations, and several other cosmic ray observations all confirmed the great magnitude of this event, and we arranged for these to be published as a collection in the Journal of Atmospheric and Terrestrial Physics. [*Gold and Palmer,* 1956].

Mr. Palmer and I were of course proud that our new instrument had brought in so quickly such a remarkable result, and had recorded it with excellent precision. However the new Astronomer Royal, the head of the observatory, Dr. R. R. Woolley, appointed less than two months before, was not pleased. He instructed me to have the instrument and its building dismantled forthwith, since "such observations were not part of astronomy and had no place in the Royal Observatory". (My subsequent move to the United States was related to this and other similar commands.)

The following conclusions could be drawn from an observation of these energetic particles: [*Gold*, 1959 b.; *Gold*, 1959 c.; *Gold*, 1959 d.; *Gold*, 1961; *Gold*, 1963]

1.) The Earth is a large part of the time in magnetic fields that connect with the Sun. It is only rarely, if ever, embedded in galactic fields that are unconnected with the Sun.

2.) The high energy particles that arrive first from a solar outburst show a small angular spread in their direction of flight, and that direction is consistent with an origin at the Sun. However, at later times this flux slowly changes to a more isotropic distribution of velocities. The field must thus be in the shape of loops, drawn out from the Sun so that the particles filling such a loop can eventually arrive even from anti-solar directions.

3.) The interplanetary magnetic fields have irregularities that cause the captive particles to develop finally an isotropic velocity distribution.

4.) The magnetic fields drawn out by gas eruptions retain their connection with the Sun for a day or more and probably often for as much as five days.

5.) The magnetic total pole-strength of the Sun cannot increase indefinitely with time. (Pole-strength is here defined as the total number of field lines emanating from the surface, irrespective of the sign.) A mechanism has to be at work that reduces this pole-strength, on an average by as much as it is increased by the individual outbursts. Loops of field blown out must eventually be cut off from the Sun, and then become independent clouds progressing into space. This cutting-off process is attributed to a reconnection near the Sun across the neutral sheet that separates the two sides of the loops. [*Gold* and *Palmer*, 1956, *Gold*, 1959, 1961, 1963]

There was another set of observations that related to these phenomena. These were the observations by Forbush, who had noted that there were times of a small decrease in the flux of galactic cosmic rays seen at the Earth, decreases that had a relationship with solar outbursts. Of course outbursts that made such field configurations as would favor solar energetic particles to get to the Earth would also tend to exclude particles from external sources. However, the effect is only small, mainly because the galactic cosmic rays are mostly of higher energy, and therefore fields that can make a region of containment of solar cosmic rays can still be penetrated by much of the galactic cosmic ray flux.

Whistlers

R. A. Helliwell

*Space, Telecommunications and Radioscience Laboratory,
Stanford University, Stanford, California*

My whistler work began quite unexpectedly during the course of radio studies at Stanford University in 1949. Observations of radio atmospherics from lightning led to the identification of dispersed, falling-tone signals that proved to be whistlers. Stimulated by *Storey*'s [1953] explanation of whistler propgation along paths extending to high altitudes within the Earth's magnetic field, observations at College, Alaska, led to the discovery of the "nose" whistler which provided the key information on the path of propagation. The nose whistler was the tool for all subsequent ground based studies of the thermal plasma distribution in the magnetosphere as well as the interpretation of triggered emissions.

A RADIO BACKGROUND

In this paper on whistlers, (see Figure 1 for a sketch of whistler paths) I will describe how I unexpectedly entered the field and how the whistler research program at Stanford developed in cooperation and in competition with others. Then I show how this study of an ancient natural phenomenon turned into a probe for remote sensing of the magnetosphere, and then led the way to a new man-made method of stimulating non-linear wave particle interactions in the radiation belts. Like many new scientific fields, the study of whistlers has had several unexpected turns, both experimental and theoretical, with more to come I am sure.

I entered Stanford in 1938, having long been interested in electricity and radio, getting a ham radio operator's license in 1934 while in Palo Alto High School. Making radio contacts was fun but I got even more pleasure just building the equipment and making it work (not always successfully). One memorable experiment was constructing a Lecher transmission line on which, with the aid of a neon bulb, I could see the standing waves excited by my transmitter. It was only several years later, as an electrical engineering student at Stanford in a course by Prof. Hugh Skilling, that I was introduced to the theory and applications of transmission lines, which further deepened my interest in radio waves. In 1939, Mike Villard, then a graduate student in E.E., started a modest program on ionospheric sounding at HF, and I was lucky enough to be taken on as an assistant. I operated a 3 frequency sounder in the small furnace room located in the basement of the Ryan High Voltage Laboratory, where my companions were assorted tarantulas and black widow spiders.

With the increasing political unrest in the world at large, there was a demand for more accurate forecasting of high-frequency propagation conditions, the principal means at that time for long distance communications, both military and civilian. The key instrument was the sweep-frequency sounder, a new innovative version of which Mike Villard was charged with building at Stanford. Again, I was ready and eager to help. In addition to operating and maintaining the sounder, I got a part time job (about 1939) with the Hoover Library on War, Revolution and Peace (now the Hoover Institution) recording Japanese propaganda from Radio Saigon, under the direction of Mike Villard. We used the wax cylinders of a standard dictating machine coupled acoustically to a Hammarlund short wave receiver located in the Ham Shack behind the Ryan High Voltage Laboratory. These recordings I delivered to the Hoover Library where

Discovery of the Magnetosphere
History of Geophysics Volume 7

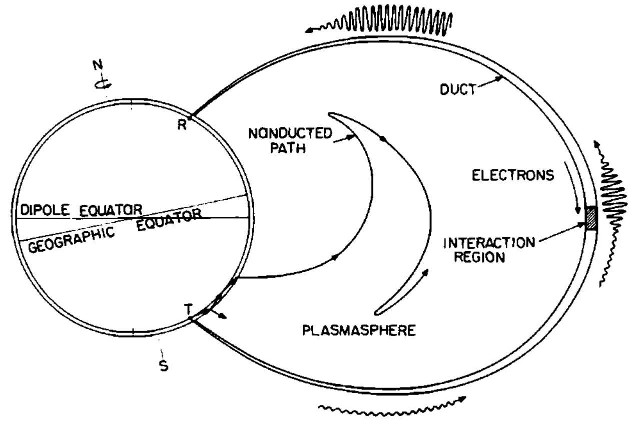


Figure 1. Typical ducted and non-ducted paths of whistler-mode signals within the plasmasphere, together with sketches of the waveforms of ducted WM signals before, during and after their non-linear interaction with cyclotron-resonant electrons in the interaction region at the dipole equator. Each traverse of the duct is called a "hop". A two-hop whistler then is one that has traveled to the opposite hemisphere and been reflected back to the starting point.

they were transcribed under the direction of Inez G. Richardson.

After graduating in June 1942, I had an opportunity to work for the General Electric Co. in Schenectady, but an interesting research project for the U.S. Navy attracted me to remain on the Stanford campus as a graduate student. The work involved using an HF Direction Finder to study the errors in the bearings of short wave transmitters, particularly those in Europe. One purpose was to help military pilots recognize faulty bearings, a frequently occurring event on high latitude routes to Europe. What we found was that the signals of BBC stations, for example, which would normally arrive from the Northeast, along the great circle path from London, would during magnetic disturbances often come from the Southeast. This anomaly resulted from back scattering of the radio signal from the ocean surface in the South Atlantic and became a source of confusion to pilots flying between Europe and the U.S.

In the meantime, I had completed an M.A. (1943) and was working towards an Engineer's Degree (1944) when I was classified 1-A in the Draft. I had applied for a deferment to continue with the research which was sponsored by the Office of Scientific Research and Development (OSRD), but the papers were apparently misplaced by the Draft Board. As the deadline approached, I rushed my E.E. thesis to completion and packed my bags, prepared to bid my family adieu. In the mean time the draft board found the deferment papers and decided that I would be more useful to the war effort by remaining at Stanford doing wave propagation research.

As the war drew to a close the need for direction finder (DF) work became less urgent but the wave propagation studies with our newly acquired C-3 ionospheric sounder, sponsored by the National Bureau Standards, were turning up interesting questions. For example, I wondered what the ionosphere would look like if the sounder frequency were reduced below 1 MHz, the lower limit of the C-3. By this time, I was attracted to the idea of a career in teaching and research and decided to pursue the Ph.D. degree in E.E. My interest in frequencies below 1 MHz coupled with the presence of the Ryan lab with its high voltage facilities led me to the idea of building a low frequency spark transmitter to operate at 100 KHz using a 200 kv charged capacitor as a power source. The antenna was to be a leftover section of a 2 inch diameter power line cable that had been designed to carry power from Hoover Dam on the Colorado River to Los Angeles, California. I built a power supply consisting of a nine-stage series-connected stack of war-surplus 20 kV oil-filled capacitors. Each capacitor was charged through a half-wave vacuum rectifier. Their filaments were heated in a series resonant circuit driven at ~500 kHz. The spark transmitter became operational in the spring of 1947.

The repetition period of the pulses was set at ~1.5 s by mechanically rotating one sphere of a spark gap. At the receiver, located about 3 miles away in a hut on old Searsville Road, the scope time base was triggered by the strong ground pulse from the transmitter. To keep the ground pulse from overloading the receiver, a converted war surplus standard (1.85 MHz) Loran receiver was fed from a home made balanced dipole antenna whose orientation could be adjusted by remote control to reduce the ground pulse without affecting the strength of the downcoming ionospheric echoes.

To operate such an unconventional transmitter, I needed a license from the FCC, whose regulations stated that an experimental transmitter such as I was proposing must have a bandwidth less than a certain value. However, nowhere in the regulations could I find a definition of bandwidth for this purpose. So I made my own definition such that the calculated sidebands of the exponentially damped pulse fell within the required range, which, not surprisingly turned out to be very much less than the -3 dB point of the actual spectrum. The license was duly granted, but in fact the spectrum was so broad that a 30 kHz wide receiver could detect ionospheric echoes on virtually any clear channel up to several megahertz. Some of the most useful data came from the next higher-order mode at 356 kHz (the higher order modes were anharmonic because of the series inductance needed to tune the antenna to 100 kHz) A consequence of this wide spectrum was a potential for interference to communication services. However by keeping the pulse repetition frequency (PRF) low (< 1/s) and performing operations late at night, there were no serious complaints.

One unexpected problem arose during unattended automatic operation late one night when the feed-through insulator on the roof of the transmitting building began flashing over (probably due to dust on the insulator's porcelain surfaces) producing a sound like that of a high-powered rifle. Alarmed neighbors called the Stanford police who arrived well armed, but fortunately no real shots were fired. Since the Ryan Lab was involved they correctly surmised that the E.E. Department might be able to explain the discharges. I took corrective action, cleaning the insulators, and the work went on.

EARLY WHISTLER EXPERIMENTS

As the presence of multiple-layer fine structure in the D and E regions became apparent in the fixed frequency records [*Helliwell et. al.*, 1951], a natural question was: what happens at frequencies even lower than 100 kHz, say 10 kHz? Because of space limitations, the very large antenna required at 10 kHz was not feasible on the campus; instead we took advantage of lightning impulses, called sferics, as very low frequency VLF sources of oblique incidence sounding. At VLF, the nighttime ionosphere has a reflection coefficient approaching one (on magnetically undisturbed days) allowing a train of up to a hundred or more multi-hop pulses (called tweeks) to be detected from a single lightning discharge (16mm film recordings of tweek waveforms were collected using a Tektronics 'scope). Since the time spacing between successive echoes depended on the height of the reflecting layer and the distance of the source, it was easy to extract both parameters from a single tweek [*P. Smith*, 1953].

It was in the course of these nighttime tweek recording sessions that Jack Mallinckrodt heard long (1-3 s) whistlers of descending frequency mixed in with the tweeks. When I saw Jack the day after one of these sessions, he reported that the recording session went well with many good tweek waveforms having been recorded. And, then he added, rather as an afterthought, that he heard on the monitor loudspeaker some descending tones, which we came to know as whistlers, mixed in with the various pops, crackles, and tweeks from lightning. Neither of us had heard of whistlers and I wondered whether the cause might be an unstable amplifier. Jack (who is an excellent musician) felt that the whistlers were of natural rather than instrumental origin. After his next observing session Jack reported more whistlers! At this point, I suggested to Jack that perhaps he had been working too hard and needed a vacation. He stoutly rejected this hypothesis, convincing me that his whistlers were real phenomena, whereupon I suggested that I join him at the next recording session on old Searsville Road. We started, as I recall, about 8:00 p.m., after sunset, so as to reduce the effects of D region absorption, that tended to smear out the waveform of the tweeks, making the interpretation difficult.

The tweeks that night were good and we paid so much attention to the quality of the wave forms that I had almost forgotten the purpose of my visit. Then it happened--a clear descending whistler came from the loudspeaker, lasting 2-3 seconds, and I cried to Jack, "Is that what you've been hearing?", and he said, "Yes, that's it!" After hearing a few more whistlers, I became a confirmed believer and asked Jack to go to the Stanford library to determine whether anyone had reported similar observations. And, they had. There were several references to whistlers [see *Helliwell*, 1965], of which two were particularly significant. The first, by H. Barkhausen [*Barkhausen*, 1919], working for the German Army in World War I, described how he listened in on Allied telephone conversations at the front lines using two widely spaced conducting ground probes connected to an audio amplifier. Often, usually in the afternoon, he heard descending audio tones, while he was engaged in eavesdropping. At first he assumed his amplifier was oscillating, but the tests he performed confirmed its stability, leading to the conclusion that he had observed a new natural phenomenon. In a later paper [*Barkhausen*, 1930] he gave an explanation of the whistler based solely on what he remembered hearing, there being no wire or tape recorders yet available. His explanation was the same as the mechanism of the tweek which has a musical sound of rapidly descending pitch ending near 1600 Hz. But without a quantitative analysis of the wave form or a dynamic spectrum analysis, the distinction between tweeks and whistlers can be somewhat fuzzy, particularly when long tweeks are mixed with short whistlers. (One observes that the ear coupled with a fertile imagination can sometimes bring false closure between observations and theory).

Later, T. L. Eckersley [*Eckersley*, 1935] using the new equations of the magneto-ionic theory developed the so-called "Eckersley approximation" for the dispersion of a lightning pulse in the Earth's ionospheric plasma, which provided a good fit to the *f-t* curves observed in England. However, he could not find a plausible model for the path of propagation since it was believed at that time that the atmosphere ended just above the F layer. This was the state of affairs when we made our serendipitous whistler observations at Stanford. After we uncovered the literature on whistlers, I became very excited about the possibilities for further research since there were so many unanswered questions. I was hoping that the answers might lead us to new tools for research on the upper atmosphere.

About this time (circa 1950), I learned of Owen Storey's observations of whistlers, carried out at Cambridge University, U.K. [*Storey*, 1953]. He showed experimental evidence that "long" whistlers and their even-order echoes originated in lightning impulses from points within a few thousand km of Cambridge whereas *short* whistlers had odd-order multi-hop delays pointing to lightning sources in the southern hemisphere. Using the Eckersley approximation, Storey showed by ray tracing that one could get a good fit to the data if the whistler path was assumed to be approximately aligned with the Earth's static magnetic field out to 3-4 Earth radii. At this distance his model required an electron density of ~400/cc, a value in stark contradiction to the then widely accepted notion of an atmosphere terminating at an altitude of a few thousand km. This disagreement was highlighted at the URSI General Assembly in 1952 at Sydney, Australia, where I heard J.A. Ratcliffe, Storey's Ph.D. advisor at Cambridge, report Storey's density prediction with the implication that it probably was wrong because it was at variance with current ideas about the vertical extent of the atmosphere. As it turned out Storey's density was in good agreement with later satellite *in situ* measurements and "nose" whistler measurements. Thus it was that Storey discovered the *thermal* magnetosphere well ahead of the discovery of the radiation belts by Van Allen in 1958.

In order to check and extend Storey's findings, it seemed to me that we would need data from locations of much higher latitude than either Cambridge or Stanford. Accordingly, I proposed (~1954) to Chris Elvey, then Director of the Geophysical Institute of Alaska (GIA) that we carry out a joint search for whistlers at College, Alaska, (L~6), where L is the geocentric distance in earth radii of the top of the field line, assuming a dipole field model. We would supply a VLF receiver and the Institute would supply the low noise site and operating personnel. In 1955, some swishy-sounding whistlers were recorded by Joe Pope at College and the tapes were sent to Stanford for analysis using our newly acquired 0-8 kHz spectrum analyzer, called the Vibralyzer (an outgrowth of a device developed for speech analysis by R. K. Potter [*Potter*, 1951] of Bell Laboratories, later to be marketed by the Kay Electric Co., as the Sonograph).

As the Stanford-GIA field work on whistlers was getting started, I studied a photocopy of Storey's thesis which Storey had kindly supplied to me. Using a homemade filter bank driving a set of neon bulbs photographed on moving photographic film, Storey was able to plot the dispersion curve in Eckersley coordinates ($f^{-1/2}$ versus time) for each whistler. In the Eckersley approximation this plot is a straight line whose origin is the instant of the lightning source. Storey's data, shown in his thesis with error bars, fitted the straight line very well except for some higher frequency points which usually showed a little extra delay than the Eckersley Law predicted. Storey himself called attention to this deviation but offered no explanation since it had little effect on his results. Because we could see that the Eckersley approximation, based on the assumption $f<<f_H$, where f_H=electron gyrofrequency, would fail at very high latitudes, such as College, I recruited in 1955 a young undergraduate physics student, T.F. Bell, to carry out a more accurate calculation using Stanford's IBM 650 card programmed computer.

The College tape arrived by mail just as Bell was finishing his calculations. I personally made spectrograms of the tape recorded whistlers from College using the Vibralyzer which was located in the same room of the Electronics Research Laboratory (ERL) where Bell was hand plotting the dispersion curves from the IBM 650 printout. The first swishy whistler spectrum I made was quite surprising because it didn't look anything like the mid-latitude whistlers that were recorded at Cambridge and Stanford. At first glance (see Fig. 2), my impression was of a noise band whose center frequency was slowly decreasing with time, in keeping with the sound. I was puzzled and annoyed because the spectrogram seemed to make no sense. At the same time Bell had arrived at his desk with plots of frequency versus group delay (f-t curves) from the computer. His hand-plotted curves were parabolic in shape, not monotonically decreasing in frequency with time as in the Eckersley approximation. Furthermore, there was a frequency (which we later called the nose frequency) of minimum time delay where the upper and lower (Eckersley-branch) branches of the whistler dispersion curve were joined, in good qualitative agreement with the new data from Alaska. As Figure 2 shows, the swishy sounding whistler consists of multiple parabolas whose nose frequencies fall with the nose time delay. Soon after this startling finding, a new approximation was derived from the magneto-ionic theory (by Bob Smith, I believe) for the nose whistler. A discovery paper was presented at the Gainesville, Florida, USNC/URSI meeting in December 15, 16, and 17, 1955 and published in JGR [*Helliwell et al.*, 1956]. Those new results provided a theoretical framework for the following IGY program on whistlers and VLF emissions, 1957-58.

About this time I was invited by Alan Shapley of the NBS Boulder Laboratories to travel to the Antarctic in the Austral summer to learn whether there might be opportunities for whistler work there during the IGY. Having long been a fan of Admiral Byrd's Antarctic explorations, it didn't take me long to accept the invitation and to assemble a portable whistler receiver. My trip via C124 to McMurdo and DC 3 to Little America V was exciting and instructive but inconclusive because of the high station noise level and lack of sensitivity of my portable receiver. But when all factors were considered it seemed that the Antarctic should be an excellent place for whistler research and in any case we needed to know the level and kind of whistler-mode (WM) activity in both polar regions. On my return from Antarctica we put together a more sensitive system which was sent to Byrd station, beginning a long term successful effort to discover Polar region phenomena at VLF [*Helliwell*, 1965; *Helliwell and Katsufrakis*, 1978].

A VLF receiver was also set up at Eights Station where in 1963 Mike Trimpi observed the so-called Trimpi Effect, consisting of a perturbation (either positive or negative) of the signal strengh of a VLF transmitter (eg. NSS, NAA, etc.) in association with a northern hemisphere whistler. At first I tried, unsuccessfully, to explain the effect by a direct alteration of the ionosphere near the transmitter by the causative sferic. Later we showed that the effect was produced by cyclotron resonant electrons scattered into the

were presented in person by their authors at the General Assembly of URSI held in 1963 in Japan.

Other important findings based on the nose whistler include the detection of the cross L electric field E through the measurement of the $\mathbf{ExB_0}$ drift of whistler ducts [*Carpenter and Stone*, 1967] which provided additional confirmation of the duct hypothesis while at the same time providing a new ground-based tool for studying the circulation patterns of the magnetospheric plasma as predicted by theory [*Axford and Hines*, 1961]. A further application of whistlers was made by Chung Park [*Park*, 1970], who measured the filling from the ionosphere of plasma tubes after they had been emptied during a magnetic storm, the filling process requiring about one week. Using the nose whistler method, Park showed for the first time that the nighttime F region was maintained primarily by downward diffusion from the already-filled plasmasphere.

CONTROLLED SOURCES OF WHISTLER MODE SIGNALS

As whistlers and ducting became relatively well understood, we asked ourselves whether VLF communication stations (e.g. NAA on 14.7 kHz) might excite detectable whistler mode signals. If so, they would have the advantage of fixed location, known radiated power (as high as a Megawatt) and known modulation pattern. Furthermore, there was virtually no transmitter cost to the experimenter. Since such echoes had not yet been reported their observation could provide further confirmation of Storey's hypothesis. With help from the Office of Naval Research, we arranged in 1957 for a special keying sequence to be transmitted by station NSS at Annapolis, Md. [*Helliwell and Gehrels*, 1958]. Observations were to be made near the transmitter's conjugate point which was close to Cape Horn.

We were lucky to obtain the services of Ernst Gehrels, a resourceful Stanford graduate student and one of Mike Villard's Ph.D. candidates who was attracted to what turned out to be a true adventure. He was to take a VLF receiver, tape recorder, loop antennas and, most important, a wind-powered generator to charge a battery, since it was expected that suitable locations would probably not have reliable commercial power available. The observation equipment was to travel with Gehrels to Santiago, Chile and then on to Punta Arenas for installation at a suitable location. Gehrels' first challenge came as he passed through customs in Santiago. His equipment could not be allowed in because it did not satisfy certain unspecified requirements. Through astute observation Gehrels discovered that by paying five dollars his passage through customs was assured. Upon arrival in Punta Arenas, Gehrels persuaded the Captain of a Chilean freighter that was heading south into Antarctic waters to take him and his equipment along and drop him off at a remote island near the conjugate point. However, bad weather prevented a landing and so Gehrels decided to leave the ship at a more northerly location off Cape Horn where he set up his station.

However we had not realized that the wind speed in that area of the world averaged about 60 mph and his Sears Roebuck wind generator was designed for mid U.S.A. farms where the wind speed was less. Never daunted, Gehrels got the station working, but before he could record any data, the generator was literally blown away. A less resourceful person might have "thrown in the towel" at that point, but Gehrels perservered, finding a reasonable site with access to power where he made some highly successful recordings, the first of their kind, of whistler mode echoes from NSS on 15.5 kHz [*Helliwell and Gehrels*, 1958].

The measured magnetospheric delays (~0.75 s) of the NSS pulses were consistent with whistlers observed on the same paths whose sources were lightning discharges on the East Coast of the U.S. These new findings gave further support to Storey's hypothesis while at the same time demonstrating that VLF transmitters could be used as sources for sounding the magnetosphere using receivers on the ground and on satellites. Later work included several other VLF transmitters such as NLK/NPG at 18.6 kHz, in Jim Creek, WA; NAA at 14.7 kHz, Cutler Maine; NPM at 19.8 kHz, Oahu. Signals were received at Ushuaia, Argentina; Green Bank, W. Virginia (a two-hop signal from NSS); Wellington, N.Z. (one-hop from NLK); Stanford, CA. and Seattle, WA, (both two-hop NLK) [*Helliwell*, 1965]. These data revealed the effects of sunrise and sunset in the D region of the ionosphere where the whistler mode signals entered the duct. By measuring the L value of the path with whistlers and the time of onset and cessation of absorption, we could estimate the points of penetration through the ionosphere of the ducted WM echoes.

The combination of whistlers and WM signals from VLF transmitters provided the evidence needed to design experimental magnetospheric sounding transmitters for use on the Antarctic ice sheet. The first was built in the period 1966-1971 near Byrd Station and called Long Wire [*Helliwell and Katsutrakis,* 1978]. Successors at Siple Station L=4.3 were *Zeus*, 1973-1977, and *Jupiter* 1978-1988 [*Helliwell,* 1988], which replaced Zeus after it was crushed by the accumulating ice burden (about five feet per year). These wave injection experiments from Siple Station revealed many new non-linear wave-particle interactions which are discussed in the following section on Triggered VLF Emissions.

An important manmade source of relatively weak WM signal is the harmonic radiation from power grids (usually at a fundamental of 50 or 60 Hz), a subject beyond the scope of this paper [see *Helliwell et. al.*, 1975]. Another controlled source of whistlers is the nuclear explosion, whose details are also beyond the scope of this chapter [*Helliwell and Carpenter*, 1963].

triggered emissions may themselves exceed the parent signal in strength, making this phenomenon of intense interest in magnetospheric physics.

After the discovery of whistler-mode signals from certain VLF communication stations signals, other VLF stations that appear on the broad band whistler tapes were found to contain examples of WM echoes, both one-hop and two-hop. In a 1959 study of Stanford IGY 2-min recordings made every hour at Wellington, N.Z. of Morse code signals from NPG on 18.6 kHz, located at Jim Creek, WA, it was discovered quite by accident, that the WM echoes often triggered discrete VLF emissions. Not only was this the highest frequency yet found to exhibit triggering, but the triggering of risers by NPG was confined almost wholly to the Morse dashes (150 ms duration) with few, if any emissions from the Morse dots (50 ms duration) [*Helliwell*, 1965]. Since the transmitted amplitudes were the same for the dots and dashes the difference in triggering was apparently due simply to the factor of three difference in the length of the dashes compared with the dots. Detailed measurements showed that the triggered riser began 70-130 ms after the start of a dash, a range that exceeded the dot length of 50 ms which confirmed the role of pulse duration. Called the "dot-dash" anomaly, this effect was not present in any known plasma instability and for me was a top priority mystery.

But there was more to come. Similar phenomena were found in 1962 by the U.S.N.S. Eltanin while cruising in South Atlantic waters near the magnetic conjugate point of NAA, which was radiating about 1 Megawatt at 14.7 kHz, well above the radiated power of NPG (~250 kW). We had equipped the Eltanin with a complete whistler station (designed and installed by L.H. (Bud) Rorden), with two large crossed loops, each in the form of a square constructed from 2 inch diameter copper pipes mounted on the foredeck of the ship. Ampex type 350 dual channel analog audio recorders provided about 55 dB dynamic range and an upper cutoff frequency of 30-40 kHz. These favorable characteristics provided good simultaneous data on whistlers and related chorus and hiss as well as the entire VLF band used for communications. The two crossed loops gave some information on signal polarization and direction of arrival of both the earth-ionosphere (e-i) waveguide signals (usually vertically polarized) and the downcoming circularly polarized ducted WM signals. An example of these data is shown in Figure 3b recorded on October 1,1962 [*Helliwell*, 1965, Fig. 7-65], along with the ground waves from all detected VLF transmitters from 14 kHz to 24 kHz as measured at Green Bank, W. Virginia (GR) in the upper panel of part (b) along with the Eltanin observations from 8 -28 kHz plotted as function of time. The slanting lines in (b) connect the source pulses with the received pulses based on a measured one-hop group delay of about 0.7 sec. This set of observations confirmed the earlier NPG data on the dot-dash anomaly but showed a much higher level of activity, possibly due to both the greater power (4:1) and lower frequency (14.7 kHz vs. 18.6 kHz) of NAA relative to NPG.

The Eltanin results provided a strong base for the idea of a controlled VLF wave-injection facility that could experimentally determine the nature of non-linear VLF wave-particle interactions. It would operate in the frequency range where wave-particle interactions (WPI) were commonly observed and would be controllable in frequency, radiated power, modulation characteristics and polarization. No such transmitter existed nor was there any readily accessible location where a reasonably efficient antenna could be erected. Solutions to these problems were later found, resulting in the creation of Siple Station on the Antarctic ice sheet near the average plasmapause location of L~ 4 and where the ice thickness of ~2 km provided an acceptable antenna efficiency of a few percent [*Raghuram et. al.*, 1974]. The first antenna was a 23 km long horizontal dipole about 5 meters above the snow surface, operating in the frequency range, 2 to 6 kHz, of greatest interest for controlled WM experiments. The receiving station was placed close to the Siple station conjugate point in Canada (Roberval, Quebec) for the first transmitter (called Zeus) and Lake Mistissini, Quebec for the second, called Jupiter. Details of the Siple Station program can be found elsewhere [*Helliwell and Katsutrakis*, 1978; *Helliwell*, 1988].

INTERPRETATION OF TRIGGERED EMISSIONS

A fundamental problem was posed by the dot-dash anomaly described above since only coherent input waves (150 ms Morse dashes from keyed continuous wave [CW] transmitters (e.g. NPG and NAA) produced risers. There was no known instability in plasma physics that could explain the effect. To meet this need Neil Brice, who was working on his Ph.D. in Stanford's Radioscience Lab., came up with an ingenious suggestion based on the concept of coherent addition of the radiation from an end-fire array of antennas. *Brice* [1963] argued that a coherent WM signal (with its wave normal aligned with the Earth's magnetic field) would tend to phase-bunch any electrons traveling in the opposite direction at the cyclotron parallel resonance velocity "$v_{//}$" given by $w_H = w + kv_{//}$, where w_H = gyrofrequency, w=angular wave frequency, k=wave number. These phase-bunched electrons would radiate more coherent wave energy which would phase-bunch more electrons, and so on.

Brice's phase-bunching process was based on the assumption of a homogeneous interaction region, whose effective length was not given. Furthermore, the mechanism for producing the change in frequency observed on triggered risers and fallers was not included in Brices's model.

The publication of the *Kennel and Petschek* [1966] model provided the landmark concept of the regulation of the radiation belts through the feedback between a growing

WM wave (assumed to be hiss) and the dumping into the loss cone of the cyclotron resonant electrons which grew the waves. The idea was that the Doppler-shifted cyclotron resonance interactions involved coupling between the transverse **B** field of the WM wave and the electron's perpendicular velocity component v_{perp}; wave growth occurred through the conversion of the v_{perp} energy of the electron to the E_{perp} of the wave which then increased in power according to Maxwell's equations. As v_{perp} dropped, the pitch angle dropped causing some electrons with small pitch angles to fall into the loss cone where they were absorbed in the ionosphere (such absorption may produce light, x-rays, enhanced ionization, and heat).

The difficulty with the KP theory was its inability to describe narrow band signal growth, since KP assumed very small wave-induced changes in the phase of an interacting electron in order to make their analysis tractable. For their waves, assumed to be random noise, (called *hiss*), this assumption seemed reasonable, but, as *Brice* [1963] showed, one could expect much larger phase changes during electron phase-bunching by coherent waves.

With this background, I became mildly obsessed with visualizing the physics of the wave-electron interaction. The key insight came in the middle of the night on Valentine's Day (Feb. 14, 1967) with the visualization of the two helices, one representing the motion of the resonant elecrons and the other the locus of the circularly polarized B field of the wave. In this picture, the electrons were radiating a changing frequency as they moved through the inhomogeneous interaction region (IR), defined as the region in which the phase angle between E_{perp} and v_{perp} remained within $\pm\ \pi/2$. In the following days, weeks (and months) I derived a differential equation for *df/dt* in terms of f and the inhomogeneity variables f_H, f_N and pitch angle which gave the *df/dt* of a triggered emission at the location of the IR. Typical asymptotic values of the emission slope were observed to be 1-2 kHz/s.

To get the *df/dt* at the receiver (ground or space) it was necessary to compute the effect of dispersion on *df/dt* as the wave traveled from the source to the receiver. For frequencies near the whistler nose, this effect could often be neglected to first order. I applied my analysis to a "hook" triggered by a Morse dash, getting acceptable agreement with the data [*Helliwell*, 1967].

By applying the phase criterion for the IR length noted above, I was able to show that for values of *df/dt* less than about 1-2 kHz/s the length of the IR was to first order independent of *df/dt* [*Helliwell*, 1970]. This meant that such rising and falling ramps could be expected to have roughly the same growth rate and saturation values as a constant frequency signal; this prediction was supported by later Siple experiments [*Helliwell and Katsufrakis*, 1978]. For slopes exceeding 1-2 kHz the IR length was determined by direct computation [*Carlson et al.*, 1985].

The dot-dash mystery was solved experimentally with the first variable pulse length transmission from Siple Station in 1973 [*Helliwell and Katsufrakis*, 1974] where it was shown that at any given time all pulses, regardless of length showed the same temporal growth rate (typically 40-250 dB/s) in agreement with the "dot-dash" observations.

SATELLITE OBSERVATIONS OF WHISTLER MODEWAVES

Our first satellite VLF experiment on Explorer VI (1959) was a piggy back NSS receiver. It was built by Bud Rorden at Stanford Research Institute on a very short time scale. Complete testing with deployable antenna was not feasible, both cost-wise and time-wise. Although the overall mission was successful, the NSS receiver produced detectable signals only from the ground up to the *E* region and was virtually deadthereafter. No satisfactory explanation was ever found, but it was suspected that the monopole antenna may not have deployed properly.

Since the Vanguard III satellite recorded short whistlers with its magnetometer, it was clear that WM signals were readily detectable within the ionosphere, as expected [*Helliwell*, 1965, Fig 4-22]. Accordingly, we teamed with SRI to produce a whistler receiver for the range 300 Hz to 100 kHz to be flown on 060-1 in 1964. We created a new type of loop antenna, with the aid of the Lockheed Co., which consisted of an inflatable torus of mylar and aluminum (a 1-turn 3 meter diameter loop that was folded into a small package to be deployed by compressed CO_2 after orbit was achieved). It was highly successful, producing the first evidence of echoing signals in the non-ducted whistler mode (see Figure 4) that generally does not reach the ground with detectable intensity. Ducted whistlers on the other hand have wave normals that are confined to a small angle around B_0 and hence can more easily exit the ionosphere whose refractive index requires that only waves close to the vertical can escape into the earth-ionosphere waveguide (this is the critical angle effect of geometrical optics) [*Helliwell*, 1965].

The superposition of many echoing components in the magnetospherically reflected whistler gave the impression that it might be some kind of emission, but ever suspicious of the possible complexities of propagation we set to work first to understand the effects of propagation on the received signals. Smith and Angerami [*Smith and Angerami*, 1968] using an expression for refractive index that included ions together with a 2-D ray tracing program developed by *Kimura* [1966] showed that non-ducted waves would typically be reflected near the location where the wave frequency f and the lower hybrid resonance frequency f_L were equal. Each reflected echo would in turn be reflected near the LHR point in the conjugate hemisphere, creating a superposition of multiple echoes at the satellite altitude which we dubbed Magnetospherically Reflected (MR) whistler. See Fig 4. Other examples of MR whistler spectra and their path interpretation are shown elsewhere [*Edgar*, 1976]. MR whistlers tended to have high wave

region acting as source and sink for the overlying plasmasphere were all identified and even in some cases explained. The whistler method being essentially an integral technique was relatively immune to localized fluctuations and was able to provide plasma models that were later shown to be in good agreement with *in situ* satellite measurements.

But a crowning achievement of the IGY, in my view, was the bringing together of the waves of the thermal plasma (eg. whistlers and VLF emissions) with the charged particles of the radiation belts, in terms of wave particle interactions (WPI) in the magnetosphere.

One of these, the Doppler-shifted cyclotron resonance, plays a key role in the precipitation of electrons from the radiation belts [*Kennel and Petschek*, 1966; *Inan et al.*, 1989] and in the generation of the whistler-mode waves [*Brice*, 1963; *Kennel and Petschek*, 1966; *Helliwell*, 1967; *Carlson et al.*, 1990, and references therein].

A practical result of the synthesis of waves and particles was recognition that non-linear WPI phenomena could be investigated in a controlled fashion by VLF coherent wave injections from a ground transmitter. Accordingly, Siple station was established for this purpose and led to the discovery of a host of new non-linear phenomena [H*elliwell*, 1988] most of which had never before been observed or predicted. Although Siple Station had a finite lifetime [1973-88] due to its being crushed by the snow cover, these 15 years of data are being archived and continue to produce new results of importance to basic plasma physics as well as for diagnostics of the magnetosphere itself. At the time of closure of Siple Station, many new questions were being raised which require more experimentation. Accordingly, a new improved wave injection facility is needed and should provide exciting opportunities for extending knowledge of cosmic plasma physics.

Acknowledgments. Much of the funding for the whistler work reported here was provided by the National Science Foundation through the Atmospheric Sciences Section and the Office of Polar Programs. For support in my early research years I owe thanks to many colleagues in addition to each of those mentioned in this paper. I am especially indebted to the late Prof. (Research) John Katsufrakis who directed the field and data reduction programs.

REFERENCES

Armstrong, W. C., Lightning triggered from the Earth's magnetosphere as the source of synchronized whistlers, *Nature*, 6121, 1987.

Angerami, J. J., and D. L. Carpenter, Whistler studies of the plasmapause in the magnetosphere, 2, Equatorial density and total tube electron content near the knee in magnetospheric ionization, *J. Geophys. Res.*, *71*, 711, 1966.

Axford, W.I., and C. O.Hines, A unifying theory of high-latitude geophysical phenomena and geomagnetic storms, *Can. J. Phys.*, *39*, 1433, 1961.

Barkhausen, *Physik. Zeit.*, *1919* (20), 401.

Barkhausen, Whistling tones from the earth, *Proc. IRE, 18* (7), 1155, 1930.

Brice, N. M., An explanation of triggered very-low-frequency emissions, *J. Geophys. Res.*, *68*, 4626, 1963.

Burtis, W.J., and R. A. Helliwell, Magnetospheric chorus: Occurrence patterns and normalized frequency," *Plan. Space Sci.*, 1007, 1976.

Carlson, C. R., R. A. Helliwell, and D. L. Carpenter, Variable frequency VLF signals in the magnetosphere: Associated phenomena and plasma diagnostics, *J. Geophys. Res.*, *90*, 1507,1985; correction, *J. Geophys. Res.*, *90*, 6689, 1985.

Carlson, C.R., R. A. Hellilwell, and U. S. Inan, Space-time evolution of whistler mode wave growth in the magnetosphere,*J. Geophys. Res.*, *95*, 073, 1990.

Carpenter, D. L., The magnetosphere during magnetic storms; a whistler analysis, Ph.D. thesis, Tech. Rep. 12, Radiosci. Lab., Stanford University, 1962.

Carpenter, D.L., Whistler evidence of a 'knee' in the magnetospheric ionization density profile,*J. Geophys. Res.*, *68*, 1675-82, 1963.

Carpenter, D. L., Whistler studies of the plasmapause in the magnetosphere, 1, Temporal variations in the position of the knee and some evidence on plasma motions near the knee, *J. Geophys. Res.*, *71*, 693, 1966.

Carpenter, D. L., Ducted whistler-mode propagation in the magnetosphere; A half-gyrofrequency upper intensity cutoff and some associated wave growth phenomena, *J. Geophys. Res.*, *73*, 2919, 1968.

Carpenter, D.L., and K. Stone, Direct detection by a whistler method of the magnetospheric electric field associated with a polar substorm, *Plan. Space Sci.*, *15*, 395, 1967.

Dunckel, N., and R. A. Helliwell, Whistler-mode emissions on the OGO-1 satellite, *J. Geophys. Res.*, *74*, 6371, 1969.

Dunckel, N., B. Ficklin, L. H. Rorden, and R. A. Helliwell, Low-frequency noise observed in the distant magnetosphere with OGO-1, *J. Geophys. Res.*, *75*, 1854, 1970.

Draganov, A.B., U. S. Inan, V. S. Sonwalker, and T. F. Bell, Magnetospherically reflected whistlers as a source of plasmaspheric hiss,*Geophys. Res. Lett.*, *19*, 233-236, 1992.

Eckersley, T.L., Musical atmospherics, *Nature*, *135*, 104-5, 1935.

Edgar, B.C., The upper and lower frequency cut offs of magnetospherically reflected whistlers, *J. Geophys. Res.*, *18*, 205, 1976.

Gurnett, D. A., The Earth as a radio source: Terrestrial kilometric radiation, *J. Geophys. Res.*, *79*, 4227, 1974.

Helliwell, R.A., Ionospheric virtual height measurements at 100 kilocycles, *Proc.IRE, 37*, 887 , 1949.

Helliwell, R. A., *Whistlers and Related Ionospheric Phenomena*, Stanford, CA: Stanford University Press, 1965.

Helliwell, R. A., A theory of discrete VLF emissions from the magnetosphere, *J. Geophys. Res.*, *72*, 4773, 1967.

Helliwell, R. A., Intensity of discrete VLF emissions, in *Particles and Fields in the Magnetosphere*, B.M. McCormac (ed.), 292, 1970.

Helliwell, R.A., VLF wave stimulation experiments in the magnetosphere from Siple Station, Antarctica, *Rev. Geophys.*, *26*, 551, 1988.

Helliwell, R. A., and D. L. Carpenter, Whistlers excited by nuclear explosions, *J. Geophys. Res.*, *68*, 4409, 1963.

Helliwell, R. A., and E. Gehrels, Observation of magneto-ionic duct propagation using man-made signals of very low frequency," *Proc. IRE*, *46* (4), 785, 1958.

Helliwell, R. A., and J. P. Katsufrakis, VLF wave injection experiments into the magnetosphere from Siple Station, Antarctica, *J. Geophys. Res.*, *79*, 2511, 1974.

Helliwell, R.A. and J. P. Katsufrakis, Controlled wave-particle interaction experiments, Paper 5 in *Upper Atmosphere Research in Antarctica*, vol.29 of *Antarctic Research Series*, L.J. Lanzerotti and C.G. Park, editors, American Geophysical Union, Washington, 1978.

Helliwell, R.A., A. J. Mallinkrodt, and F. W. Kruse, Jr., Fine structure of the lower ionosphere, *J. Geophys. Res.* , *56*, 53, 1951.

Helliwell, R.A, J. H. Crary, J. R. Pope, and R. L. Smith, The 'nose' whistler -- a new high latitude phenomenon, *J. Geophys. Res.*, *61*, 139, 1956.

Helliwell, R. A., J. P. Katsufrakis, and M. Trimpi, Whistler-induced amplitude perturbations in VLF propagation, *J. Geophys. Res.*, *78*, 4679, 1973.

Helliwell, R. A., J. P. Katsufrakis, T. F. Bell, and R. Raghuram, VLF line radiation in the earth's magnetosphere and its association with power system radiation, *J. Geophys. Res.*, *80*, 4249, 1975.

Kennel, C. F., and H. E. Petschek, Limit on stably trapped particle fluxes, *J. Geophys. Res..*, *71*, 1, 1966.

Kimura, I., Effects of ions on whistler mode array tracing, *Radio Sci.*, *1*(3), 269, 1966.

Park, C. G., Whistler observations of the interchange of ionization between the ionosphere and the protonosphere, *J. Geophys. Res.*, *75*, 4249, 1970.

Potter, R.K., Analysis of audio-frequency atmospherics, *Proc. IRE*, *39* (9), 1067-69, 1951.

Raghuram, R., R. L. Smith, and T. F. Bell, VLF Antarctic antenna: Impedance and efficiency, *IEEE Trans. Ant. & Prop.*, *AP-22*, 334, 1974.

Smith, P., E.E. Thesis, Dept. of Electrical Engineering, Stanford Univsersity, 1953.

Smith, R. L., Propagation characteristics of whistlers trapped in field-aligned columns of enhanced ionization, *J. Geophys. Res.*, *66*, 3699 -3707, 1961a.

Smith, R.L., Properties of the outer ionosphere deduced from nose whistlers, *J. Geophys. Res.*, *66*, 3709-16, 1961b.

Smith, R. L., and J. J. Angerami, Magnetospheric properties deduced from OGO-1 observations of ducted and non-ducted whistlers, *J. Geophys. Res.*, *73*, 1,1968.

Storey, L.R.O., An investigation of Whistling Atmospherics, *Phil. Trans., Royal Soc. (London) A*, *246*: 113, July 9, 1953.

R. A. Helliwell, Space, Telecommunications, and Radioscience Laboratory, Stanford University, Stanford, CA 94305.

The Opportunity Years: Magnetic and Electric Field Investigations

James P. Heppner

Hughes STX Corporation, Greenbelt, Maryland

The author reviews his personal research activities and participations in space programs in the NRL and NASA environments between 1954 and 1970. Emphasis is placed on the origins and initial stages of (1) early space projects directed toward magnetic field investigations, and (2) the electric field projects initiated at the Goddard Space Flight Center in the mid-1960s.

1. SETTING THE STAGE

The stage for my space physics career was set during a 20 month leave from graduate studies at the California Institute of Technology following receipt of an MS in geophysics. In 1950 the thought of a job in Alaska was very appealing for a student sabbatical. On arrival at the Geophysical Institute of the University of Alaska I found that most of the institute's staff had resigned as a consequence of local political skirmishes and I was immediately given responsibility for the auroral aspects of contracts with both the Central Radio Propagation Laboratory of the National Bureau of Standards and the U.S. Army Signal Corps. Both projects were directed toward studying the effects of aurora on radio wave propagation and I knew virtually nothing about either aurora or the ionosphere. However, at age 22 one is not frightened by ignorance and I readily became fascinated with the challenges involved. What followed was two winters of visual observing with all-sky sketch mappings, continuous zenith photometry, and intermittent height measurements using photographic parallax from a short, 7 km baseline. The second winter photometric measurements were extended to ionospheric midpoints between E-W and N-S trans-Alaska transmitting and receiving stations but the descriptive visual mappings carried out through all dark, non-overcast hours provided the most useful information for analyses. The immediate results [*Heppner et al.,* 1952] were in the form of relating different sporadic-E densities to auroral forms, finding a close relationship between "blackout" r-f absorption and pulsating auroras which were found to usually occur at heights < 110 km, and the identification and naming of the "slant-Es" ionosonde signature. More importantly, a unique auroral database was acquired along with lasting mental images of auroral behaviors, both morphological and unusual.

Returning to Caltech I envisaged a thesis topic within the framework of the geophysics faculties traditional interests in the solid Earth. Instead Professors B. Gutenberg, C. Dix, G. Potapenko et al. not only approved but strongly encouraged a continuation of my auroral research and suggested that Oliver Wulf (Chemistry Dept.) would be an excellent advisor. This was particularly relevant because my objective was to make a detailed study of the magnetic disturbance accompanying different auroral forms and their diurnal sequence of occurrence and O. Wulf was known to have a long-standing interest in "magnetic bays" (also called "polar elementary storms" by C. Birkeland and now called "substorms" which is unfortunate because their occurrence is not dependent on "storms" characterized by sudden commencements and enhancements of the ring current or Dst index). Precedents for the study were non-existent and within the range of widely divergent auroral theories I could not find any with predictive relevance; thus the study was not burdened with existing concepts. Although two publications were spun-off during the study [*Heppner*, 1954,1955] the principal result in the form of a pattern of auroral behavior in magnetic local time (MLT),

Discovery of the Magnetosphere
History of Geophysics Volume 7

centered on coincident "auroral breakups" and negative bay onsets, was not published, simply because I became too engrossed in new work on leaving school. Later, the Defence Research Board of Canada requested permission to publish the entire 1954 thesis as a DR report [*Heppner*, 1958] for which I was both flattered and grateful. The thesis conclusion that aurora and its morphology had to be related to electric fields set the stage for my initiating an electric field program within NASA in the mid-60's.

In later years I revisited and updated the thesis study on two occasions. The first was at an aurora and airglow conference in Keele, England in 1966 at which I presented [*Heppner*, 1967] a motion picture showing at 2.5 minute increments the simultaneous magnetic variations at 25 high latitude observatories for 16 consecutive days. Amongst other features this illustrated the stability of the MLT vs. invariant latitude (INL) two-cell disturbance pattern over a wide range of disturbance levels. It further illustrated my disagreement with a then popular current system proposed by *Akasofu et al.* [1965] which treated the eastward evening auroral currents as being merely a return flow from a more poleward westward electrojet which encircled the pole, i.e., a one-cell system.

The second revisit was prompted by an invitation to contribute to a memorial volume of Geofysiske Publikasjoner honoring Professor Leiv Harang [*Heppner*, 1972c]. This provided an opportunity to not only acknowledge Harang's influence on my own studies but also to make other scientists more aware of his very basic and important contributions which had, unfortunately, been ignored by most investigators of high latitude current systems.. In effect my identification of negative bay (i.e., substorm) onsets and current reversals with a space-time discontinuity in auroral activity could be regarded as putting structure on a foundation previously laid by Harang [1946, 1951]. Thus it was appropriate to name this spatial feature, which shifts irregularly in position within a range of several hours at and preceding midnight, the "Harang Discontinuity." It is now gratifying to see this name being widely used [see e.g., *Fukushima,* 1994]. The Harang paper [*Heppner*, 1972c] also permitted an update relating the previous studies to the electric fields that were by then being measured by Ba^+ cloud motions and double probe instruments (discussed later).

2. EARLY ROCKET EXPERIMENTS

Very few laboratories were active in upper atmosphere research in 1954. Thus I was fortunate in finding a job at the Naval Research Laboratory following a letter to E. O. Hulburt, NRL's Director, that asked critical questions regarding the Bennett-Hulburt auroral theory based on magnetic self-focusing of ion streams from the Sun. Hulburt's gracious acceptance of criticism and Willard Bennett's enthusiastic demonstrations of his "terrella" experiments are engraved in my memory. As hired within the Rocket Sounding Branch, headed by Homer Newell, my first assignment was to assist Charles Johnson in fabricating and testing Bennett r-f mass spectrometers for measuring both positive and negative ion compositons from Aerobee rockets. This grass roots training in rocket instrumentation was invaluable for future endeavors and the flights after a year of lab work were highly successful [*Johnson and Heppner*, 1955, 1956].

Packard and Varian's [1954] announcement of free nuclear induction in the Earth's magnetic field was of considerable interest at NRL where there was a recognized need for a magnetometer with much greater accuracy than achieved in the pioneering rocket measurements of *Singer et al.* [1951]. In view of my interests in current systems I was given the opportunity to look into possibilities for developing this new technique into an operational magnetometer. This began a highly effective working relationship with scientists and engineers at Varian Associates that extended beyond the development of the proton precession magnetometer to the development of alkali vapor optical pumping magnetometers several years later. For an initial flight test of the proton magnetometer it was apparent that the Aerobee payload capability would not be fully utilized. This gave me the opportunity to pursue another objective, i.e., obtaining the altitude distributions of oxygen and sodium nightglow emissions, using modified versions of the photometers I had previously used in Alaska for auroral monitoring. This ambitious undertaking, flying two new experiments in the same payload, was greatly facilitated by the assistance of Leslie Meredith, my section supervisor who was also a valuable coworker in this and future endeavors. The pre-flight testing of this payload was arduous, in part because of the vacuum tube technology still in effect and the high levels of r-f interference encountered with dual, high power level, telemetry systems but also as a result of the special environments required for realistic tests of both experiments. Fortunately, the effort paid off. The flight in July 1956 was highly successful. The magnetometer performed flawlessly [*Heppner et al.*, 1958] and accurate altitude profiles of nightglow emissions were obtained [*Heppner and Meredith,* 1958]. This terminated my involvement in airglow research but rocket experiments with the proton magnetometer were continued with flights from Ft. Churchill, Canada during the IGY. The simplicity of proton magnetometers also rapidly made them popular for studies of ionospheric currents by other government laboratories and universities. Larry Cahill at the State University of Iowa and later at the University of New Hampshire was particularly active in this arena along with his graduate student Nelson Maynard. Later, within my GSFC group, rocket investigations of ionospheric currents were rejuventated by NRC research associates, T. Neil Davis and Keith Burrows, using Rb-vapor magnetometers to measure S_q and equatorial electrojet currents as well as auroral electrojets [e.g., *Davis et al.*, 1965; *Davis et al.,* 1967; *Burrows et al.*, 1971]. A review with tabulations and references through 1964 is given in *Heppner* [1968].

3. VANGUARD 3

In 1955 NRL was intensely engaged in selling Project Vanguard as the means of meeting the countries' commitment (announced in July 1955) to launch scientific Earth satellites during the IGY. My contribution at this stage was limited to outlining a scientific experiment whereby magnetic field measurements above the ionosphere could be used to determine the existence of an equatorial ring current at much greater altitudes during magnetic storms. This became one of two sample experiments used in NRL's proposals which were primarily concerned with vehicle and tracking feasibility issues. NRL received the go-ahead on Vanguard in the fall of 1955 and by 1956 the competition for satellite experiments was underway. Our successful rocket flight of a proton magnetometer in July 1956, noted above, demonstrated instrument feasibility and placed us in an excellent position for this competition. In February 1957 the "IGY Technical Panel for the Earth Satellite Program" made its selection of experiments to fly in the first full-size Vanguard satellites. We received position three in this queue, behind Lyman-alpha and cosmic ray experiments [for a detailed history, see *Green and Lomask,* 1970]. Our first opportunity was Satellite Launch Vehicle-5 (SLV-5) for which we had designed a 13-inch fiberglass sphere with a protruding tube to hold the polarizing/sensing coil away from the instrument, telemetry, and command electronics contained in the sphere. Following the early rocket flights, suitable transistors had become available and designs were greatly simplified. The launch looked good until the second stage pitched wildly and separated the satellite from the vehicle giving 400 seconds of meaningless signal.

The next and last opportunity became Vanguard 3. For this launch, September 1959, the magnetometer was again prime but solar X-ray, Lyman-alpha, and environmental experiments were added with additional space using a 20-inch sphere. The Explorer 1 radiation belt discovery had given new meaning to the ring current search but the belt electrons saturated and nullified the X-ray and Lyman-alpha measurements. Magnetometer operations were carried out almost exactly as planned over an 85 day period during which accurate measurements were obtained by command over the altitude range 510-3750 km within the meridian strips covered by Minitrack stations [*Heppner et al.,* 1960]. The Vanguard experiment also involved the installation of magnetic obsevatories at nine Minitrack stations to provide a disturbance reference for the satellite measurements. For this we developed a vector proton magnetometer based on applying homogeneous bias fields to a proton precessional magnetometer. With the speed and excellence of NRL's shops and John Stolarik and Ivan Shapiro's dedication these installations were completed by the spring of 1958 [*Shapiro et al.,* 1960].

Inasmuch as the analysis of the Vanguard measurements, scattered in space and time, depended critically on referencing the measured values to a common magnetic field model we had an obvious need for skills in using computers for spherical harmonic analyses. Fortunately, Joe Cain was interested and available for employment. This initiated his subsequent career in developing models of the Earth's main field and the Vanguard measurements became an important element in refining main field models for the next 10 years. The dominant handicap in analyses to reveal ring current variations was the uncertainty involved in evaluating the accuracy of the satellite's location as provided by the GSFC's orbit determination group. The technique for orbit determination involved analysis of the Minitrack data in weekly increments. The consequence was that plots of measured minus computed field values showed step-like displacements at weekly intervals, suggesting a very high precision within each week but making it obvious that the absolute accuracy of locations was uncertain [e.g., see *Heppner et al.,* 1960]. With time and many reworks of the location information, analyses finally demonstrated that the equatorial source of Dst variations had to be above Vanguard altitudes [*Cain et al.,* 1962].

As a by-product Vanguard 3 also provided the first measurements in space of whistlers and rising tone VLF emissions. The magnetometer's coil was a very effective antenna and the amplifier band-pass was well matched to whistler frequencies. Approximately 1200 whistlers were recorded during the two-second proton precession readout periods that followed approximately 4000 responses to ground commands. When statistically analyzed [*Shapiro et al.,* 1964] the occurrence and dispersion characteristics of these fractional-hop signals agreed very well with extrapolations from studies of whistlers from the Earth's surface.

4. NRL TO NASA TRANSITION

In late 1958 and early 1959 most of NRL's space effort was transferred to NASA to form the Beltsville Space Center which was soon, May 1959, renamed the Goddard Space Flight Center. We continued to work in NRL facilities and our preparations for Vanguard 3, now a NASA satellite, were unaffected. There was, however, an air of great expectation for new opportunities and resources [See *Rosenthal,* 1968 for a history of GSFC's early years]. In 1958 prior to my official transfer I was one of a small number of NRL scientists detailed part-time to NASA to layout a start-up space science program. Under Homer Newell's supervision the objective was sound science but also included impressing congressional committees. The fields and particles programs were, as one might anticipate, oriented toward initial projects for the new Beltsville Space

have been from its location on some mission planning chart. Within GSFC I was designated Project Manager and Tom Skillman served as Payload Coordinator. Many years later I was reminded that I was GSFC's first Project Manager for an in-house mission. The payload structure and all supporting electronic systems were designed and built within GSFC. Thus, line managements throughout GSFC played key roles and the Project Manager was obliged to work through management chains when there were problems. Fortunately, very few administrative problems were encountered in the positive atmosphere that prevailed.

Early in the design phase two events altered the course of the P-14 project. One was a meeting with Homer Stewart, representing the NASA Administrator. In this meeting I was told what I already knew: that the odds of hitting the Moon with a Thor-Delta launch were very low, that tracking facilities would not be able to verify a close approach, etc. My counter argument was that a Moon encounter would be a no-cost bonus to add to the magnetospheric crossing and interplanetary objectives; hence we should try. The over-riding argument was that it would be embarrassing to NASA and the country to aim for the Moon and miss. I accepted that argument with little choice and agreed that P-14 would be launched in an orbit that could not be interpreted as an attempt to hit the Moon. Later I wondered if Stewart, on assignment from Caltech, wasn't influenced by JPL's desire to go to the Moon with the Ranger program. Subsequently I was given the opportunity to fly a Rb-vapor magnetometer on Rangers 1 and 2. For this effort John Stolarik had the engineering responsibility and he did an excellent job working with the JPL team. Unfortunately both launches, in August and November 1961, were failures. A full account is given in *Hall* [1977]. On reading this document I found that I had given JPL numerous headaches over magnetic cleanliness: confirmation that John had done his job well !

The second event was the addition of a modified Faraday cup plasma probe designed for interplanetary measurements by Bruno Rossi's group at MIT. Rossi had very effectively lobbied through various channels to obtain a flight opportunity and following the rejection of our lunar objective, P-14 became an obvious candidate [See Rossi's autobiography, *Rossi,* 1990]. I was initially reluctant to accept this addition, primarily because I envisaged magnetic contamination jeopardizing our weak field objectives but also because we were making every effort to reduce weight to achieve a large distance from the Earth. The logic of joining the field and plasma measurements was, however, overwhelming. Meetings with Herbert Bridge and Frank Scherb, who were to carry out the plasma probe integration, also provided assurances regarding contamination and weight.

As a warm-up to flight test a Rb-vapor magnetometer and the new telemetry and ground systems designed for P-14, we scheduled a 4-stage Javelin flight from Wallops Island. This flight, December 12, 1960, was completely successful from the standpoint of magnetometer and systems performance but results were never published. The reason was that the rocket went northward from the predicted eastward trajectory and was lost to facility radars. Furthermore, from piecing together the available information, particularly from telemetry antenna tracking, it was quite possible that the payload and 4th-stage impacted in Boston harbor; a possibility that we did not wish to publicize.

P-14, renamed Explorer 10, was launched on March 25, 1961 following a trajectory that traversed the late evening geomagnetic tail in the southern hemisphere to an apogee at 46.6 Re. A variety of constraints such as solar radio noise at receiving sites, look angles for the optical aspect system and the plasma cup, and launch angle restrictions led to this trajectory. The result was numerous crossings of the geomagnetic cavity surface (i.e., the magnetopause) attrributed to diurnal changes in the tilt of Earth's magnetic axis relative to the solar wind as well as cavity expansions and contractions. The first crossing was in reasonable agreement with theoretical models for the cavity boundary [e.g., *Spreiter and Briggs,* 1962] but the total sequence of crossings suggested a slightly more conical or flared configuration than the nearly cylindrical tail predicted by the models. Earlier models predicting tear-drop and other geometries were clearly ruled out. One of the principal questions in data interpretations revolved around the extent to which closeness to the cavity surface would alter the properties of the interplanetary field and solar wind plasma. As there were no previous measurements for comparisons, both groups (MIT and GSFC) tried to look at all alternatives in interpretations [See *Bonetti et al.*, 1963; *Heppner et al.*, 1963].

The Vanguard 3 nemesis of uncertainties in orbit determination reappeared in the analysis of Explorer 10. This had little affect at large distances but near the Earth initial determinations with large residuals led to radiation belt ring current effects of several hundred nanoteslas. Our discomfort with this result, presented at a meeting in Kyoto, Japan [*Heppner et al.,* 1962], prompted us to put considerable pressure on the orbit determination personnel to continue work on the problem. Consistency was finally found using primarily minitrack data and this reduced diamagnetic ring current effects near 3 Re to 10 to 50 nanoteslas depending on the choice of reference field. During this period Norman Ness, who had joined our group near launch time, became an effective partner in arguing the orbit issue and also undertook the task of obtaining fluxgate calibrations from the Rb-vapor magnetometer. As reviewed in *Ness* [1996] this was his introduction to space exploration.

Explorer 10 was followed by Explorers 12 (August 1961) and 14 (October 1962) carrying fluxgate magnetometers with Larry Cahill as Principal Investigator. These measurements [reviewed by *Cahill,* 1995] outlined the dayside magnetopause and field configurations in the near-

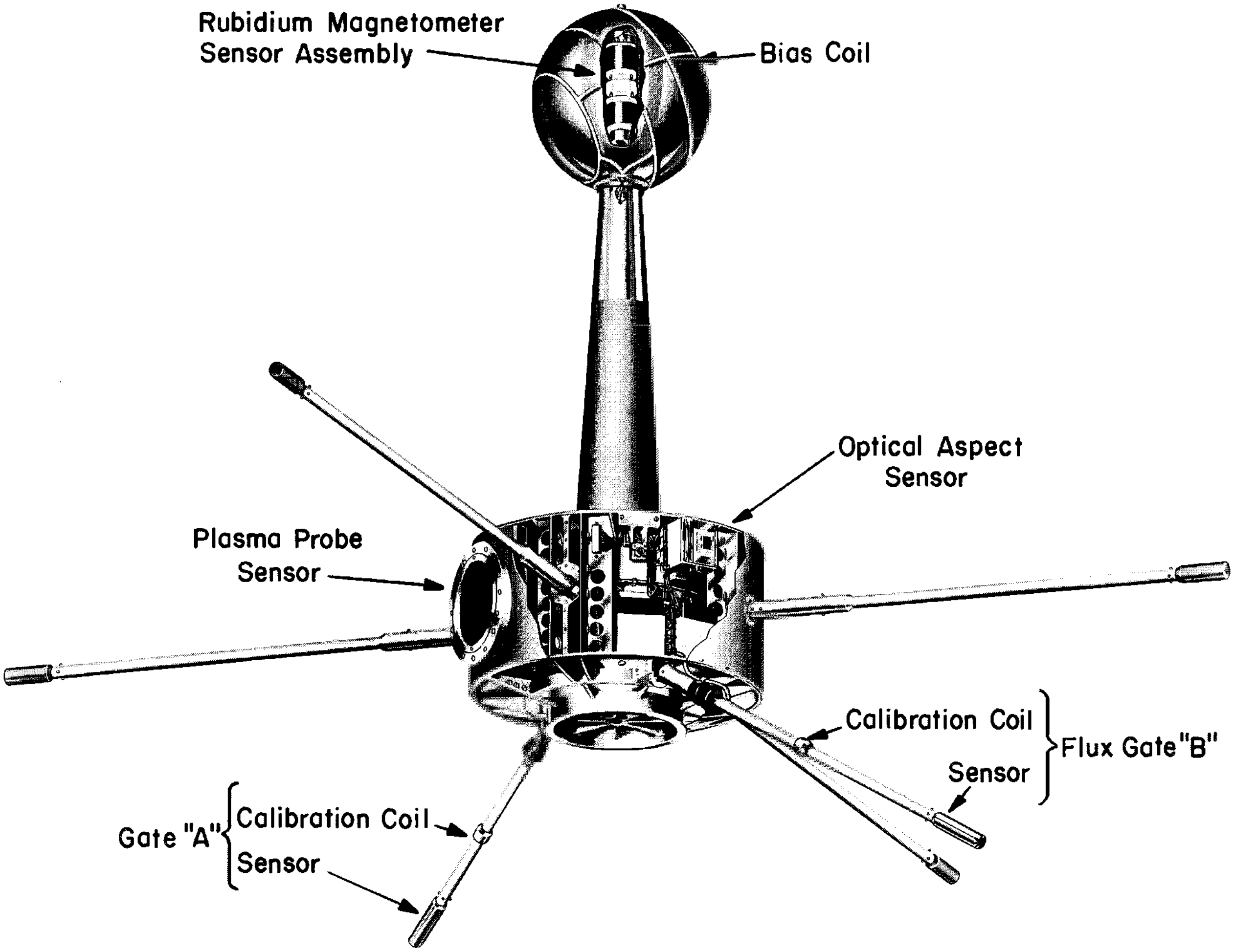


Figure 2: Explorer 10

tail. Explorer 18 (IMP-1, November 1963) followed with identification of the bow shock [*Ness et al.,* 1964] and mappings of the geomagnetic tail and thus completed an initial observational picture of the magnetosphere's configuration.

6. ORBITING GEOPHYSICAL OBSERVATORIES (OGOs)

The period following Explorer 10 and Rangers 1 and 2 was one of multiplying opportunities and a need for dividing work loads. As a first cut I took responsibility for OGO magnetic field experiments and Norman Ness became the PI for magnetic field measurements on what was to become the IMP series [See *Ness* 1996; *McDonald,* 1996]. In a later cut Joe Cain became the PI for the Polar Orbiting Geophysical Observatory (POGO) measurements on OGOs 2, 4, and 6 and I retained the PI role for the Eccentric Geophysical Observatory (EGO) measurements on OGOs 1, 3, and 5. Engineering was also divided into teams with Clell Scearce and Joe Seek supporting IMPs, Harry Farthing and Walter Folz supporting POGOs, and Tom Skillman supporting EGOs with the backing of two new post-doc research associates from Australia, Brian Ledley and Malcolm Campbell. IMPs 1, 2, and 3 and OGOs 1, 3, and 5 used combinations of fluxgate magnetometers and Rb-vapor magnetometers with bias fields for vector measurements. OGOs 2, 4, and 6, mapping the main field in low Earth orbit used Rb-vapor magnetometers for scalar measurements only; vector measurements in strong fields were ruled out by the need for extremely accurate orientation information as well as analog accuracy limitations.

My personal involvement in OGOs 2, 4, and 6 and main field mapping effectively ended after two related diversions: the World Magnetic Survey Program [*Heppner,* 1963] and the NASA-U.S.S.R. attempt to cooperate in space research in 1962-63. Following a widely publicized exchange of letters between Presidents Kennedy and Khruschev a U.S. delegation (H. Dryden, D., Hornig, and J. Townsend) was established to implement the "Bilateral Space Agreement of June 8, 1962 between the Academy of Sciences of the U.S.S.R. and the National Aeronautics and Space Administration of the U.S." Initially magnetic fields were designated as one of three areas of cooperaton and I was given the opportunity to serve as a technical advisor. Enjoyable meetings were held in Geneva, Rome, and Madrid but the outcome for magnetic fields eventually degenerated to exchange of surface observatory data. Satellite mapping with the exchange of measurements from absolute magnetometers was rather suddenly ruled out. My speculation was that Soviet authorities suddenly realized that magnetic field measurements could reveal the accuracy of their satellite tracking capabilities. Right or wrong, I believe this possibility impressed our delegation and possibly gave magnetic field mapping new status in the space program.

OGO 1 (also called OGO A and EGO 1), launched September, 5, 1964, was, for magnetic field studies, the most successful failure I can recall. The 3 section, 22 foot boom holding the magnetometers failed to deploy with the consequences: (a) that the spacecraft became spin stable with a 12 second period rather than being 3-axis stabilized, (b) the Rb-vapor magnetometer was left against the spacecraft body in high field gradients where it couldn't work, and (c) the 3-axis fluxgate was left in its undeployed position several feet from the spacecraft. The fortunate aspect was that the Z-axis of the dual range (0 - 30 and 0 - 500 nT) 3-axis fluxgate was aligned with the highly stable spin axis such that X and Y measurements could be made independent of spacecraft fields, and spacecraft fields in the Z direction were very weak. Accurate measurements over the range 4 to 24.5 Re were obtained for more than 20 months.

New findings from OGO 1 were published in a 54 page JGR paper [*Heppner et al.*, 1967] after extensive cutting of figures and a much longer manuscript, as requested by the editor. We realized at this point that we should have submitted at least 4 separate papers. The grouping of results did, however, lead to this paper being one of the first space research papers to be honored as a "Citation Classic" in *Current Contents* [*Heppner*, 1979]. Two of the results are, to me, particularly memorable. One was the construction of a model magnetic field profile of the cross-sectional structure of the bow shock. This resulted in part from the large number of crossings but primarily from OGO 1's sampling rates which were very high at that time. The high sampling rates also revealed shock associated waves. Coherent waves at frequencies near 1 Hz in the satellite reference frame were particularly prominent superimposed on the shock and in packets sunward from the shock. Masahisa Sugiura, who had much earlier joined our group and became a valuable co-worker in analyses, almost immediately identified these waves as being in the whistler mode.

The second most memorable result was the discovery of events we called "tail collapse" events immediately following negative bay onsets in the near midnight auroral zone. Today these are called "dipolarization" events and substorm onsets. Using my previous studies of auroral break-ups and bay onsets, my interpretation was that the events had to originate on the lowest latitude auroral shells and propagate outward in L-coordinates. Taking this one step further I speculated that the cause was probably short-circuiting in the auroral ionosphere with a readjustment of the electric field and electrostatic particle acceleraton along magnetic field lines. When first presented at an international ESRO magnetospheric colloquim in Stockholm in November 1965 [*Heppner,* 1967] Jim Dungey and Ian Axford enthusiastically greeted the "tail collapse" observations as support for their concept that sudden bay onsets were a consequence of reconnection with rapid merging of field lines in the geomagnetic tail. Arguments contrary to this interpretation were later presented in *Heppner et al.*, [1967] but for years to follow Axford tried to convince me that the initial appearance on the lowest latitude auroral shells was simply analogous to a break in a dam when the water pressure behind the dam passed some threshold.

Following OGO 1 my participations in satellite magnetic field experiments became intermittent primarily because I had by then become highly committed to our long antenna, double probe and chemical release, barium cloud programs for measuring electric fields. M. Sugiura, B. G. Ledley, and T. L. Skillman capably continued with both magnetometers functioning properly on OGO's 3 and 5 [e.g., *Sugiura et al.,* 1971].

7. INITIAL ELECTRIC FIELD MEASUREMENTS

7.1. *Double Probe Measurements*

Stemming from my thesis studies I was sometimes bothered by the thought that our magnetic field measurements were merely showing the effects of electric currents and not the cause. To study cause one should be measuring electric fields. The opportunity to do something about this first appeared in the form of Tom Aggson. When Tom joined our group in 1964 he made me aware that he wanted to do something new. My response was a challenge to find a technique for measuring weak electric fields in and above the ionosphere as opposed to the field mill techniques that had previously been attempted. A month or so later Tom came into my office with a pail of water with a cork and four protruding wires in an XY configuration floating on the surface. In answer to my question, what's this, he replied, these are double probes and you measure the floating potentials. I'm sure Tom enjoyed my initial confusion in trying to find some meaning in the word "floating" as applied to the water, but once the irrelevance was clarified he showed me what he had done on paper. He had clearly demonstrated that a high impedence voltmeter measuring the potential difference between two long cylindrical booms placed end to end would measure the difference in plasma potential between the two mid-points of the booms (i.e., $V_1 - V_2 = \mathbf{E} \cdot \mathbf{d}$ where **E** is the electric field in the plasma and **d** is the distance between mid-points). In addition, he had analyzed a variety of effects that could cause errors with the general conclusion that they would not be restrictive for most objectives, unless some unspecified condition caused large (e.g., 10 volt) potential differences between the spacecraft body and the booms.

The next step was to look for immediate flight opportunities that would not require hardware developments or a drawn out period of obtaining flight approvals from experiment selection committees. A piggyback experiment on the Advanced Technology Satellite, ATS-1, with long gravity gradient booms thus became our first objective and a proposal was rapidly put together in July 1964 and updated with additional detail in February 1965 [*Aggson and Heppner,* 1964, 1965]. Knowledge of these proposals spread rapidly and we received numerous inquiries for information, particularly from potential future competitors. Rather than take the time to prepare a journal paper on the technique, we simply supplied copies of our proposals. In fact, with Tom's reluctance to write, the technique did not reach the open literature until 1969 [Aggson, 1969] in the proceedings of a NATO Advanced Study Institute in 1968 and this did not include the details presented in proposals. However, by then we could show results from 1966 rocket flights [*Aggson*, 1969] and OV1-10 [*Heppner,* 1969] launched in December 1966. The proposal era extended to proposals for ATS D and E in September 1965; POGO satellites (OGOs 4 and 6) in October 1965; EGO-3 (OGO 5) in December 1965; and IMP spacecraft H, I, and J in October 1966.

During this period Nelson Maynard became a valuable addition to our team and took responsibilities for the piggyback flight we received from the Air Force on OV1-10. OV1-10 unfortunately suffered a short-circuit to one boom which negated dc measurements but the ac channels from the other boom supplied a wealth of data on the global distribution of electric field irregularity structures and their high latitude occurrence boundaries [*Heppner,* 1969; *Maynard and Heppner,* 1970]. Our ATS flight was a total loss as a consequence of an upper stage vehicle failure. However, these early flights, conducted at very little cost, demonstrated the simplicity of the cylindrical double probe approach. If Tom had initially taken the alternative of using spherical probes, satellite flights would have had to wait for the development of long booms with internal wiring. Asymmetric shadowing of spherical sensors by the holding booms would also have presented problems in gravity gradient configurations with the measurement axis not normal to the sunline.

The limitations of the double probe technique first appeared with OGO 5. Beyond the plasmapause where plasma densities dropped below 100/cm^3 it was predicted that the photoelectron flux from the probes and spacecraft would, with exceptions for select conditions, be greater than the ambient plasma flux and that under these conditions the probes and spacecraft would be positively charged. Analyses did not, however, predict the complex interactions between the sheath surrounding the spacecraft and the electric field probes as observed by OGO 5. Initially it was thought that the use of electric field baselines (i.e., mid-point separations, **d**) much greater than the 20 meters used on OGO 5 would solve this problem. This turned out to be only partially correct. IMPs 6 and 8 later showed that the influence extended to dimensions of the order of 100 meters. Adequate explanations were not found but it appeared likely that the large sheath dimensions came from the high energy tail of the photoelectron distribution. Interpretations of measurements were further complicated by the fact that the interactions between the sheath plasma and the probes were sensitive to density and temperature changes in the ambient plasma and these variations could be incorrectly interpreted as being electric field variations. The complexity of the sheath overlap problem became even more evident when it was observed that the sheath did not collapse when the spacecraft potential was small or zero, thus indicating that control of the potential would not solve the problem. Despite these problems some measurements, at select locations and times, were believed to be valid. These were associated with solar wind and bow shock fields, fields near the magnetopause, and transient fields on nightside auroral L-shells [e.g., *Aggson and Heppner,* 1977; *Maynard et al.,* 1982].

In contrast to the sheath problems encountered in the outer magnetosphere, the OGO 6 near Earth polar orbiting measurements in 1969 were completely reliable and highly accurate. They provided a wealth of data and firmly established the fundamental two cell configuration of high latitude convection. Analysis of this data continued for a number of years with an eventual development of model convection patterns [*Heppner,* 1977]. The need for plural models rather than a single model came from the much earlier finding [*Heppner,* 1972b] that the distribution of polar cap electric fields was dependent on the Y component of the interplanetary magnetic field, B_y. Dawn-dusk asymmetries in the distributions reversed with reversals in the sign of B_y and were of the opposite sense in the northern and southern hemisphere polar caps (i.e., distributions skewed to the dawn meridian in the northern hemisphere were accompanied by distributions skewed to the dusk meridian in the southern hemisphere, and vice versa). This behavior has ever since been a key element in developing reconnection models of the solar wind - magnetosphere interaction.

Preceding the model studies, above, the OGO 6 data was used to look for large scale electric field changes that might explain the onset of substorms [*Heppner,* 1972a]. The findings in this case were negative in that no unusual distributions or intensity changes were found preceding the time of a substorm onset. Effects, such as turbulence near the late evening polar cap boundary were evident after, not before, the onsets. Thus the substorm onset problem remained elusive, as it is today.

7.2 *Measurements Using* Ba^+ *Clouds*

Parallel in time with the development of the double probe technique at GSFC, the Max-Planck Institute (MPI) under Professor R. Lust was developing a chemical release technique for creating barium ions that could be optically tracked to determine the vector velocity, **v**, and thus the electric field, **E**, from the relationship $\mathbf{E} = -\mathbf{v} \times \mathbf{B}$ where **B** is the magnetic field [*Foppl et al.,* 1965, 1968; *Haerendel et. al.,* 1967]. I was greatly impressed for several reasons: one, simultaneous ion cloud and double probe measurements could be used to prove the double probe technique; second, double probe and ion cloud measurements would be complementary rather than competitive for most scientific objectives; and third, it was apparent that the ion cloud approach could be uniquely powerful in mapping the time-space electric field characteristics over an extensive region through use of multiple releases. The latter reason was particularly compelling for detailed studies of field configurations in regions such as the Harang discontinuity and polar cap boundaries and I was convinced that we had to exploit this tool as part of an electric fields program. Accordingly I began looking into sources for fabricating chemical canisters that would essentially be copies of the MPI chemical technology. Simultaneously, I discussed this with Gene Wescott, who had applied for an NRC research associateship at GSFC, and he was immediately interested in taking a lead role in this new program. In less than 2 months following Wescott's arrival in September 1966 we issued a competitive procurement action for fabricating chemical release payloads. In January 1967 a contract was granted to the Astro-Met Division of the Thiokol Chemical Corporation and in March 1967 we conducted our first flight test from Wallops Island with three successful releases. Support from the Astro-Met group was particularly advantageous because this same group was under contract with GSFC's sounding rocket division for rocket field services. Thus there was a considerable cost savings in remote site operations through use of a common crew. Thiokol also developed assembly line efficiencies in producing canisters and payloads. With Thiokol producing the payloads our efforts at GSFC were directed to the two other aspects of a chemical release program: (1) assembling systems for synchronized imaging from multiple observing sites, and (2) developing computerized tools for tracking and triangulating the released clouds relative to star backgrounds. Mary Miller was initially given responsibilities for item (2) and with time, and Wescott's return to the University of Alaska in 1970, she looked after observing logistics as well as the analysis of cloud motions.

Eleven Ba^+ clouds were released in our first auroral belt measurements from Andennes, Norway in 1967 [*Wescott et al.,* 1969]. Positive bay, negative bay, and Harang discontinuity conditions were all encountered. In all cases the ion cloud motions revealed electric fields perpendicular to the ionospheric current, clearly demonstrating for the first time that auroral electrojets were Hall currents. Along with this expected result, the clouds also indicated agreement with the double probe finding [*Aggson,* 1969] that within narrow auroral arcs the electric field drops to low values.

Our next objective was to obtain measurements within the polar cap. For this we obtained approvals from the U.S. Air Force, Canadian Air Force, and the National Research Council of Canada to use the Pin-Main DEWline (Distant Early Warning) radar site on the Cape Perry Arctic coast as a launch site and other radar stations in the Pin sector for camera sites. This was the beginning for our future use of DEWline sites for numerous launches that were very efficiently conducted in winter months by freezing launchers into several feet of ice.

The three Pin-Main launches with 12 releases in March 1969 were probably the most scientifically significant set of launches in our program. The simultaneous directions of the electric field, from the anti-solar cloud motions, and the magnetic disturbances, from our temporary observatories at the camera sites, could not be attributed to ionospheric Hall currents or other elements of the ionospheric conductivity tensor. This clearly demonstrated that the traditional global current patterns, showing a closure of auroral belt currents across the polar cap, were incorrect. From this one could conclude that the auroral electrojets had to complete their circuits via field aligned currents into the magnetosphere, and further, that the field aligned currents were responsible for the polar cap magnetic variations [see *Heppner et al.,* 1971 for discussion and additional references].

From the early releases, above, we also found that we were obtaining unique information on thermospheric winds at altitudes 200 - 320 km by tracking the neutral strontium clouds accompanying each barium release. To expand this information we added TMA/TEA, and later lithium, trail releases in our payloads to cover the 90 to 200 km range. This by-product information on high latitude wind profiles from 90 to 320 km at numerous high latitudes and local times was eventually summarized in *Heppner and Miller* [1982].

8. APOLOGIES AND COMMENTS

While writing the above account of my personal involvement in space research prior to 1970, I was

Heppner, J. P., Note on the occurrence of world-wide s.s.c.'s during the onset of negative bays at College, Alaska, *J. Geophys. Res., 60,* 29-32, 1955.

Heppner, J. P., A study of relationships between the aurora borealis and the geomagnetic disturbances caused by electric currents in the ionosphere, Ph.D. thesis, Calif- ornia Institute of Technology, 1954; Report No. DR 135 Defence Research Board, Canada, 1958.

Heppner, J. P. and L. H. Meredith, Nightglow emission altitudes from rocket measurements, *J. Geophys. Res., 63,* 51-65, 1958.

Heppner, J. P., J. D. Stolarik, and L. H. Meredith, The Earth's magnetic field above WSPG, New Mexico, from rocket measurements, *J. Geophys. Res., 63,* 277-288, 1958.

Heppner, J. P., High latitude magnetic disturbances, in *Aurora and Airglow,* edited by B. M. McCormac, pp. 75- 92, Reinhold Publishing Corp. New York, 1967.

Heppner, J. P., J. D. Stolarik, I. R. Shapiro, and J. C. Cain, Project Vanguard magnetic field instrumentation and measurements, *Space Research I,* edited by H. K. Kallmann Bijl, pp. 982-999, North-Holland Publishing Co., Amsterdam, 1960.

Heppner, J. P., N. F. Ness, T. L. Skillman, and C. S. Scearce, Magnetic field measurements with the Explorer 10 satellite, *J. Phys. Soc. of Japan, 17,* 546-552, Supplement A-2, 1962

Heppner, J. P., The world magnetic survey, *Space Sci. Rev., 2,* 315-354, 1963.

Heppner, J. P., N. F. Ness, C. S. Scearce, and T. L. Skillman, Explorer 10 magnetic field measurements, *J. Geophys. Res., 68,* 1-46, 1963.

Heppner, J. P., Recent measurements of the magnetic field in the outer magnetosphere and boundary regions, *Space Sci. Rev., 7,* 166-190, 1967

Heppner, J. P., M. Sugiura, T. L. Skillman, B. G. Ledley, and M. Campbell, OGO-A magnetic field observations, *J. Geophys. Res., 72,* 5417-5471, 1967.

Heppner, J. P., Satellite and rocket observations, in *Physics of Geomagnetic Phenomena,* Vol. 2, ed. by S. Matsushita and W. H. Campbell, pp.935-1036, Academic Press, Inc., New York, 1968

Heppner, J. P., Magnetospheric convection patterns inferred from high latitude activity, in *Atmospheric Emissions,* edited by B. M. McCormac and A. Omholt, pp. 251-266, Van Nostrand Reinhold Co., New York, 1969.

Heppner, J. P., J. D. Stolarik, and E. M. Wescott, Electric field measurements and the identification of currents causing magnetic disturbances in the polar cap, *J. Geophys. Res., 76,* 6028-6053, 1971.

Heppner, J. P., Electric field variations during substorms: OGO-6 measurements, *Planet. Space Sci., 20,* 1475-1498 1972a.

Heppner, J. P., Polar cap electric field distributions related to the interplanetary magnetic field direction, *J. Geophys. Res., 77,* 4877-4887, 1972b.

Heppner, J. P., The Harang discontinuity in auroral belt ionospheric currents, *Geofysiske Publikasjoner, 29,* 105-120, 1972c.

Heppner, J. P., Empirical models of high latitude electric fields, *J. Geophys. Res., 82,* 1115-1125, 1977.

Heppner, J. P., Commentary on "OGO-A Magnetic Field Observations," *Current Contents, 19,* No. 1, 14, 1979.

Johnson, C. Y. and J. P. Heppner, Night-time measure- ment of positive and negative ion composition to 120 km by rocket-borne spectrometer, *J. Geophys. Res., 60,* 533 1955.

Johnson, C. Y., and J. P. Heppner, Daytime measurement of positive and negative ion composition to 131 km by rocket-borne spectrometer, *J. Geophys. Res., 61,* 575, 1956.

Maynard, N. C. and J. P. Heppner, Variations in electric fields from polar orbiting satellites, in *Particles and Fields in the Magnetosphere,* edited by B. M. McCormac, pp. 247-253, D. Reidel Publishing Co.,, 1970.

Maynard, N. C., J. P. Heppner, and T. L. Aggson, Turbulent electric fields in the nightside magnetosphere, *J. Geophys. Res., 87,* 1445-1454, 1982.

McDonald, F. B., IMPs, EGOs, and Skyhooks, *J. Geophys. Res., 101,* 10521-10530, 1996.

Ness, N. F., C. S. Scearce, and J. B. Seek, Initial results of the IMP-1 magnetic field experiment, *J. Geophys. Res 69,* 3531-3569, 1964.

Ness, N. F., Pioneering the swinging 1960s into the 1970s and 1980s, *J. Geophys. Res., 101,* 10497-10509, 1996.

Packard, M. and R. Varian, Free nuclear induction in the Earth's magnetic field (Abstract), *Phys. Rev., 93,* 941, 1954.

Rosenthal, A., *Venture into Space: Early Years of Goddard Space Flight Center,* 354 pp., NASA Center History Series, NASA SP-4301, Washington, D. C., 1968.

Rossi, B. *Moments in the Life of a Scientist,* 181 pp., Cambridge University Press, 1990.

Shapiro, I. R., J. D. Stolarik, and J. P. Heppner, The vector field proton magnetometer for IGY satellite ground stations, *J. Geophys. Res., 65,* 913-920, 1960

Shapiro, I. R., J. D. Stolarik, and J. P. Heppner, Data report on whistlers observed by Vanguard III, *NASA Technical Note, NASA TN D-2313,* 27 pp., Goddard Space Flight Center, Greenbelt, MD, 1964.

Singer, S. F., E. Maple, and W. A. Bowen, Jr., Evidence for ionospheric currents from rocket experiments near the geomagnetic equator, *J. Geophys. Res., 56,* 265-281, 1951.

Spreiter, J. R. and B. R. Briggs, Theoretical determination of the form of the boundary of the solar corpuscular stream produced by the interaction with the magnetic di- pole field of the Earth, *J. Geophys. Res., 67,* 37-51, 1962.

Sugiura, M., B. G. Ledley, T. L. Skillman and J. P. Heppner, Magnetospheric field distortions observed by OGOs 3 and 5, *J. Geophys. Res., 76,* 7552-7565, 1971

Wescott, E. M., J. D. Stolarik, and J. P. Heppner, Electric fields in the vicinity of auroral forms from motions of barium vapor releases, *J. Geophys. Res., 74,* 3469-3487, 1969.

J. P. Heppner, 6201 Green Valley Road, New Market, Maryland 21774

The Magnetosphere is Brought to Life

Colin O. Hines

15 Henry Street, Toronto M5T 1W9 Canada

An account is given of scientific and institutional changes encountered by the author and by some of his colleagues during the birth of the magnetosphere as an object of scientific investigation. The role of the 1961 Axford-Hines paper on convection of the high-latitude geomagnetic field lines and associated phenomena is highlighted.

1. THE PRE-NATAL STATE

The magnetosphere, as yet unnamed, entered the world of scientific investigation in the early 1950s in response to the whistler studies of Owen Storey, published in 1953 but given currency a year or two before that.

Prior to his work, the ionosphere was effectively terminated as an object of scientific concern at or immediately above the level of maximum ionization of the F layer, there having been no means of examining it at greater heights. True, a ring current had been hypothesized by Sydney Chapman and V. C. A. Ferraro to form at several earth radii within the geomagnetic field at times of magnetic storm, but it was viewed as a transient and isolated phenomenon. True too, scintillations of radio stars had been attributed to fluctuations of electron density, but whether in Earth's near vicinity or in interplanetary space was uncertain. And, in what came closest to heralding things to come, David C. Martyn had suggested an explanation of the ionospheric equatorial (or geomagnetic) anomaly on the basis of an electrodynamically driven uplifting of equatorial ionization above the F peak and its subsequent diffusion downward along magnetic field lines to off-equatorial latitudes; but the process, if operative, could not be examined in situ.

Nor could theorists contribute much to invigorate studies above the F peak, for they had little information on temperature or chemical transition (specifically, to a hydrogen-dominated regime) on which to base upward extrapolations beyond the observed domain. The theory of thermal escape of molecules from Earth's gravitational field, and so of the exosphere, was already begun; but it concerned only the neutral gas, or so it seemed at the time, and had nothing to do with the magnetosphere proper.

2. THE BIRTH

Into this virtual vacuum were thrust the findings of *Storey* [1953]. His detailed studies of whistler characteristics, most notably their form of dispersion, permitted no rational explanation other than propagation nearly along geomagnetic field lines, arching up to heights of a few earth radii from hemisphere to hemisphere, through a medium surprisingly rich in electron concentration. Hydrogen ions were postulated as partners to the electrons, first as invaders from the sun [*Storey*, 1953] and then as a direct extension of the normal ionosphere [*Dungey*, 1954]. Suddenly the region was populated by interesting plasma and interesting processes, and was itself susceptible to scientific investigation. The essential features of the magnetosphere—ionization and the geomagnetic field, extending together upward from the F peak—were now in place. The magnetosphere had arrived.

3. MY OWN EARLY INVOLVEMENT

My own research career had begun slightly earlier, in 1948, at the Radio Propagation Laboratory (RPL) of Canada's Defence Research Board (DRB) in Ottawa, Canada. RPL was an accidental step-child of war-time operations in the field of radio propagation, a home in which to locate some of the senior scientists of the war years. It was headed

Discovery of the Magnetosphere
History of Geophysics Volume 7

at the time by Frank T. Davies, a good-hearted but occasionally curmudgeonly Welshman, with James C. (Jim) Scott as second in command.

My appointment to RPL was as a summer student, following my third year of university. DRB and RPL, themselves only recently created and so still mere novices in the business, had yet to find an effective way of handling summer students. My own assignment turned out to be the microfilming, for an incipient RPL library, of a large number of journal papers deemed to be relevant to ionospheric studies: the very latest technology (of its day) for information storage and retrieval.

I maintained my spirits intact through the summer, but only by taking the opportunity of my frequent visits to the library of the National Research Council of Canada (NRC) to do some reading in matters closer to my own heart, mainly in fundamental physics. Aside from that, the job itself was a pitiful introduction to what I had been informed was scientific research.

When, toward the end of August, I was told that I must submit a formal letter to terminate my employment, I had no hesitation in writing, "It is with great pleasure that I tender my resignation, effective ..." The secretary to whom I handed it, Rita Richard (later Langille), refused to accept it in that form; and, since she was such a sweet person, I rewrote it according to her prescription.

The fortunate outcome of this byplay was that Ms Richard drew my plight to the attention of Jim Scott, who had just returned from a summer in Switzerland serving on a frequency-allocation committee or suchlike, and he was suitably appalled. A theorist himself, he immediately assigned me to a problem he had been meaning to look into: the magnetic susceptibility of an ionized gas. Within a week I was able to report to him an answer, noting that an ionized gas turned out to be diamagnetic. This may have been well known in some quarters already at that time, but to him (as to me) it was a surprise and an interesting one.

He then assigned me to another problem, the designing of a nomogram for the convenient calculation of radio-wave dispersion in a dissipative ionized medium. (These were the days before computers—before even desk calculators, other than mechanical, often hand-driven. In my previous summer's employment I had even been given a "cylindrical slide rule" of five-figure accuracy as my calculating device.) Despite the fact that a Ph.D. had worked all summer on this project and had failed to achieve the wanted nomogram, I managed to design an acceptable product before my time was up. Thus began a long and happy relationship with Scott, whom I have ever since considered to be my scientific mentor (though he did once tell me that I would never be a great scientist, because my interests were too diverse).

The following summer produced quite a different story, as did the year following my master's degree. Scott treated me from the outset as a productive researcher and guided me into what became fruitful problems concerning propagation in an ionized medium. In the last of these, he directed me to Hannes Alfven's recently published volume, Cosmical Electrodynamics, and suggested that I would do well to come to an understanding of hydromagnetic theory. This I did, but primarily in the sense of understanding hydromagnetic waves as an extension of ionospheric magneto-ionic theory, a route not taken by Alfven (who, indeed, assumed always an isotropic conductivity). This study became a catalyst for much of my later work.

It was with this background that I arrived in Cambridge, England, for doctoral studies during 1951-53. I must have overlapped with Storey, but in fact I do not recall meeting him or learning of his work at that time, despite the fact that he was a student in J. A. (Jack) Ratcliffe's group, with which I formed a loose affiliation. My own studies, after a brief fling with Ritz's theory of relativity, centered on motions in the ionosphere. These turned out not to be explicable as hydromagnetic waves, suggested for my consideration by Ratcliffe, but rather as atmospheric gravity waves, which I have studied on-and-off through most of my research years since then. (In *Hines* [1989], I have outlined my entry to and early involvement in this field, and the consequences.)

After a year at University College, London, for post-doctoral theater-going, I returned to RPL in the summer of 1954, to find that it had been relocated to Shirley Bay, west of Ottawa, had been renamed the Radio Physics Laboratory, had been extended by what was soon to become the Communications Laboratory (CL), and had been united administratively with the Electronics Laboratory (EL), located east of Ottawa, into the Defence Research Telecommunications Establishment (DRTE) with Davies now Chief Superintendent and Scott now Deputy Chief Superintendent. I was allowed to finish my gravity-wave studies as they then existed, but was soon reoriented into meteor studies under Peter A. Forsyth. [See *Hines*, 1989.]

A year or two later, Storey joined the RPL staff, though I have no idea how he came to be recruited or why he would have accepted. He began establishing an observational program of whistler studies, but only on a modest scale, for he was not given support at the level appropriate to his accomplishments and prospects. He suggested to me that I might look into the propagation of whistlers when account was taken of ions as well as electrons. This work led to the recognition that the geomagnetic guidance of whistlers broke down at the hybrid frequency given by the geometric mean of electron and ion gyrofrequencies. [*Hines*, 1957*a*.]

At about this time, E. N. (Gene) Parker, who had recently introduced the solar wind, took issue with the long-standing concept of the terrestrial ring current. His objection was based on the existence of the high plasma concentrations that had been revealed by whistlers. He argued that any magnetic effects the hypothesized ring current might produce at ground level would arrive only after diffusion through this plasma, which diffusion would take too long to produce the observed behavior. I took exception to his view, arguing instead that the magnetic effects would propagate as hydromagnetic waves, with delay times of a few seconds at most [*Hines*, 1957*b*].

Further discussion with Storey led to our jointly publishing a somewhat revised version of this view [*Hines and Storey*, 1958]. It took into account the possible effects of multiple reflections along field lines, and it produced an irregular rise of the observable disturbance in times of the order of a few minutes, as observed. The heart of the issue was resolved in due course [*Hines and Parker*, 1960], once Parker accepted that the ring current would be formed within the geomagnetic domain rather than outside it as he had previously been thinking. Indeed, he and A. J. (Alex) Dessler went on to develop a theory of the ring current—one that, incidentally, stressed its diamagnetic nature. (An otherwise similar model had already been suggested by S. Fred Singer, who called upon Alfven's "guiding-center" drifting of low-energy ionization within the geomagnetic domain.)

My involvement with Storey, extending to that with Parker, had thrust me into magnetospheric studies, just then gaining impetus from the detection of energetic trapped particles by James A. Van Allen and his colleagues. This thrust was given further impetus by certain institutional changes that proceeded simultaneously.

4. INSTITUTIONAL REORGANIZATION

Early in 1958, Davies was ordered into DRB Headquarters (DRB/HQ) in order to "serve time," as he thought of it, alongside militarily oriented bureaucrats rather than in the research ambiance that he so much enjoyed at DRTE, until then his personal fiefdom. Scott was named to succeed him. Forsyth, only recently named as Superintendent of RPL, was on the point of departure for the academic world, and Scott named me to succeed him. This was an act of courage (or frivolity) on his part, for I was suspect to many both because of my relative youth and because of my theoretical predilections. However, had I not been given the appointment, I would have looked elsewhere for a more appropriate home: the constraints of research conducted in a defense organization had proven to be unpalatable to me, at least unless I was the one implementing or, better, circumventing them. I thought I saw an opportunity of building a research group of such strength that DRB would simply have to put up with it (or give it away to NRC, its more natural home at the time). I was quite prepared to dedicate a decade of my life to the project.

Putting first things first, upon my appointment being announced I offered to Storey whatever support he wanted that I might now be able to provide. Unfortunately, he had already made plans to go elsewhere—the U.S. National Bureau of Standards (NBS), as it was then, in Boulder, Colorado—and could not be dissuaded. For me, it was a poor start.

I fared better when I sought out Jules A. Fejer, then working in South Africa, who I supposed might be looking for an escape route from that then benighted country. This proved to be the case, and I was able to wheedle out of my superiors a salary for him that, though inappropriately low for his talents, was adequate to his needs. (The in-house battle over the offer to be made him derived in part, no doubt, from the suspicion in which I was held. The powers-that-were ultimately agreed to a salary somewhat less than my own, where I had been demanding one considerably greater than the pittance I was then being paid. The smirks with which my request had been met were turned to blushes when, too late to have an immediate effect, a letter arrived from Ratcliffe in support of Fejer's pending appointment. After extolling Fejer, Ratcliffe stated that the only concern he had was that Fejer would find himself in something of a scientific wasteland, with no one to talk to of a calibre suitable to his talents. I do believe that my own stock rose a little in consequence, at least as a manager if not as a scientist, and in due course both Fejer and I received appropriate increments.)

Aside from questions of personnel, there was the question of program. RPL had hitherto been operated as a collection of groups that had grown with their own individual interests, but usually under the protective coloration given them by Canada's position relative to the auroral zone, which made for radio-communications problems in the military. It had been argued that any basic research related to the ionosphere—particularly the disturbed ionosphere—must be suitable for DRB's purposes. I resolved to take this position a step further and actually make the disturbed ionosphere the focus of our research activities, no longer just in name but now also in fact.

A group headed by T. R. (Ted) Hartz, including George C. Reid as consultant, formed the observational nucleus, one centered on VHF radio-star observations that revealed such things as polar-cap absorption. A second, headed by J. S. (Jack) Belrose, and including R. E. (Ron) Barrington, who took over such whistler studies as were left after Storey's departure, dealt with observations at lower frequencies. (It was to this group that Fejer most closely attached himself, on arrival.) A third group, headed by Walter J. Heikkila, was concerned with scattering of radio waves in the troposphere; but he came to be persuaded that a better future lay in the development of rocket techniques for auroral observations to be made at Churchill, Manitoba, a DRB base and the site of many rocket launches during the International Geophysical Year, then in progress.

But the ultimate heart of the program, as I saw it, was to be what I called the Synoptic Studies Group. Its mission would be to view with dispassion all these local studies, meld them with complementary studies conducted on site in CL and others reported in the journal literature, and so produce a comprehensive integration of the diverse and loosely correlated types of information that were building in our area of concern. The initial scientists in this group were Doris Jelly and Clare Collins, but they were soon joined by Irvine Paghis, an administratively senior scientist, formerly of EL by way of DRB/HQ, who wanted an opportunity to return to research of whatever nature RPL might offer him. I accepted him as head of the group, and he quickly turned to geomagnetic fluctuations as his personal field within its broader context. Within a year, the group was extended again when Louise Herzberg, wife of the Nobel laureate but a well-established scientist in her own right, sought out a position at RPL (leaving behind one at the Dominion Observatory), conceiving RPL to be "the place to be" for the most exciting research activity of the day.

As part of RPL's role, an extension course was given at Carleton University, in suburban Ottawa, by our staff members (augmented by Bert C. Blevis and Ray Montalbetti from CL). It later provided the basis for a text on the upper atmosphere [*Hines et al.*, 1965], of which there were virtually none at the time.

Not long after Heikkila was reoriented to rocket studies, Scott called me into his office and asked whether I would like RPL to advance to satellite-borne studies. There was a proposal (initiated, I believe, by Scott's deputy, John H. Chapman, and Eldon S. Warren of CL) for a "topside sounder," a miniaturized ionosonde that would be little bigger than a baseball and would be fitted into someone else's satellite. It was supposedly a trivial thing to build but one having great scientific potential. I replied that I would be interested, but all in good time: Heikkila should first get a sound grounding with rockets (as he was then doing), if only because the project would undoubtedly grow into something much more complex, especially with the requirement to marry it to someone else's payload. Thus I turned down the opportunity to develop what became the first Alouette satellite—and a good thing, too, since it ultimately grew to consume much of the resources of DRTE, even after those resources had been expanded by DRB out of political necessity. (Had I been its godfather, no doubt the skeptics would have railed at my naivete and cancelled the whole project at its first demand for an order-of-magnitude leap in finances and personnel.) I was, nevertheless, trotted off by Scott, in company with Warren, to add "scientific respectability" when he made his first pitch to NASA, I believe early in 1958. The story of Alouette and its successor spacecraft has been surveyed, minus the foregoing details, by *Hartz and Paghis* [1982].

5. MY SCIENTIFIC REORGANIZATION

My scientific career underwent a major bifurcation at about the same time, following upon a totally unexpected phone call from Millett G. Morgan of Dartmouth College. He had undertaken to edit a special issue of the Proceedings of the Institute of Radio Engineers commemorating the International Geophysical Year, whose ending was then in sight. Among the papers he had commissioned was a review of motions in the ionosphere, with Martyn (the natural choice) as author, but Martyn had at a late date reneged. Morgan had been in a quandary for a suitable replacement, had consulted Scott (a long-time friend), and was now acting on Scott's advice that I should be invited. It was with some dismay that I accepted, for the schedule was tight by then and I still had my new job as Superintendent of RPL to get under control. My acceptance changed my scientific life.

I have described elsewhere [*Hines,* 1989] how preparation of the review led me back into studies of atmospheric gravity waves, now in a much broader context than I had previously considered. Those studies have constituted the principal focus of my subsequent scientific life. But a second area opened as well, leading in time to my paper with Ian Axford [*Axford and Hines*, 1961] on magnetospheric convection induced by interaction with the solar wind.

It began simply enough. The most basic motion of the ionosphere, requiring at least a passing comment in my review, was rotation. Oddly enough, I could find virtually nothing that had commented on this topic: it seemed to be taken simply as "a given." My search turned up only a remark by J. W. (Jim) Dungey in a rather obscure report [*Dungey*, 1954], to the effect that the entire ionospheric milieu would corotate with the earth, under the hydromagnetic constraint that all ionization on a geomagnetic field line must move together. (Or, as he was soon to say [*Dungey*, 1958], "If a rotating star (or planet) had no magnetic field, the surrounding gas would tend to rotate with the star, but the rate of rotation would be controlled by the viscosity and [the angular velocity] would decrease gradually with the distance from the star. When there is a magnetic field, on the other hand, the rate of rotation is controlled by the law of isorotation.") In saying so he was responding in part to the new reality imposed by Storey's discovery.

So far so good. But I was well aware that the hydromagnetic constraint could not impose itself through the underlying nonionized atmosphere—a fact subsequently impressed forcefully on the scientific community by Tom Gold as a preamble to his discussion of magnetospheric convection [*Gold*, 1959], but not yet in common currency. My own understanding derived from earlier work by Martyn. He had described the tidal generation of electric fields in the "dynamo" region of the ionosphere, their communication upward along equipotential geomagnetic field lines, and the driving of the F-region "motor" in response to those fields. This was nothing but a classical electrodynamic description of what was soon to be described with the convenience of hydromagnetic terminology: the overlying ionization was constrained to convect as if frozen to convecting field lines. And the field lines were free to convect (if one chose to think of them as being in motion—not at all a necessity) since they lost their individual identities (as defined by the ionization "frozen" on them) below the ionosphere. By the same token, the ionization was free to corotate with Earth, or not, depending on other forces acting on it: it was not in fact forced into corotation by hydromagnetic means.

It seemed clear that ionization on low-latitude field lines would corotate, since atmospheric viscosity would combine with ion-neutral collisions to tend to set both ends of these field lines into corotation, and there was no countervailing force. But the high-latitude field lines were a different matter: if they extended out into the interplanetary medium, their corotation would imply a corotation of that medium, indefinitely into space. Conversely, if that medium was reluctant to corotate with Earth, it would impose along the field lines an inhibition of corotation that would extend right down to the dynamo layer. The net effect would be some compromise between the forces imposed from without and the viscous forces imposed from within—a compromise whose consequences could not be forecast without further analysis. In my review paper [*Hines*, 1959], I simply pointed out the problem, but I developed it formally in a subsequent paper [*Hines*, 1960*a*].

The later history of this idea is sketched in Postscript 40 of *Hines* [1974]. Though the idea itself faded from sight, and may be of no observational significance, I believe that it must lie buried somewhere inside the numerical computations that now dominate the theoretical study of magnetospheric convection, its message overwhelmed by the more complicated patterns of motion that have come to be of greater concern. For me, however, its important consequence was that it thrust me into thinking of motions above the dynamo layer in terms of the "frozen fields" of hydromagnetics.

Hydromagnetic thinking was soon being thrust on the ionospheric community at large by Tom Gold, who took the long-delayed but highly appropriate further step of putting a name—the magnetosphere—to the extended upper region of geomagnetic control [*Gold*, 1959].

Gold, coming from an astrophysical background in which magnetic field lines could be identified once-for-all by the ionization "frozen" to them (save for "merging" processes), took pains to point out the breakdown of that concept imposed in Earth's case by its nonconducting atmosphere. Having freed his magnetosphere from freezing to the earth, he went on to suggest a certain class of diurnal variations that might be engendered by a form of thermal convection of the inner magnetosphere, with possible consequences to geomagnetic variations and trapped radiation. At the same time, he noted the possibility of similar motions at higher latitudes having consequences for aurorae.

Gold was fully aware of the availability of an equivalent classical description of such motions, but he seems to have been unaware that it had already produced, in the concept of an F-region "motor" driven by electric fields from the E-region "dynamo," frozen-field motions precisely of the type he was invoking. Indeed, he took no cognizance of the partially conducting dynamo region, where induced currents would have acted to inhibit the thermal convection he was postulating. Further, he paid no heed to the fact that some of the observations he was seeking to explain were already reasonably well described by dynamo theory. In consequence Martyn, who was preeminent in such studies, felt called upon to invite comment on the earlier work when Gold presented his views orally at the International Symposium on Fluid Mechanics in the Ionosphere [*Transactions*, 1959, p. 2083]. Gold responded with a neat segue, redirecting attention simply to the great vault of the magnetosphere throughout which motions of this class must now be recognized to occur.

Gold's idea of a thermally driven magnetospheric convection met a dead end: Martyn's dynamo effects, suitably updated as our understanding of tidal forcing has improved, have retained their dominance in the analysis of large-scale magnetospheric plasma motions on low- and mid-latitude field lines. But the encompassing concept of the magnetosphere had been established, the convenient concept of hydromagnetic convection within it had been put in place for use when later needed, and the prospect of further application at higher latitudes had been introduced.

(The term "convection" is often employed to imply thermal convection, as in Gold's usage. When Axford and

I came to write our 1961 paper, this usage bothered me, and for some time I considered using "advection" as a less prejudicial word. But I ultimately opted for "convection" on the grounds that its Latin root words might emphasize the concept of the field lines being carried with the ionization.)

6. INSTITUTIONAL RE-REORGANIZATION

My attention at this time was split between the publication of my comprehensive gravity-wave thesis [*Hines*, 1960*b*] and my commitment to the research program I had reorganized within RPL. That commitment was directed mainly toward a search for a similarly comprehensive framework on which we might in time synthesize the wide-ranging elements that constituted Earth's response to solar disturbances.

As one step in this process, I arranged for a two-day symposium at RPL in the summer of 1959, designed to bring together many of the foremost North American researchers in the field (e.g., Sydney Chapman, E. H. Vestine, J. W. Chamberlain; Parker, whose views on the ring current would have made for interesting discussion, was one of my hoped-for invitees but unfortunately was unable to attend). The conference went well, I believe, both scientifically and as an advertisement for RPL's new role in the field.

But, on the afternoon preceding its social evening, I was called into the office of Frank Davies, recently returned from DRB/HQ as Chief Superintendent of DRTE once again. He had taken to expounding on how he now, as never before, appreciated the responsibilities incumbent on a defense scientist, and how DRTE must play its part in providing staff for DRB/HQ on a rotating basis (as had been forced on him) and in supporting the military with research of immediate consequence to their needs. Basic research into the disturbed ionosphere did not fall into that category, in this new view. Moreover, I was too much oriented toward basic research and too little of a management type to meet his needs, he said. (He reached this latter judgment despite also arguing that, of all the initiatives undertaken during Scott's tenure, including the Alouette satellite, the only ones appropriate to DRTE were the formation of the rocket program and of the Synoptic Studies Group! He later recanted his adverse judgment of me as a manager, but too late.) Accordingly, he invited me to step aside while he replaced me as Superintendent of RPL with Paghis, who had served him well in earlier years at EL. He would ensure that I would be given suitable status, and perhaps even one or two people to form a separate theoretical group within DRTE—a small bauble for me that would not get in the way of his over-all plan. I protested, but to no effect: he had his mind set, and that was that.

Making no secret of my dismay once the symposium had been completed, I took a week's leave of absence, in part to scout employment possibilities at the University of Toronto. On the morning of my return, upon boarding the DRTE bus I was pulled aside by Davies, who turfed out his seat-mate and obliged me to sit beside him. He announced that A. Hartley Zimmerman, Chairman of DRB and therefore his boss, had instructed him to keep his hands off me and to leave me and RPL in peace; he would, of course, comply. (I suppose, but do not know, that Paghis had got word through to Zimmerman as to Davies' plans and my reaction.)

This was a happy indication that RPL had already established credentials sufficient to its needs, but the entire incident left a bad taste. Moreover, in the months to come, it became clear that the Alouette topside sounder, no longer a baseball-sized instrument on someone else's vehicle but a full-blown satellite in its own right, was taking over EL and making financial and personnel demands within DRTE (and DRB, for that matter) that could not be resisted: RPL would be, if not strangled by its growth, at least prevented from taking new initiatives of its own.

The only exception would be in the purely theoretical area: Davies' idea of a theoretical group within RPL had by now been embraced wholeheartedly by Zimmerman, and I was asked to establish such a group as a separate unit of DRB, as an additional responsibility. Six positions were to be made available initially, more if things worked well. It was to be located at DRTE but responsible directly to the Chairman, who had visions of it developing in time into a Canadian equivalent of the Institute of Advanced Studies in Princeton. Thus arose the DRB Theoretical Studies Group (DRB/TSG), initially comprising me as Head, a reluctant Fejer (who preferred and indeed maintained direct contact with RPL experiments) and Stan Mack, a scientist of some seniority who had served his time in DRB/HQ and was being allowed to "recharge his batteries" (as a popular saying then went) by conducting certain studies in general relativity that were close to his heart. Reid, though invited to join, had already made plans to move on to NBS in Boulder and chose not to change them.

Thus things continued for much of the ensuing year. In the end, however, I decided that my long-term hopes for RPL must be abandoned and that I should redirect my organizational energies, now in decline, solely to the DRB/TSG, which would take on a visibly separate existence outside of RPL. I terminated my RPL position at the end of June, 1960, and went off to symposia and a holiday in Europe to wash my mind of my RPL experience. Ironically, it was in the course of this trip that I was led into the model of high-latitude magnetospheric convection that would provide the comprehensive framework I had hoped to find, the one on which RPL's studies of the disturbed ionosphere might be expected to be synthesized.

7. THE 1961 AXFORD-HINES PAPER

The first of my symposia was in Kiruna, Sweden, and was directed specifically at disturbed conditions. Among the presentations was one made on behalf of *T. Neil Davis* [1960], who was unable to be present himself. It concerned motions of auroral structures observed in Alaska: westward pre-midnight and eastward post-midnight, as had been widely reported previously by others, but now joined by equatorward in the midnight region at somewhat higher latitudes. The whole picture, as I sketched it in my program, had the appearance of a sweep across the polar regions, followed by bifurcation of the flow at auroral latitudes. I could not resist the urge to complete the pattern in my notes, making the auroral streamlines—if that was what they were—coalesce again on the (unobserved) sunward side, and producing two loops of apparent circulation.

I was conditioned at the time to a "closed" magnetospheric model. It had been presented by *Francis S. (Frank) Johnson* [1960], in part in response to my own arguments (outlined above) concerning inhibition of rotation if the polar field lines extended indefinitely into interplanetary space. Johnson argued that those field lines would instead be carried into a closed geomagnetic tail and so would be free to rotate with the earth (or to counterrotate, at their equatorial crossing), contrary to my suggestion. With the thought that Davis' pattern of motion might represent the low-altitude end of a field-line convection, I mapped that pattern up the arching field lines and down into the equatorial section of the tail, producing the sketch shown here as Figure 1, in tidied-up form.

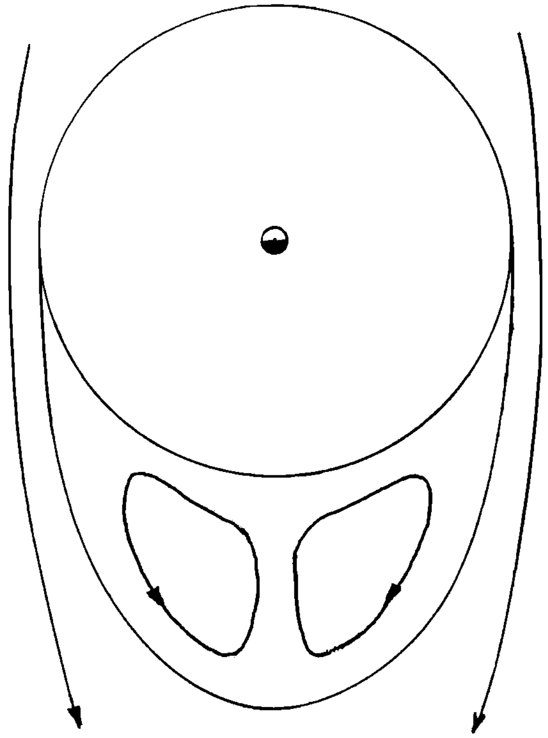


Figure 1.

I had also been recently exposed (by Gold and Hermann Bondi, in response to a cocktail-party query I had put to them) to the idea that the front-to-back asymmetry inherent in Johnson's model must imply some sort of viscous-like interaction between the solar wind and the magnetospheric plasma. And now my sketch was showing just the pattern of convection that would accord with such a process: convection directed away from the sun on the outer flanks, where any such interaction would be imposed. The magnetosphere was, to me, suddenly alive with new-found motion; and that motion—yet to be combined with rotation—would affect any number of processes at ionospheric heights that had previously maintained separate identities!

In the course of my European jaunt, while visiting friends in Cambridge, I met a young officer of the Royal New Zealand Air Force, W. Ian Axford. He was about to be seconded to RPL for a period, his superiors having found nothing better to do with him. (He had been aimed in my direction at RPL by Scott, who was by this time serving as DRB liaison officer in London. The circumstances were as I have described them elsewhere [*Hines*, 1986].)

Axford had recently completed a Ph.D. degree at Manchester, on shock waves in interstellar plasmas, and was being given some exposure to things ionospheric in Ratcliffe's group. Knowing nothing of him in advance, other than his militaristic background (howsoever that might have been tempered by his stay in Manchester), I had been intending to leave him to his fate in RPL; but this brief meeting was sufficient to convince me that I should pry him loose, as a recruit to DRB/TSG, once I got back to Ottawa. Paghis, who had succeeded me as Superintendent of RPL, resisted such a move but at last acquiesced on condition that I would "raid" RPL no further.

Even before the transfer was formalized, Axford began working with me in developing the convection idea and its extensive implications for the ionosphere, aurorae, trapped radiation, geomagnetic storms, and so forth. As to the basic model, he was soon arguing that, though I had drawn the impressed pattern as if it was confined to the counter-rotating tail, the low-latitude corotating part of the magne-

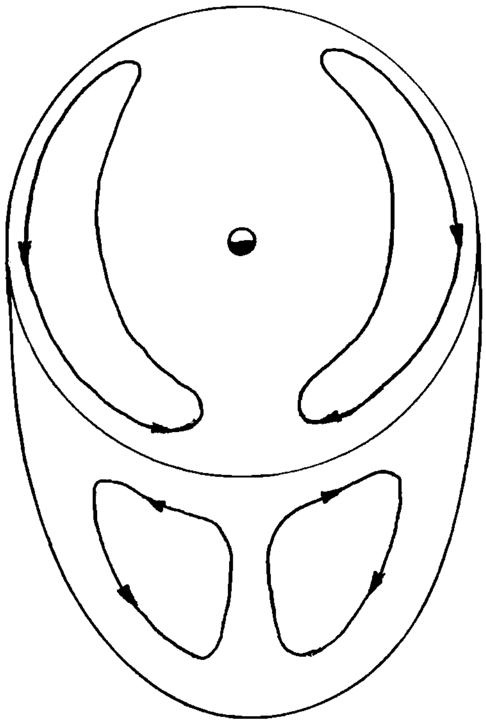


Figure 2.

tosphere would undergo a similar convection, similarly induced by the hypothesized viscous-like action on its own flanks. For several weeks, we were sketching a four-celled convection as in Figure 2. Indeed, this was the point we had reached when I went, with Fejer and Reid, to the US-URSI meeting in Boulder a few months later (December 1960, if I recall correctly).

One of the sessions there, chaired by Alan H. Shapley, ran out of material before its time was up, and Shapley asked if there were any unscheduled topics of interest that might usefully fill the remaining time. I took the opportunity to sketch the model we had developed, complete with its four convective cells, in what constituted our first public announcement of what we were onto. Since the four-celled convective pattern implied a four-celled pattern of Hall currents (and associated Pedersen currents, closed by field-aligned Birkeland currents) and a corresponding Ds geomagnetic variation, I enquired as to whether any such four-celled patterns of currents were known to anyone present. They were not, a fact that left me in a quandary as to whether the lower-latitude convection cells would really exist. (Four-celled patterns have since been reported, but I believe them to have no relationship to ours except possibly as transient phenomena.)

I ultimately resolved this quandary in mid-flight a month later, on a gravity-wave trip, when I realized that the line I had first drawn to distinguish the counter-rotating tail from the corotating inner magnetosphere was but an imaginary figment of magnetospheric book-keeping. Even if there were some initial tendency for two distinct pairs of convection cells to be established, it should soon be overcome by viscous-like forces at the figmentary boundary: the two pairs of cells would simply merge into a single pair, and so Figure 3 was born. This, with rotation added as in Figure 4, completed our modeling of the convection itself. One might well ask, in retrospect, how it took so long for us to reach this point, but that is of course the nature of research.

Simultaneous with these developments, and following on them, were our integration of a variety of observed phenomena into the picture and our inference of other processes that had to be implicit (such as perturbations in the motion and distribution of trapped radiation, the energization and precipitation of auroral-energy particles, and acceleration of the neutral gas at ionospheric heights, which in turn would modify the current systems of the dynamo region). We also broadened our starting point, with the recognition that any mechanism that would set up the two-celled circulation would serve all our ultimate purposes. *Dungey's* [1961] mechanism of field-line merging and di-

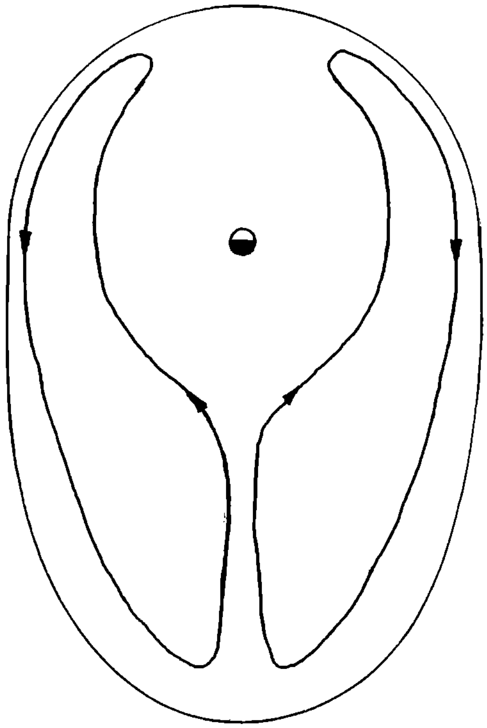


Figure 3.

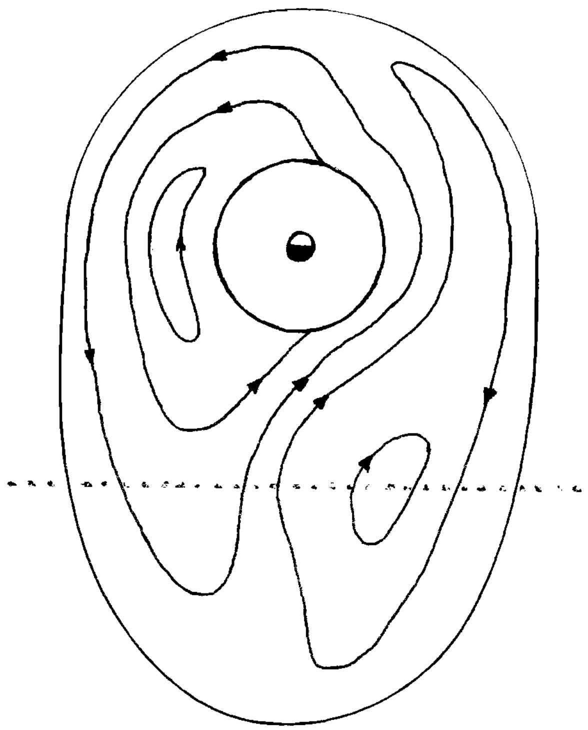


Figure 4.

rect driving by the solar wind, of which we learned during this period, became our ally as a likely contender.

(*Fejer* [1961] produced yet another mechanism, though I believe it would have been transitory and appropriate to storms only [*Hines*, 1962]. By the time Axford and I came to publication, we were convinced that our model was as important in quiet times as in disturbed times; and so, while we mentioned Fejer's mechanism, we did not incorporate or otherwise adopt it. Both Fejer and Reid—for different reasons—came close to being invited to appear as co-authors, in what would then have been a Hines et al. paper; but this would have done an injustice to Axford's role, and so I dropped the thought. Upon then adopting alphabetical ordering of Axford's name and mine, I opened the way to a reverse injustice: J. H. Piddington—who may have felt slighted by our comments on his work and would have attributed the slight to me—in one paper referred to our work as "Axford et al." This seems to have titillated Michael Gadsden, Secretary-General of IAGA in the late 1980s, who invited me to reminisce in the IAGA News on the origins of the Axford-Hines paper, conditional on my permitting him to identify the author as "Al." This I did, thereby leading to the forerunner of my *Eos* reminiscence [*Hines*, 1986].)

The first formal presentation of our theory took place in the early months of 1961, at the annual DRB symposium. This symposium was designed to permit DRB scientists to present for discussion even their highly classified research and was therefore conducted under deep security wraps. Zimmerman insisted that I should make the presentation—his first showcasing of the DRB/TSG—and so I did. Forsyth, who was one of the invitees, had been asked to serve as discussion leader. He made highly flattering remarks: flattering not only to Axford and me for our work, but to DRB for permitting and supporting such a high calibre of basic research. I am sure that Zimmerman could not have been more pleased.

Soon after that, I was invited by Henry G. Booker and Gold to Cornell, to present our work at a seminar. I had Axford join me, thereby initiating contacts that led him to appointments at Cornell and in due course La Jolla, California.

Our first formal presentation to the scientific world at large was made by Axford the following spring at the Washington meeting of the American Geophysical Union (AGU). At the URSI meeting five months earlier, William E. Gordon had, in concert with Booker, initiated plans whereby AGU would lay claim to the magnetospheric realm. The mechanism was to be a special session of invited speakers. A suitable set of speakers—including me, speaking on my long-ago union of magnetoionic and hydromagnetic theory—had been arranged. But it became clear to me, as Axford and I continued our development following the URSI meeting, that the session would be remiss if it did not include our work on convection. I proposed to Gordon that he invite Axford for the purpose, which he did. Axford made a masterful presentation, and our synthesis was almost immediately assimilated into the magnetospheric thinking in North America. A few months later, I performed the corresponding function on the international scene at the International Conference on Cosmic Rays and the Earth Storm in Kyoto, Japan.

Journal publication of our work, when it came [*Axford and Hines,* 1961], was almost anti-climactic (or, in truth, post-climactic). The big picture was already accepted, and a race to fill in the details was already begun.

It often seems to me nowadays that our journal publication must have gone virtually unread. Much of its message—the basic patterns of convection, the associated electric fields, and their various consequences—became so much a part of the "revealed truth" of magnetospheric studies that our original formulation was ignored, while much of the detailed development was lost from sight only to be reformulated later by others, as if for the first time, often with a change of nomenclature that disguised the equivalence.

Whether or not this impression is valid, it is certainly true that our paper is now cited most often for its alleged assertion that a viscous interaction drives the convection, and as an alleged alternative to the Dungey "merging" mechanism. Anyone who takes the trouble to actually read our paper will find that this is but a perversion of the truth. We were careful to write only of "a viscous-like component" of the interaction, and we specified that "By this we mean simply that some of the momentum of the solar wind is transferred across the boundary of the magnetosphere to the ionization within. The nature of this momentum transfer is, for present purposes, of minor importance; its existence, or the existence of an equivalent mechanism, is crucial." We included, explicitly, reference to Dungey's as one of the mechanisms that would serve our purpose. And we said, "Regardless of its manner of generation, however, the convective system we discuss has consequences of far-reaching import, and it is these that we wish to emphasize in the present paper."

Nevertheless, current discussion would lead to the impression that one must choose between Dungey on the one hand and Axford and Hines on the other, and that the choice to be made is that of mechanism alone. But, as one now-prominent member of the geomagnetic community said to me some time after hearing my complaint on this score, "I went and actually read your paper. You guys didn't say at all what people say you said!"

Lost in the debate over mechanisms is what we actually said—what our actual contribution to the field was: that of providing the basic structure on which future assimilations would be built, and taking some steps toward making and anticipating such assimilations. I feel that we produced, not a paradigm shift, but a new paradigm where none had existed before.

(Axford—who, incidentally, embraced Dungey's mechanism more strongly than I did at the time—takes a more philosophical view than I am now displaying, holding that people forget who did what in any event. But, in a survey of the history of the magnetosphere such as this, I feel obliged to state the history as I see it.)

8. THE AFTERMATH

Axford and I joined in only two further papers, of a relatively trivial nature [*Axford and Hines,* 1962, 1964]. With his Ph.D. work as background, he (and independently P. J. Kellogg) soon predicted the formation of a shock front on the sunward side of the magnetosphere [*Axford*, 1962], one whose reality is now fully accepted. And, after leaving Ottawa, he went on to study, with students, many further aspects of magnetospheric and general plasma theory which need no documenting by me. My own further research in the area was limited to a reconciliation of the hydromagnetic and particle-drift descriptions of energization [*Hines*, 1963*a*], a topic left somewhat cloudy by Gold through his assumption of a single scalar ratio of specific heats (rather than distinct values for compression along and across magnetic field lines).

In the spring of 1962, things began to unwind. Axford was soon to depart for New Zealand in response to his Air Force obligations. Fejer, whose wife had slipped on a patch of Ottawa's winter ice and broken an arm, was being recruited away to sunny California. Stan Mack, whose relativistic research had come to naught, was returning to harness as a real defense scientist (in Victoria, British Columbia), and I had a choice of appointment offers open to me. With the DRB/TSG on the point of collapse, I made one further effort, not magnetospherically related, to keep it in being.

I decided to test Zimmerman on his long-term intent with respect to an institute. I obtained applications for employment from both Ivor Robinson (a distinguished relativist and cosmologist, then at Cornell; more recently at the Southwest Center for Advanced Studies, University of Texas at Dallas) and the somewhat younger Roger Penrose (now Rouse Ball Professor of Mathematics at Oxford, well known—probably to all readers of this account—for his tiling theory, his cosmological studies with Stephen Hawking and his recent book, *The Emperor's New Mind*). It would have been difficult to start a Canadian Institute of Advanced Studies on a higher plane.

Zimmerman responded well, taking a personal interest in moving their appointments forward over bureaucratic objections. But the latter bothered me because they revealed the tenuous nature of the support available within DRB for an institute: we might be but a heartbeat away from having the whole thing collapse. To test that I was not misreading the situation, I requested and was granted an interview with Zimmerman and his own Management Committee, including the Vice Chairman (who was really the man responsible for the internal operations of DRB), the Chief Scientist, the Chief of Personnel, the Chief of Security, and the Chief of Establishments. I asked for clarification of their perceptions of the mandate granted me and the TSG.

Zimmerman—not a scientist himself but rather an industrial engineer, however visionary—made a few pious remarks but otherwise kept rather quiet. One of the others fell asleep (the meeting was held just after lunch). Another suggested the TSG was intended to be a way-station for that favored project of the day, recharging the batteries of various DRB scientists as they moved through it from one more important position to another. The discussion did not get much above this level, and I felt all my enthusiasm—or what was left of it—drain away.

My final thrust was made by proposing that NRC and Atomic Energy of Canada, at nearby Chalk River, Ontario, be invited to join with DRB in forming a truly national TSG—or Institute of Advanced Studies—and that it be relocated onto the campus of Carleton University in Ottawa. (The barbed-wire fences that surrounded DRTE did not seem to me conducive to recruiting the sort of people I had in mind.)

This proposal was considered in due course by the Management Committee and rejected. I was informed later (by Scott) that the Vice Chairman had said, "Hines almost pulled it off. If only he had proposed that NRC join in, we probably would have gone for it." Such was the management team on which I was dependent!

Upon the rejection of my proposal, I had to inform Robinson and Penrose that the deal was off: I was moving to the University of Chicago, there to lick my wounds. In September, 1962, the DRB/TSG came to an end. Later that same month, Alouette 1 was launched. It performed magnificently, giving a major focus to DRTE over the ensuing years and providing the start of Canada's long-term space program. I cannot begrudge it its success.

At the University of Chicago, my research continued along its bifurcated path. Problems associated with gravity waves seemed to suit both my students and me more than did those associated with the magnetosphere, so they got the bulk of my attention. I nevertheless did have three excellent post-docs (Atsuhiro Nishida, Fritz Neubauer, and Norbert Skopke) who addressed magnetospheric problems and went on to make major contributions to the field, to the research programs of their countries of origin, and even internationally. There were also visiting scientists, notably (in the magnetospheric context) Henry Rishbeth and Keith D. Cole, each of whom spent an extended period with me and my group.

Beyond this, my further contributions were primarily of a tutorial nature designed to disseminate concepts of magnetospheric convection in greater depth and to a wider audience. They are listed as part of the Appendix to my 1974 collection of papers [*Hines*, 1974], in which collection the interested reader will find further bits of historical reminiscence. My own favorite of these papers [*Hines*, 1964] contains what I believe to be the most compact proof of the "frozen field" theorem of hydromagnetics, describes in detail the means by which two superimposed convective systems (e.g., rotation and the Axford-Hines convection) may be combined accurately graphically—a procedure I had employed in our 1961 paper but one that has been rendered redundant by the growth of numerical analyses—and integrates into a hydromagnetic context the driven motions of Martyn's long-ago F-region "motor." As updated in Postscript 3 of the collected papers, it also describes how

Colin Hines during his stay at the University of Chicago, 1962-1967.

the closed-magnetosphere illustrations of the Axford-Hines paper can be employed in application to an open magnetosphere, namely by (mentally) folding the "tail" portion of the diagram upward (and equally downward) out of its original plane, along the dotted line in Figure 4 here, to represent a cross-section of the open tail. I mention this because one prominent member of the geomagnetic community, who had not recognized the equivalence, rhapsodized once it was revealed; others may feel likewise.

I made my own final comment on driving mechanisms in an invited review on the magnetopause [*Hines*, 1963b], when I suggested that the Dungey mechanism might well dominate at times of magnetic storms but might yield some of its prominence to other processes of momentum transfer at magnetically quiet times. I believe that this view is still held by some.

Then, having seen the magnetosphere not only born and named but brought into an excitingly active life, I turned my research efforts fully to other concerns, all the while watching its development to maturity with continuing interest.

REFERENCES

Axford, W. I., The interaction between the solar wind and the Earth's magnetosphere, *J. Geophys. Res.*, *67*, 3791-3796, 1962.

Axford, W. I., and C. O. Hines, A unifying theory of high-latitude geophysical phenomena and geomagnetic storms, *Can. J. Phys.*, *39*, 1433-1464, 1961.

Axford, W. I., and C. O. Hines, On the thinness and orientation of auroral arcs, *J. Geophys. Res.*, *67*, 2057-2058, 1962.

Axford, W. I., and C. O. Hines, Comments on "A hydromagnetic theory of geomagnetic storms" by J. H. Piddington, *Planet. Space Sci.*, *12*, 660-661, 1964.

Davis, T. N., The morphology of the polar aurora, *J. Geophys. Res.*, *65*, 3497-3500, 1960.

Dungey, J. W., *Electrodynamics of the Outer Atmosphere*, Scientific Rept. No. 69, Ionospheric Research Lab., Penn. State Univ., 1954.

Dungey, J. W., *Cosmic Electrodynamics*, Cambridge University Press, 1958.

Dungey, J. W., Interplanetary magnetic fields and the auroral zone, *Phys. Rev. Lett.*, *6*, 47-48, 1961.

Fejer, J. A., The effects of energetic trapped particles on magnetospheric motions and ionospheric currents, *Can. J. Phys.*, *39*, 1409-1417, 1961.

Gold, T., Motions in the magnetosphere of the Earth, *J. Geophys. Res.*, *64*, 1219-1224, 1959.

Hartz,, T. R., and I. Paghis, *Spacebound,* Ministry of Supply and Services Canada, 1982.

Hines, C. O., Heavy-ion effects in audio-frequency radio propagation, *J. Atmos. Terr. Phys.*, *11*, 36-42, 1957a .

Hines, C. O., On the geomagnetic storm effect, *J. Geophys. Res.*, *62*, 491-492, 1957b.

Hines, C. O., Motions in the ionosphere, *Proc. Inst. Radio Eng.*, *47*, 176-186, 1959.

Hines, C. O., On the rotation of the polar ionospheric regions, *J. Geophys. Res.*, *65*, 141-143, 1960a.

Hines, C. O., Internal atmospheric gravity waves at ionospheric heights, *Can. J. Phys.*, *38*, 1441-1481, 1960b.

Hines, C. O., Ionospheric disturbances at auroral latitudes, *J. Phys. Soc. Japan 17*, *Supp. A-I*, 308-313, 1962.

Hines, C. O., The energization of plasma in the magnetosphere: Hydromagnetic and particle-drift approaches, *Planet. Space Sci.*, *10*, 239-246, 1963a.

Hines, C. O., The magnetopause: A new frontier in space, *Science*, *141*, 130-136, 1963b.

Hines, C. O., Hydromagnetic motions in the magnetosphere, *Space Sci. Rev.*, *3*, 342-379, 1964.

Hines, C. O., *The Upper Atmosphere in Motion*, *Geophys. Monogr. Ser.*, vol. 18, AGU, Washington, D.C., 1974.

Hines, C. O., Origins of the 1961 Axford-Hines paper on magnetospheric convection, *Eos Trans. AGU*, *67*, 634, 1986. (Reprinted in *History of Geophysics*, vol. 4, edited by S. Gillmor, pp. 97-98, AGU, Washington, D.C., 1990.)

Hines, C. O., Earlier days of gravity waves revisited, *Pure App. Geophys.*, *130*, 151-170, 1989.

Hines, C. O., I. Paghis, T. R. Hartz, and J. A. Fejer, eds., *Physics of the Earth's Upper Atmosphere*, Prentice-Hall, 1965.

Hines, C. O., and E. N. Parker, Statement of agreement regarding the ring-current effect, *J. Geophys. Res.*, *65*, 1299-1301, 1960.

Hines, C. O., and L. R. O. Storey, Time constants in the geomagnetic storm effect, *J. Geophys. Res.*, *63*, 671-682, 1958.

Johnson, F. S., The gross character of the geomagnetic field in the solar wind, *J. Geophys. Res.*, *65*, 3049-3051, 1960.

Storey, L. R. O., An investigation of whistling atmospherics, *Phil. Trans. Roy. Soc. (London)*, *A246*, 113-141.

Transactions, Transactions of the International Symposium on Fluid Mechanics in the Ionosphere, *J. Geophys. Res.*, *64*, 2041-2091, 1959.

Ray Tracing Technique Applied to ELF and VLF Wave Propagation in the Magnetosphere

Iwane Kimura

Faculty of Information Science, Osaka Institute of Technology,
1-79-1 Kitayama, Hirakata-city, Osaka 573-01, Japan

The ray tracing technique which has contributed to the study of whistler mode wave propagation in the magnetospheric plasma is reviewed, to stress its significance in the history of the discovery of the magnetosphere. It has been closely associated with plasma distribution surrounding the earth, later called the exosphere, plasmasphere, and magnetosphere. Many whistler mode wave phenomena observed on the ground and on-board satellites have been clarified by the analyses of ray tracing. In the first part of this article, the works made mainly before 1975 were described from the point of view of the history. In the second part of this article, the author attempts to describe how the ray tracing technique has been revised since 1975. First of all, to do more accurate ray tracing a more realistic geomagnetic field configuration than the dipole model has been adopted. The ray tracing of ion cyclotron waves and in hot plasmas are the further development. The ray tracing technique has been used for the exploration of plasma environment by means of various VLF/ELF wave phenomena. However, this technique is an inverse problem, because we have to give plasma density distribution in advance for doing ray tracing. Recently there has been developed a very positive method by using the ray tracing technique tomographically to determine the global electron density distribution at least in the plasmasphere. This is one of the most significant applications of ray tracing technique related to the magnetosphere.

1. INTRODUCTION

Though the discovery of whistling atmospherics or whistlers dates back to the 19th century [*Helliwell,* 1965], the theoretical clarification of the whistlers related with the discovery of the magnetosphere was made by Storey's work published in 1953.

He showed that from Eckersley's theory, or Appleton's equation of the magneto-ionic theory, VLF waves whose frequency below the electron cyclotron frequency (f_H), can propagate nearly along a background geomagnetic field line with group velocity much slower than the light velocity. For frequencies much lower than f_H, the group velocity is inversely proportional to the plasma frequency (f_p) and is proportional to the square root of

Discovery of the Magnetosphere
History of Geophysics Volume 7

the product of f_H and the wave frequency (f). A low frequency part of wave energies generated from lightning flashes penetrates the ionosphere and it is guided by the geomagnetic field and comes back to the opposite hemisphere, if there are enough electrons in the space where the wave propagates. The deviation of the wave energy from the direction of wave normal and also the characteristics of dispersion are due to the anisotropic nature of magnetized plasma. In general, the mode of the right-hand polarized electromagnetic waves with frequencies lower than both the electron cyclotron frequency and the plasma frequency, is called the whistler mode.

According to Storey, for the existence of whistling atmospherics the space surrounding the earth at least up to several earth's radii must be filled by a sufficient concentration of electrons (naturally together with ions), much higher than so far expected, and the dispersion characteristics or delay time characteristics of the phenomena would give rise to the valuable information of plasma concentration in the space surrounding the earth. This was a sensational point, the discovery that the space surrounding the earth is filled by plasma. Therefore it could be said that the so-called space age started from 1953, when Storey's work was published.

Following the initial work of whistlers by Storey, observations of whistlers have been widely made all over the world and soon after, satellite observations were added from 1959 [*Cain et al.*, 1961]. Theoretically on the other hand, ray tracing study for the ray paths of whistlers was started by Maeda and Kimura in 1955 and the first result was published in 1956.

By the way, Maeda was a graduate of Kyoto University, and had been working in research laboratories of the government, where he had mainly investigated, for long, topics associated with radio wave propagation in the ionosphere. In 1953, he was invited to become a professor of Kyoto University in charge of radio wave communication engineering. In 1954, Maeda, for the first time, took care of 8 senior students for graduation research. Just before this time, Maeda had learned, at an international meeting, Storey's work on the interpretation of whistling atmospherics. He felt it very interesting, so that he assigned this topic to one of his students, named Kimura. This was the start of Kimura's connection with whistlers and magnetospheric study.

In several months Kimura wrote a graduation thesis, which was just a report, explaining the theoretical part of Storey's paper. After graduation, he entered the graduate course for master's degree, where he continued the investigation of the whistlers, and started to develop a method of ray tracing for whistlers under the guidance of Maeda as mentioned previously.

In the following of the present paper, the initial phase of the works using ray tracing technique nearly over 10 years from the start of ray tracing is first reviewed. In corporation with the results of ground and satellite observations the ray tracing technique has contributed much to the clarification of plasma concentration in the plasmasphere and magnetosphere. The distribution of plasma concentrations in the plasmasphere and magnetosphere is directly concerned with ray tracing of waves propagating in the space, which is therefore very briefly described in the subsequent section of the paper. Various VLF and whistler mode wave phenomena observed by mainly satellite borne receivers were interpreted by the results of ray tracing. Such typical phenomena are also described.

In the second part of this paper the recent works are very briefly described in the sense of how the ray tracing technique has developed and contributed to advanced and more sophisticated purposes associated with the physics of the plasmasphere and magnetosphere. This includes, for example, a revision of the background geomagnetic field for ray tracing from the dipole to the IGRF nondipolar model, ray tracing technique for the ion dominant modes, the effects of plasma temperature on ray paths, and finally a tomographic application of the ray tracing technique for the exploration of a global electron density distribution in the plasmasphere.

The works using the ray tracing technique were also reviewed in the past by the present author [*Kimura*, 1985; 1989].

2. RAY TRACING STUDIES OF WHISTLERS

In his paper, Storey described short whistlers, long whistlers and whistler trains; one hop, two hop and multiple hop echoes of whistlers respectively. For long whistlers and whistler trains to be detected at the source point, the ray path from the source hemisphere to the opposite hemisphere should be symmetric with respect to the geomagnetic equator. The initial motivation of Maeda and Kimura's ray tracing for whistlers [1956 and 1959] was to know how can the ray path of a whistler be symmetric with respect to the geomagnetic equatorial plane.

In an anisotropic medium like a magnetized plasma, the wave normal direction is refracted by generalized Snell's law due to inhomogeneity of the medium, so that a change of the instantaneous direction of the wave normal is determined by the spatial derivatives of the medium parameters. On the other hand the ray or the group velocity is always directed perpendicular to the refractive index surface at the direction of the wave normal. For ray tracing we have to trace simultaneously the change of the ray direction as well as the change of the wave normal direction. This principle of the ray tracing technique in an anisotropic medium was studied by Haselgrove [1955] to derive the first order differential equations in the form to be suitable for numerical integration using an electronic computer.

Maeda and Kimura, however, performed their ray tracing by a graphical method, based on Fermat's principle, instead of the above mentioned numerical integration, because no electronic computer was available at that time. Although the method was not so accurate, a characteristic of whistler ray paths was clarified. For example, for the sources located at relatively low geomagnetic latitudes, the ray path arrives in the opposite hemisphere at a latitude higher than the source latitude (so-called polar creep). On the other hand, for the sources located at high latitudes, the ray paths become more symmetric with respect to the geomagnetic equator, although the ray paths are still deviated appreciably from the geomagnetic field lines (see Figure 1).

In 1961, Yabroff made the first ray tracing by using a digital computer, doing numerical integration of first order differential equations derived by Haselgrove [1954].

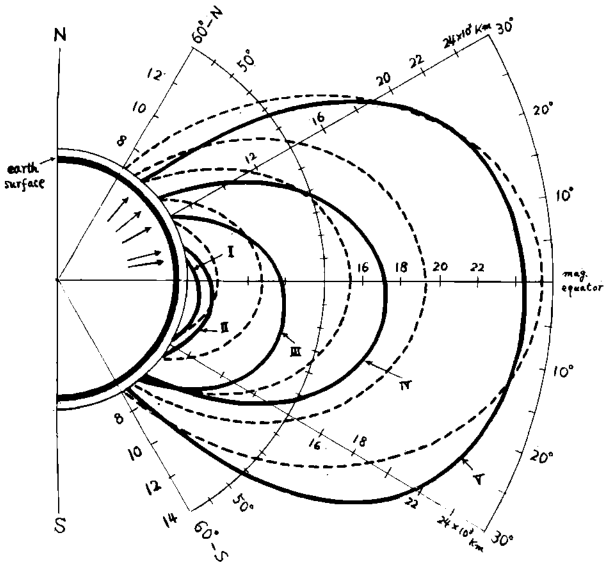


Figure 1. One example of ray paths for several VLF frequencies calculated by Maeda and Kimura [1956].

Her calculation confirmed the asymmetric ray path characteristics, found by Maeda and Kimura [1956; 1959], especially the polar creep for whistlers starting from relatively low latitudes and more symmetric ray paths for relatively higher latitude sources, although there was some discrepancy in detail between the results of Maeda and Kimura's and those of Yabroff, which was ascribed to an accumulation of error in the former's calculation.

Considering the asymmetric ray paths calculated based on so-called magneto-ionic theory and the existence of whistler trains whose ray paths should be symmetric with respect to the geomagnetic equatorial plane, there must be a mechanism which makes the ray paths more symmetric than the results of ray tracing. It was the multi-path nose whistlers that stimulated Smith's study of field aligned ionization ducts [*Helliwell et al.,* 1956], which enable the whistlers to propagate along a field aligned symmetric ray path [*Smith,* 1960; 1961; *Scarabucci and Smith,* 1971]. This mechanism is just the same as the trapped loss-free propagation of a light beam through an optical fiber cable. No clear experimental evidence of the existence of the field-aligned ducts has so far been clearly detected, from the point of view of a direct measurement of electron density enhancement along the geomagnetic field line in the plasmasphere. The existence has, however, been indirectly proved later by a careful examination of ducted and nonducted whistlers observed by a satellite. According to OGO-1 data analyzed by Smith and Angerami [1968], the equatorial separations between ducts ranged from 50 to 500 km, and the equatorial thicknesses were about 400 km.

Kimura went to Stanford University in 1964, where he developed a three dimensional ray tracing computer program [*Kimura,* 1966] for the aim to study the effect of ions on whistler propagation in the exosphere. He used, for the first time, a so-called diffusive equilibrium (DE) model, which was the newest physical model, at that time, of the exosphere (inside the plasmasphere) developed at Stanford University by Angerami and Thomas [1964].

Kimura's first attempt at applications of his ray tracing program was to confirm Smith's idea [1964] on the subprotonospheric (SP) whistlers, which were considered to be reflected back around an altitude of 1000 km by crossing a geomagnetic field line. He could confirm Smith's idea and multiple echoing SP whistler paths could be produced by assuming an existence of a latitudinal gradient as well as a negative gradient in the radial direction of electron density in the ionosphere. Figure 2 (a) and (b) illustrate one example of SP whistlers ob-

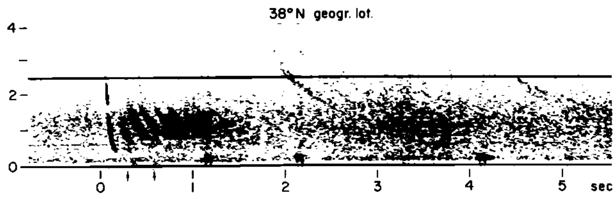


Figure 2a. SP whistlers observed by Alouette satellite [*Carpenter et al.*, 1964].

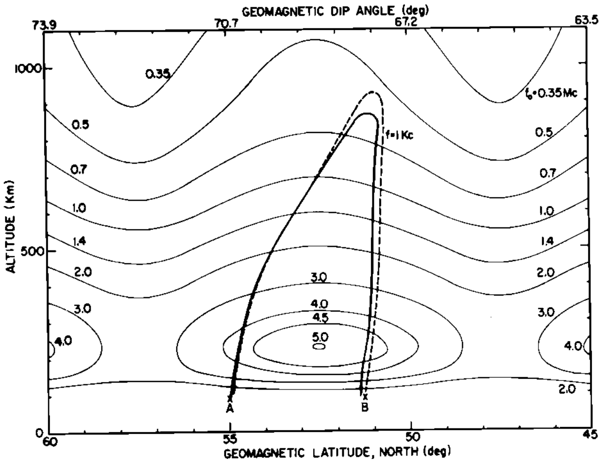


Figure 2b. Kimura's result of ray tracing [*Kimura*, 1966].

served by Alouette satellite [*Carpenter et al*, 1964] and Kimura's result of ray tracing, respectively.

A reflection of waves due to the wave energy crossing perpendicularly across the geomagnetic field lines takes place due to a closed refractive index surface with respect to the wave normal angle, when the wave frequency becomes smaller than the local lower hybrid resonance (LHR) frequency. This fact was first pointed out by Hines [1957]. Such a reflection of the ray paths is called 'LHR reflection'.

Kimura also found that whistler waves can be multiplically reflected outward by the LHR reflection mechanism when the wave propagates down from above and the local LHR frequency becomes higher than the wave frequency. Special species of whistlers which were produced by such a multiple LHR reflection, were detected by the OGO-1 satellite, and was called 'magnetospherically reflected (MR)' whistlers [*Smith and Angerami*, 1968]. Figure 3 (a) and (b) show one example of the paths calculated by ray tracing [*Kimura*, 1966] and actual dynamic spectrum of MR whistlers observed by OGO-1 [*Smith and Angerami*, 1968], respectively.

Later, this MR reflection was analytically described by Lyons and Thorne [1970], and they found that whistler mode waves are accumulated within a low latitude region along a magnetic field line, where the local LHR frequency is close to the wave frequency. They considered that such an accumulation of VLF wave energy can be a cause of relatively steady hiss band almost continuously observed within the plasmasphere by Dunckel and Helliwell [1969] and Russell et al.[1969].

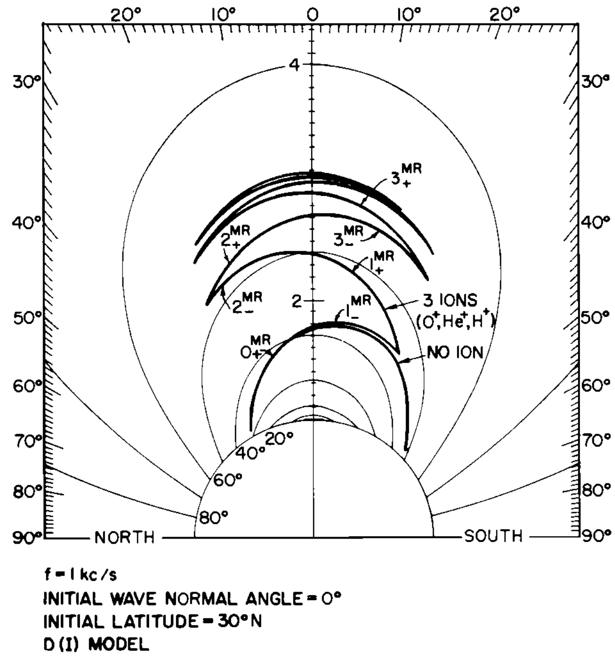


Figure 3a. MR reflections disclosed by ray tracing [*Kimura*, 1966].

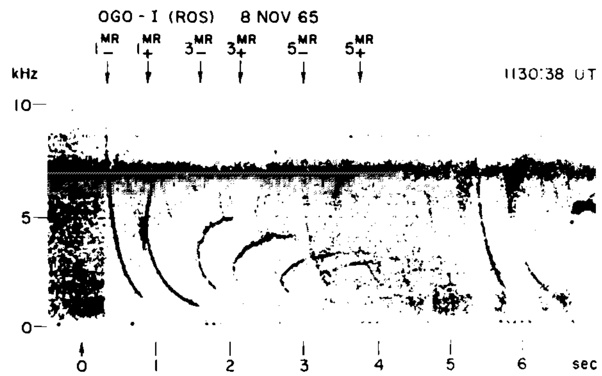


Figure 3b. MR whistlers observed by OGO-1 [*Smith and Angerami*, 1968].

3. MODEL OF THE ELECTRON DENSITY DISTRIBUTION

In the earlier stage of ray tracing, the electron density was assumed to decrease with altitude in an exponential law [*Maeda and Kimura,* 1956; 1959, *Yabroff,* 1961], and often in the function r^{-3}, which is based on the relation that f_p^2/f_H becomes constant, and is called the gyrofrequency model. This model was sometimes used in calculating dispersion of whistlers trapped in the field aligned ionization ducts. These functions were adopted from the reason that the calculated dispersions can be satisfactorily made to agree with those observed, by choosing appropriate parameters. From these results, a rough image of the electron density distribution in the magnetosphere was obtained.

The diffusive equilibrium (DE) model was investigated by Angerami and Thomas [1964], in which they took account of thermal diffusion of ions along a field line and gravity toward the earth's center. Motion of protons, and helium and oxygen ions were included. By their model, electron density distribution along geomagnetic field lines could be theoretically determined, with some parameters such as ion temperatures, and relative ion concentrations at a reference altitude, say 1000 km. We have information on these parameters, so that Kimura [1966] could use this model for ray tracing in the plasmasphere.

The existence of the plasmapause, the outer boundary of the plasmasphere, was disclosed by Carpenter [1963], from a series of nose whistlers, called 'knee whistlers' observed at a polar station, Eights. The process of his discovery of the plasmapause was based on the field aligned ducted propagation, which was thought to be a reasonable assumption for whistlers observed on the ground. Therefore ray tracing was not necessary. However, once the existence of the plasmapause was known, we had to take account of the effect of the plasmapause on the ray paths for whistler mode waves propagating in the magnetosphere.

Aikyo and Ondoh [1971] devised a model with the plasmapause for their ray tracing, in which they assumed that inside the plasmapause the diffusive equilibrium model [*Angerami and Thomas,* 1964] is used and in the outside they adopted a collisionless model [*Eviator et al.,* 1964].

Walter and Scarabucci [1974] adopted the DE model throughout the magnetosphere, but they made the electron density at the reference level latitude-dependent, the density being reduced at latitudes above L value corresponding to the plasmapause.

4. VARIOUS WHISTLER MODE PHENOMENA ELUCIDATED BY RAY TRACING

So far several whistler mode phenomena disclosed by ray tracing were introduced. In the following some other interesting results will be described.

LHR noise [*Brice and Smith,* 1965] is excited by whistlers and detected by satellites in near-earth orbits. This phenomenon was interpreted to result from the trapping of whistler energy in the valley of LHR frequency altitude profile [*Smith, et al.,* 1966] with the ray path appearing as a series of a figure 8 as shown in Figure 4.

The Nu whistlers, whose dynamic spectra appear as a reverse of the character ν, were detected by the OGO-1 satellite, together with the MR whistlers. Smith and Angerami [1968] concluded that the Nu whistlers belong to a group of the MR whistlers, and are composed of, for example, a pair of 2_-^{MR} and 2_+^{MR}, which are referred to Figure 3(a).

When there is an appropriate amount of horizontal gradient of electron density at the reference altitude level, the decreasing rate of the density toward high latitudes, that is, the negative density gradient in the radial direction increases so that a favorable condition is produced for the wave normal direction of whistlers to be kept within a small angle with a geomagnetic field line throughout the ray path to the opposite hemisphere.

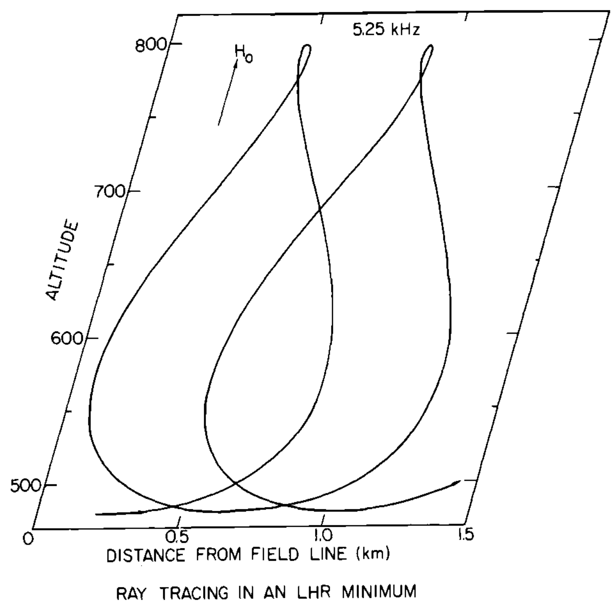


Figure 4. Trapping of rays in the LHR duct, calculated by ray tracing [*Smith et al.,* 1966]

This is called pro-longitudinal propagation (PL-mode) [*Scarabucci,* 1969].

On the other hand, normally, the wave normal direction is bent outward when a whistler mode wave passes over to the other hemisphere, finally approaching to the resonance cone angle. This type of propagation is sometimes called 'pro-resonance' (PR) propagation. If the local LHR frequency becomes larger than the wave frequency, an LHR reflection takes place. For waves with frequencies always higher than the local LHR frequency, the waves can reach lower ionosphere altitudes in the PR propagation mode. Walter and Angerami [1969] detected, by the OGO-2 and 4 satellites at altitudes below 1000 km, a group of whistlers in the PR propagation as well as the PL propagation at the same time, both arising from the same lightning sources. The PR whistlers had rising traces in frequency higher than the maximum LHR frequency along the ray path down to the satellite, and showed a rapid increase in travel time as a function of satellite latitude as compared with the PL components of whistlers observed at the same time. The PR components of whistlers were called 'walking trace (WT)' whistlers.

Taking account of a latitudinal or horizontal gradient in electron density, such as due to the equatorial anomaly, the ray tracing resulted in a focusing and defocusing of the ray paths for waves traversing over the magnetic equator. These characteristics were used to interpret an abrupt amplitude cutoff of signals from VLF transmitters [*Scarabucci,* 1970] and a drifting whistler cutoff phenomenon – striations [*Kimura,* 1971], both observed near the magnetic equatorial region by the OGO-4 satellite.

Various ray tracing trials were made, such as that by Walter and Scarabucci [1974], who found a condition of electron density distribution for getting whistler mode ray paths completely symmetric with respect to the geomagnetic equatorial plane. This condition will explain a possibility of multiple echoes or whistler trains without the help of field aligned ionization ducts.

5. RECENT DEVELOPMENTS OF RAY TRACING TECHNIQUE AND THEIR APPLICATIONS

5.1. *Geomagnetic field model for ray tracing*

For long, the ray tracing for whistler mode waves in the magnetosphere has been performed in the dipole magnetic field. However, it has been known that the actual geomagnetic field is deviated from the dipole model, and so-called IGRF (International Geomagnetic Reference Field) is recommended to be used for more realistic geomagnetic field except in the deep magnetosphere and the geomagnetic tail region.

One big problem introducing the IGRF model in ray tracing was that the electron density model, like the DE model, is defined along a geomagnetic field line. Namely, in order to know the plasma density at a certain point on the way of ray tracing, we have to refer to the plasma density at the reference altitude on the field line passing through the point, where the density is given by a latitude or L dependent function. In case of the dipole model, all field lines can be defined analytically from the location of the point in space, but in the IGRF model, the geomagnetic field line to the reference altitude from a point in space must be independently calculated by field line tracing. This process is a very heavy burden in doing ray tracing.

Kimura et al. [1985], devised the ray tracing program adaptable to the IGRF model, in which they took the following procedures, namely before the ray tracing, (1) using the IGRF model a magnetic field line tracing is performed to the reference altitude from a sufficient number of grid points evenly distributed in the space where ray tracing is later made, (2) at each grid point all informations about the field line passing through the point are registered, and (3) in the process of ray tracing the necessary quantities relating to the magnetic field line passing through each point on the ray path are calculated by interpolation from those at the nearest eight grid points. The procedure of interpolations takes much smaller computer time compared with that of integration for field line tracing, and it was checked that such an approximated procedure can be adopted for obtaining enough accurate ray paths by choosing a good number of grids.

As a result of ray tracing under the IGRF model, ray paths are generally deviated from the initial dipole meridian plane and the effect of non-dipolar magnetic field is found to be substantial. In the ray tracing used in the subsection 5.4, the above mentioned IGRF model is adopted in the computer program.

5.2. *Ray tracing of ion dominant modes*

Most dominant effect of ions from the point of view of ray tracing is 'LHR reflection', as mentioned earlier. For frequencies much lower than the LHR frequency, the wave propagation is directly associated with the ion cyclotron frequencies. The proton whistlers were first detected by Gurnett et al. [1965] by a satellite, as was interpreted from the modes in the multi-component plasma [*Smith and Brice,* 1964]. From the point of view of ray tracing, the ion cyclotron whistlers take place in the lower ionospheric region and are mostly observed by low earth orbiting satellite, so that the ray tracing is not always necessary.

For the ELF waves propagating in the magnetosphere around the ion cyclotron frequency, the ray tracing technique can play a powerful role. However, the algorithm of ray tracing is the same as that for whistler mode waves, except that the left-handed mode is taken for ion cyclotron waves and the right-hand polarized mode is taken for the magnetosonic mode or for the branch connected with the whistler mode.

Two types of interesting ELF emissions were often detected in the vicinity of the geomagnetic equatorial region, by a semi-polar orbiting satellite, Akebono, which was launched in 1989. The first type of emissions has frequency spectra that are limited below the proton cyclotron frequency and above the second LHR frequency for three component ions (H^+, He^+, and O^+) (to be called 'proton mode'), and below the helium ion cyclotron frequency and above the third LHR frequency (to be called 'helium mode'). Their wave normal directions observed by Akebono always lie in the geomagnetic meridian plane, with nearly a resonance cone angle. These waves are interpreted as the ion cyclotron mode. Ray tracing for this mode was started from the geomagnetic equatorial plane with the initial wave nor-

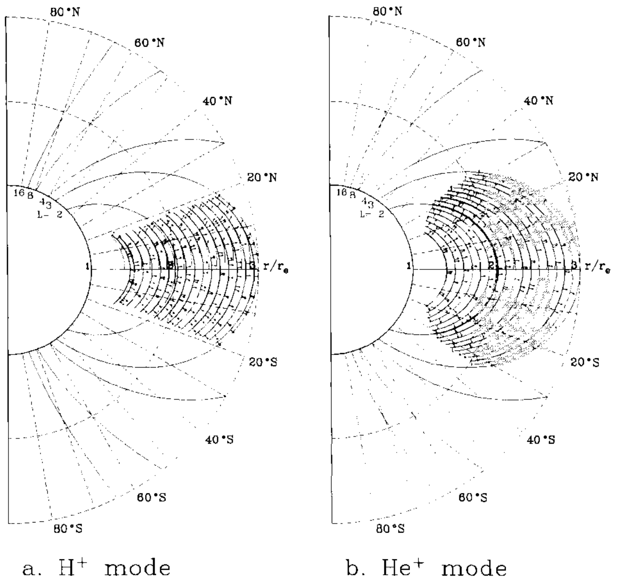


Figure 5. The results of ray tracing for (a)proton mode and (b)helium mode [*Kasahara et al.*, 1992].

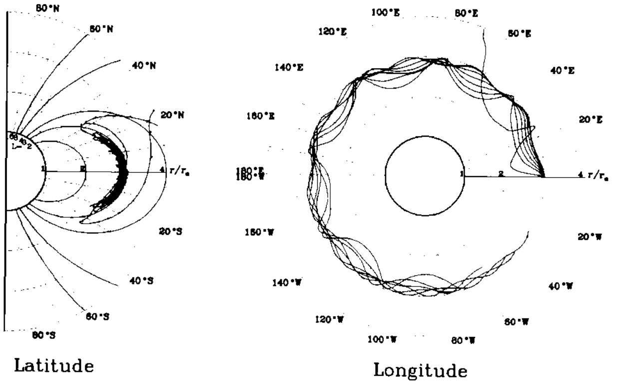


Figure 6. Calculated ray paths of magnetosonic waves which traverse around the earth in the vicinity of the equatorial plane, where a sharp density gradient is located around L=3 [*Kasahara et al.*, 1994].

mal directed along the geomagnetic field line. It was found that the rays traverse almost along the field lines and are reflected back due to LHR reflection, so that the wave energy is trapped within a relatively low geomagnetic latitude region in the geomagnetic meridian plane with the wave normal direction nearly the resonance cone angle, as illustrated in Figure 5(a) and (b) for the proton mode and the helium mode respectively. These results completely agree with the characteristics of the detected region and observed wave normal directions for the above mentioned emissions [*Kasahara et al.,* 1992].

The second type of emissions also detected by the Akebono satellite in the equatorial region was those with a wide frequency band and some times a harmonics structure covering below and above the proton cyclotron frequency. For these emissions, the wave normals were directed nearly perpendicular to the meridian plane. Ray tracing was made for the right-hand polarized magnetosonic mode (connected with the whistler mode) with the wave normal direction nearly perpendicular to the meridian plane, starting from the geomagnetic equatorial plane [*Kasahara et al.,* 1994]. Rauch and Roux [1982] had made a similar ray tracing for the magnetosonic mode in a plasma density distribution monotonically decreasing with radial distance and found that all rays go outward. Kasahara et al.[1994], on the other hand took account of the existence of a sharp density gradient around a distance where the emissions were detected, such as the plasmapause, and found that the ray paths are trapped within a torus with the sharp gradient as the outer boundary, circulating around the equatorial plane, as shown in Figure 6.

5.3. *Effects of hot plasma in ray tracing*

In a warm or a hot plasma, whistler mode waves suffer cyclotron damping. Another effect of thermal motion of particles is a modification of plasma dispersion from the cold plasma case. Growth and damping through the ray paths have been calculated, for e.g. plasmaspheric hiss. However, most of them were separately made from ray tracing, that is ray paths were calculated in the cold plasma, and growth or damping was calculated by integrating along the paths the imaginary part of refractive index due to high energy particles.

Real hot plasma ray tracing can be made by modifying the dispersion relation taking account of isotropic Maxwellian hot plasma, the ray paths could be traced of so-called Z mode continually through the electron cyclotron harmonic mode [*Hashimoto et al.*, 1987].

A hot plasma ray tracing for the magnetospherically reflected whistlers was also made by Thorne and Horne [1994].

5.4. *Tomographic approach of determining the global electron density distribution in the plasmasphere by ray tracing technique*

For most of ray tracing, we wish to find the plasma density distribution in the plasmasphere or magnetosphere model, so as to well explain observed ELF and VLF wave phenomena. This is however an inverse problem, that is we have to know a plasma density distribution in advance for doing ray tracing. The destiny was found to be reconciled in the following work, where we have positively used the ray tracing technique to find appropriate plasma parameters to construct a global electron density distribution in the plasmasphere, using wave data observed by the Akebono satellite.

Namely, Omega navigational signals transmitted from eight stations over the world could have been detected continuously over more than 30 minutes, and their wave normal directions are determined by three orthogonal loop antennas and a pair of crossed dipole antennas onboard Akebono. The delay time from the source to Akebono and *in-situ* electron density can also be measured on-board. Using these three quantities obtained from the wave measurements an appropriate set of plasma parameters based on the DE model with an ion temperature gradient along geomagnetic field lines can be determined, by changing the plasma parameters until these quantities calculated by ray tracing and field line tracing agree with those observed along the satellite trajectory. The algorithm of calculation to find the parameters is non-linear least square fitting method, so-called modified Marquart method [*Sawada et al.*, 1993; *Kimura et al.*, 1995; 1996a].

In Figure 7, four electron density profiles on the equatorial radial line, thus obtained on four different days in December, 1989 are illustrated. The profiles on 891206 and 891211 show a compression of electron density after a very high magnetic activity ($K_p = 6$ on 891206) and the profiles on 891218 and 891222 show those of quiet conditions long after the high activity on 891206. As shown in this figure, a global electron density profile in the plasmasphere around the meridian of Akebono orbit can be determined for each orbit of the Akebono trajectories, as far as at least one Omega signal was detected continuously for more than half an hour [*Kimura et al.*, 1996b].

In these procedures to obtain a global electron density profile, we use a number of Omega ray paths from the source to the satellite, so that this is a kind of tomographic approach and the ray tracing technique using the IGRF model is effectively utilized.

6. CONCLUDING REMARKS

This paper was prepared as a historical review as well as a review of the recent development of ray tracing techniques used in the plasmasphere and magnetosphere of the earth.

Needless to say, plasma wave phenomena are closely associated with the physics of the plasmasphere and

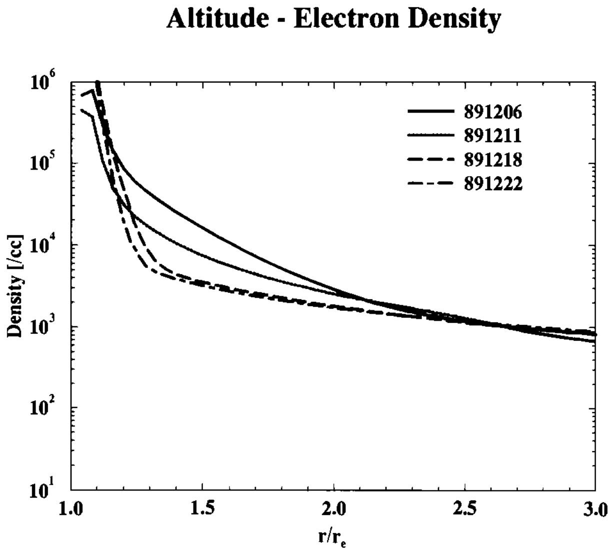


Figure 7. Several global electron density profiles (the densities along the equatorial radial line are shown) on different days in December, 1989 [*Kimura et al.*, 1997].

Kimura, I., Drifting whistler cut-off phenomena –Stria- tions– observed by POGO satellites, *Space Res.*, *XI*, 1331–1336, 1971.

Kimura, I., Whistler mode propagation in the earth and planetary magnetospheres and ray tracing technique, *Space Science Review*, *42*, 449–466, 1985.

Kimura, I., Ray paths of electromagnetic and electrostatic waves in the earth and planetary magnetospheres, *AGU Monograph 53, Plasma Waves and Instabilities at Comets and in Magnetospheres*, 161–171, 1989.

Kimura, I., T. Matsuo, M. Tsuda, and K. Yamauchi, 3-D ray tracing of whistler mode waves in a non-dipolar magnetosphere, *J. Geomag. Geoelectr.*, *37*, 165–180, 1985.

Kimura, I., A. Hikuma, Y. Kasahara, A. Sawada, M. Kikuchi, and H. Oya, Determination of electron density distributions in the plasmasphere by using wave data observed by Akebono satellite, *Adv. Space Res.*, *15(2)*, (2)103–(2)107, 1994.

Kimura, I., A. Hikuma, Y. Kasahara, and H. Oya, Electron density distribution in the plasmasphere in conjunction with IRI model, deduced from Akebono wave data, *Adv. Space. Res.*, *18(6)*, (6)279–(6)288, 1996.

Kimura, I., K. Tsunehara, A. Hikuma, Y. Z. Su, Y. Kasahara, and H. Oya, Global electron density distribution in the plasmasphere deduced from Akebono wave data and the IRI model, *J. Atmosph. Terr. Phys.*, , in press, 1997.

Lyons, L. R., and R. M. Thorne, The magnetospheric reflection of whistlers, *Planet. Space Sci.*, *18*, 1753–1767, 1970.

Maeda, K., and I. Kimura, A Theoretical Investigation on the Propagation Path of the Whistling Atmospherics, *Rep. Ionos. Space Res. Japan*, *X(3)*, 105–123, 1956.

Maeda, K., and I. Kimura, Calculation of the propagation path of whistling atmospherics, *J. Atmos. Terr. Phys.*, *15*, 58–65, 1959.

Rauch, J. L., and A. Roux, Ray tracing of ULF waves in a multicomponent magnetospheric plasma: Consequences for the generation mechanism of ion cyclotron waves, *J. Geophys. Res.*, *87*, 8191–8198, 1982.

Russell, C. T., R. E. Holzer, and E. J. Smith, OGO-3 observations of ELF noise in the magnetosphere, 1, Spatial extent and frequency of occurrence, *J. Geophysical Res.*, *74*, 755–777, 1969.

Sawada, A., T. Nobata, Y. Kishi, and I. Kimura, Electron density profile in the magnetosphere deduced from the in-situ electron density and wave normal directions of Omega signals observed by the Akebono (EXOS-D) satellite, *J. Geophysical Res.*, *98(A7)*, 11267–11274, 1993.

Scarabucci, R. R., Satellite observations of equatorial phenomena and defocusing of VLF electromagnetic waves, *J. Geophys. Res.*, *75*, 69–84, 1970.

Scarabucci, R. R., and R. L. Smith, Study of magnetospheric field oriented irregularities in the mode theory of bell-shaped ducts, *Radio Sci.*, *6(1)*, 65–86, 1971.

Smith, R. L, Guiding of whistlers in a homogeneous medium, *J. Res. NBS-D, Radio Propagation*, *64D(5)*, 505–508, 1960.

Smith, R. L., Propagation characteristics of whistlers trapped in field-aligned columns of enhanced ionization, *J. Geophys. Res.*, *66(11)*, 3699–3707, 1961.

Smith, R. L., Explanation of subprotonospheric whistlers, *J. Geophys. Res.*, *69*, 5019–5021, 1964.

Smith, R. L., and N. M. Brice, Propagation in multicomponent plasmas, *J. Geophys. Res.*, *69*, 5029–5040, 1964.

Smith, R. L., I. Kimura, J. Vigneron, and J. Katsufrakis, Lower hybrid resonance noise and a new ionospheric duct, *J. Geophys. Res.*, *71(7)*, 1925–1927, 1966.

Smith, R. L., and Angerami, J. J., Magnetospheric properties deduced from OGO 1 observations of ducted and nonducted whistlers, *J. Geophys. Res.*, *73*, 1–20, 1968.

Thorne, R. M., and R. B. Horne, Landau damping of magnetospherically reflected whistlers, *J. Geophys. Res.*, *99(9)*, 17,249-17,258, 1994.

Walter, F., and Angerami, J. J., Nonducted mode of VLF propagation between conjugate hemispheres; Observations on OGO's 2 and 4 of the 'walking- trace' whistler and of Doppler shifts in fixed frequency transmissions, *J. Geophys. Res.*, *74*, 6352–6370, 1969.

Walter, F., and R. R. Scarabucci, VLF ray trajectories in a latitude-dependent model of magnetosphere, *Radio Science*, *9 (No.1)*, 7-15, 1974.

Yabroff, I. W., Computation of whistler ray paths, *J. Res. NBS.*, *65D*, 485–505, 1961.

Iwane Kimura, Faculty of Information Science, Osaka Institute of Technology, 1-79-1 Kitayama, Hirakata-city, Osaka 573-01, Japan

Music and the Magnetosphere

Carl E. McIlwain

Department of Physics, and Center for Astrophysics and Space Sciences, University of California at San Diego

The beginning of the space age is usually associated with the time of the first satellite launches. Some of us were fortunate enough to have already been pursuing the exploration of space. This paper attempts to document some of the extraordinary opportunities for exploration and discovery Professor James A. Van Allen offered a young music student at the State University of Iowa (SUI) during the period 1954 to 1959.

I arrived in Iowa with the intention of studying the physics of music in early 1954. I had a degree in music, and at the time, the only thing required to be a graduate student in anything at the State University of Iowa was to have a bachelor's degree. Well, I had a bachelor's degree, so I signed up as a physics graduate student, and enrolled in undergraduate physics, mathematics, and chemistry courses. For three years at the North Texas State College School of Music, I had helped my flute teacher and mentor, Dr. George Morey, by teaching the secondary flutists. Upon my graduation, he arranged a student teaching position for me at SUI, his Alma Mater. Shortly after arrival, I played my flute for Himie Voxman, the Chairman of the SUI Music Department, and obtained a chair in the SUI orchestra (as second flutist, first chair positions being reserved for music majors). Later Chairman Voxman informed me that the flute teaching position I was supposed to take over had disappeared. The previous graduate student had decided to join the Iowa faculty, and would be continuing her role of teaching the flute students. Professor Betty Bang later became the President of the National Flute Association, and continues her illustrious career at the University of Iowa to this day.

At that time, Frank McDonald had recently arrived from Minnesota with a position as Research Associate. There were other Van Allen graduate students: Les Meredith was just finishing; Joe Kasper, Bill Webber, and Ernie Ray were well underway; Larry Cahill would arrive shortly; and George Ludwig was about to finish his undergraduate studies. Obviously, I was just beginning.

Dr. Van Allen was on sabbatical leave at Project Matterhorn at the time I arrived. Upon return, he found that he had some new graduate students, so he called them into his office, and put a list of potential projects and topics needing attention on the blackboard. Van Allen had developed a technique of getting extra altitude out of rockets by launching them from balloons, and he called the combination rockoons. He had the idea of launching rockoons, not only with just six-inch rockets, but also with easier-to-use and cheaper rockets that were only three-inches in diameter. One of the potential projects was to miniaturize the proton precession-magnetometer, which had just been invented. Larry Cahill thought that was interesting, so he took over that project of fitting the magnetometer into the three-inch rockets.

Another project was to use these small rockets to make a latitude survey of cosmic rays, i.e., to obtain the spectrum of cosmic rays by measuring the magnetic latitude dependence and using the Earth's magnetic field as a spectrometer. Not knowing any better, the flute player said, "OK, I'll do that." (Dr. Van Allen probably did not know that this student was fresh out of music school.) The idea was to launch the three-inch Loki rockets from Skyhook balloons from a Navy ship during a voyage to Thule, Greenland in 1955. Even in those days, money was needed to do things. Therefore, Dr. Van Allen wrote a proposal to NSF. Figure 1 shows a copy of that proposal. Only one and a half pages,

Discovery of the Magnetosphere
History of Geophysics Volume 7

asking for only $2000. The estimated costs were later slightly reduced to make room for another item: fifteen percent university overhead. Dr. William H. Pickering, Director of JPL, helped obtain surplus Loki Phase I rockets, and that reduced the launch costs.

Carl McIlwain

(General background on our obligation to NSF)

17 February 1955

Mr. Ross C. Peavey
National Science Foundation
Washington 25, D. C.

Dear Mr. Peavey:

At the 22 January 1955 meeting of the Technical Panel on Rocketry (of the U. S. National Committee for the International Geophysical Year) I drew attention to the fact that none of the budgeted amount of $120,000 for fiscal year 1955 had yet been made available for the procurement of rockoons for a vital phase of the I.G.Y. rocket program. I further pointed out that there were a number of new possibilities available for smaller and less expensive rockets for this program and that, at the least, a thoroughgoing study of these possibilities should be made as soon as possible.

It was proposed by Dr. Whipple and agreed by the Panel that an amount of $2,000 should be allocated to the State University of Iowa for this purpose, the work to be done under my supervision. The proposed use of these funds is as follows:

(a) Trajectory and design calculations and preparations of reports	$800
(b) Travel	400
(c) Procurement of several small rockets and balloons for tests of new small rockets, and preparation of test instrumentation	800
Total	$2,000

It is proposed that the results of design and performance studies be presented to the Technical Panel on Rocketry at or before its next scheduled meeting on 28 April 1955. It is further proposed that the flight appraisals be made in July 1955.

Sincerely yours,

J A Van Allen

J. A. Van Allen

cc. Business Office, State University of Iowa,
Members Technical Panel on Rocketry
and Special Committee on Small Rockets for the I.G.Y.

Figure 1. Proposal for the first Loki rockoon program.

After designing an adapter for the three-quarter inch tip of the Loki rocket permitting a three-inch-diameter instrumentation, the next task was to try to package a Geiger tube, voltage sources, and telemetry system in a lightweight stack capable of withstanding the 270 g acceleration of the Loki rocket. The individual components were tested to 1000 g in a centrifuge and to temperatures below minus fifty degrees Centigrade (it is very cold at balloon altitudes).

Hoping to detect the 'soft radiation' discovered on a previous rockoon expedition [*Meredith et al,* 1955; *Van Allen,* 1957], the nose cones were made of aluminum only 32 thousandths of an inch thick. Transistors were unknown, so the signal circuits, the transmitters, and high-voltage supplies all required vacuum tubes. The transmitter coils were wound by hand and then adjusted for proper frequency (74 Mhz) and maximum power (one to two watts). The adapter included electrical insulation so the rocket motor served as one half of a dipole antenna. The broad lawn in front of the Old Capitol was used to simulate "space" for testing the transmitters.

The instrumentation assembly line for my ten little three-inch rockets occupied one of the lab benches in the basement of the Physics Building (Figure 2). The entire instrumentation assembly weighed 6.8 pounds.

Frank McDonald, Joe Kasper, and George Ludwig were also on the 1955 expedition. Joe Kasper served many roles, and also volunteered to help me with the contents of the classes being missed while on the ship. George Ludwig was helping Frank McDonald with instrumentation for the large six-inch Deacon rockets. George's presence on the expedition was very helpful for everyone. We four proceeded to go to the U.S.S. Ashland (LSD-1) at Norfolk, Virginia, and found, to our dismay, that some of the three-inch rockets had arrived without their high-altitude fins, and only had their one-inch stubby fins for low-altitude launches. Frank McDonald and I ran into a Norfolk hardware store for some sheet aluminum, came back, cut some fins, and bolted them on. We took the precaution to use a hand drill, rather than an electric drill, to make the screw holes in order to avoid any possibility of accidental ignition.

Thus in August 1955, I found myself on a ship west of Greenland launching balloons carrying rockets. There were many experiences during this expedition, including coping with the ship's rolling up to 45 degrees in the large waves left behind by hurricane Ione. Some of the launches were made during periods with substantial rolling (Figure 4), but it was calmer for the Loki rockoon launches (Figure 5). Figure 3 is a page from my notebook at the time of the first Loki launch. It is easy to imagine my elation when Ludwig reported that he had seen a rocket trail. So it *did* ignite ... that had been the first question: would it ignite? (with good cause, the first rockoon in 1952 did not ignite). But then there was great consternation: we could not find the signal. What happened? After an agonizing 20 seconds, I found the signal at a different frequency, and began recording the data from my first rocket. As an ex-music student, I was obviously elated to be receiving real cosmic ray counts.

Figure 13. A half minute time exposure of Nike-Cajun II6.27F from nine miles away.

transition from normal to zero rates. Did it just drop suddenly as in a failure, or did it smoothly rise to higher rates, go into saturation, and finally give only zeros? Unfortunately, only tiny fragments of data were available as only the scattered Minitrack Stations were being used at the time. We had to wait, not only until Explorer II--- it went into the ocean--- but until Explorer III was launched with the tape recorder that George Ludwig had developed. Two film strip copies of its first readout, recorded in San Diego on March 28, 1958, were sent. One was sent to Van Allen, who had gone to Washington after the launch. The other was sent to Iowa, where Assistant Professor Ernie Ray, Joe Kasper, and I promptly grabbed the reel of film, put it on a microfilm reader, and anxiously began looking for a transition. *And there it was.* So we knew at once that there was something of very high intensity out there. I immediately took the spare payload, and put it in front of an x-ray machine (a 250 kV DC machine I had installed for calibrating my rocket instrumentation) where I generated what became known as a Van Allen r vs R plot [Figure 8 in *Van Allen*, 1958]. The results showed that fluxes that would ideally produce more than 35,000 counts per second, instead drove the rate to zero. We knew we had measurements of an exciting new phenomenon.

Simultaneously, having no tools other than his slide rule in his Washington hotel room, Van Allen [1983] bought graph paper and a ruler at a local drug store, and carefully plotted the counting rates for the entire 102 minutes of data. At 3:00 AM, he had turned in for the night "with the conviction that our instruments on both Explorers I and III were working reliably and giving reproducible results but that we were encountering a mysterious physical effect of a real nature" [*Van Allen,* 1983 p66]. Returning to Iowa, Dr. Van Allen proudly showed Ernie Ray and me his graph. I then showed him my x-ray machine results. He instantly agreed that the satellites were encountering very high fluxes.

Van Allen announced the discovery to the world at a National Academy Meeting in Washington on May 1, 1958 [*Van Allen* et al, 1958; *Berland,* 1962; *Hanle and Chamberlain* 1981 p58]. It was clear that a spacecraft was needed to go up and study this new phenomenon. The preceding fall, Nicholas Christofilos had asked "What would happen if we set a high-altitude atomic bomb off; would it

with the bomb blasts. Explorer IV proceeded to measure a great deal of what was to be known about the radiation belts for some time. Perhaps the majority of what was known for the first two years about the radiation belts was from Explorer IV [*Van Allen et al,* 1959a]. It did most of the mapping, much of the composition estimates, and was up in time to measure the results of the high-altitude nuclear explosions.

In February 1959, there was a classified workshop at the Lawrence Livermore Laboratory. Earlier, Edward Teller had asked Ted Northrop to "see if you can find out how particles drift in longitude." Nobody knew. We knew that particles spiraling around the magnetic field lines would bounce and be trapped, but did not know how they would drift around the Earth. Ted Northrop found the key: the Rosenbluth longitudinal invariant. At the workshop, he gave an impromptu seminar on the invariant to Dr. Van Allen, me, and other interested people. Later, the invariant was described in the open literature (Northop and Teller, 1960).

This invariant formed the basis for devising a way of mapping trapped radiation, the B,L coordinate system [*McIlwain,* 1961]. There was the fortunate circumstance of Ted Northrop finding exactly what was needed, even though it was then known only in a few classified circles. The nuclear explosions gave markers on magnetic shells, which told us where the particles were drifting [*Van Allen et al,* 1959b], and provided one of the first confirmations that adiabatic invariants really worked. Explorer IV thus provided a firm observational basis for the B,L coordinate system [*Van Allen,* 1962].

In conclusion, it is now recognized that radiation belts are an important and common aspect in many parts of our universe. We at the State University of Iowa who were involved with the Explorer I, III, and IV spacecraft were exceedingly lucky be there to help produce mankind's first view of this wondrous new phenomenon. The ex-musician would have liked more time to perform music, but he has never had regrets concerning his Iowa transformation.

Acknowledgments. Dr. James A. Van Allen cannot be thanked too much for providing the many opportunities for research and personal development, and for his continued support over the years. Mary, my wife of 44 years, also deserves great credit for her assistance and encouragement. I thank Richard Maheu and Stephen Kerr for their help in preparing this paper. The National Science Foundation and the Office of Naval Research provided partial financial support for the research projects and we had operational support from the U. S. Navy and the U. S. Army. The operational support included the use of facilities, transportation, food, clothing, and photographers.

REFERENCES

Berland, T., *The Scientific Life*, Coward-McCann, Inc., New York, 1962.

Davis, L. R., D. E. Berg, and L. H. Meredith, Direct measurements of particle fluxes in and near auroras, *Proc. Cospar Space Sci. Symposium,* North Holland Pub. Co., Amsterdam, 721-35, 1960.

Hanle, P. A., and V. D. Chamberlain, *Space Science Comes of Age,* National Air and Space Museum , Smithsonian Institution Press, Washington, D. C., 194pp, 1981

McDonald, F. B., IMPs, EGOs, and Skyhooks, *J. Geophys. Res., 101*: No. A5, 10,521-30, 1996.

McIlwain, C. E., Cosmic ray intensity above the atmosphere at northern latitudes, *M.S. thesis, University of Iowa*, 55pp, 1956.

McIlwain, C. E., Scintillation counters in rockets and satellites, *Institute of Radio Engineers Transactions in Nuclear Science, N57*, 7:159, 1960a.

McIlwain, C. E., Direct measurement of protons and electrons in visible aurorae, *Proc. Cospar Space Sci. Symposium,* North Holland Pub. Co., Amsterdam, 715-20, 1960b.

McIlwain, C. E., Direct measurement of particles producing visible aurorae, *J. Geophys. Res. 65*, 2727-47, 1960c.

McIlwain, C. E., Coordinates for mapping the distribution of magnetically trapped particles, *J. Geophys. Res., 66*, 3681, 1961.

Meredith, L. H., M. B. Gottlieb, and J. A. Van Allen, Direct Detection of soft radiation above 50 kilometers in the auroral zone, *Phys. Rev., 97*, 201-5, 1955.

Meyer, P., E. N. Parker, and J. A. Simpson, Solar Cosmic Rays of February, 1956 and their propagation through interplanetary space, *Phys. Rev., 104, No. 3*, 768-783, 1956.

Northop, T. G. and E. Teller, Stability of the adiabatic motion of charged particles in the Earth's field, Phys. Rev., *117*, 215, 1960.

Sullivan, Walter; *Assault on the Unknown*, McGraw Hill, 460pp, 1961.

Van Allen, J. A., and C. E. McIlwain, Cosmic ray intensity of high altitudes on February 23, 1956, *J. Geophys. Res. 61*, 569-570, 1956.

Van Allen, J. A., Direct detection of auroral radiation with rocket equipment, *Proc. Nat. Acad. Sciences, 43*, 57-62, 1957.

Van Allen, J. A., G. H. Ludwig, E. C. Ray, and C. E. McIlwain, Observations of high intensity radiation by satellite 1958 Alpha and Gamma, *Jet Propulsion,* 588-592, 1958.

Van Allen, J. A., C. E. McIlwain, and G. H. Ludwig, Radiation observations with satellite 1958 Epsilon, *J. Geophys. Res. 64*, 271-286, 1959a.

Van Allen , J. A., C. E. McIlwain, and G. H. Ludwig, Satellite observations of electrons artificially injected into the geomagnetic field, *National Academy Sciences Proceedings 45*, 1152-66, 1959b; and *J. Geophys. Res. 64*, 877-891, 1959b.

Van Allen, J. A., Dynamics, composition and origin of the geomagnetically-trapped corpuscular radiation, *Transactions of the International Astronomical Union Vol. XIB*, pages 99-136, 1962.

Van Allen, J. A., Origins of Magnetospheric Physics, *Smithsonian Institution Press*, 144pp Washington, D.C. 1983.

Van Allen, J. A., Early rocket observations of auroral bremsstrahlung and its absorption in the mesosphere, *J. Geophys. Res, 100,* 14485-97, 1995.

C. E. McIlwain, Center for Astrophysics and Space Sciences, University of California San Diego, La Jolla, California, 92093-0111 USA.
(e-mail: cmcilwain@ucsd.edu)

Adventures With the Geomagnetic Field

E. N. Parker

Enrico Fermi Institute and Departments of Physics and Astronomy University of Chicago, Chicago, Illinois 60637

The magnetic field of Earth and its outward extension into the varying solar wind provide a number of puzzles in classical physics, ranging from the generation of the field in the liquid metal core to the various forms of agitation driven by the solar wind. The effects are described by Newton's equations of motion for the atoms, ions, and electrons and Maxwell's equations for the electromagnetic fields. The problem is to identify the primary physical concepts and effects in the observed activity so that the overall complexity can be understood in terms of a few principles from Newton and Maxwell. In that way the geomagnetic field, along with the rest of the astronomical universe, provide an ongoing challenge and adventure for the interested physicist. Prior to the Space Age progress was slow but substantial over many decades, working with such indirect phenomena as the observed behavior of comet tails, records of worldwide geomagnetic fluctuations, records of the activity of the Sun, etc. The Space Age made it possible to obtain direct *in situ* information, greatly accelerating the scientific inquiry. So the pursuit of magnetospheric physics has been a continuing fascination and adventure throughout my professional career, with some of my earliest and latest research papers devoted to the topic and with several basic puzzles still unresolved.

I. THE EARLY YEARS

My adventures in the physics of the terrestrial magnetosphere began with the $\alpha\omega$ dynamo theory for the origin of the dipole geomagnetic field (Parker, 1955b). That work was based on the fundamental conjecture by Walter M. Elsasser (1945, 1946) that the geomagnetic field must be induced by the motions in the convecting electrically conducting metal core. That is to say, the geomagnetic field is not a consequence of a ferromagnetic interior, or a thermoelectric effect, or an intrinsic property of angular momentum. Elsasser proceeded to show that the magnetic field in the core is principally azimuthal, generated from the dipole component by the nonuniform rotation of the core. The question, then, was how to generate the axisymmetric dipole component from the dominant axisymmetric azimuthal magnetic field. It was my good fortune to work as a research associate with Professor Elsasser at the University of Utah for two years (1953-1955) thereby assimilating his ideas, on which my own work was based. After learning the basic concepts of magnetohydrodynamics from expositions by Lundquist (1952) and Elsasser (1954), I stumbled through the strictures on axisymmetric field generation imposed by Cowling's (1934) theorem and finally realized that the several convection cells in the core of the rapidly rotating Earth must be cyclonic and that the cyclonic convection provided the necessary net circulation of magnetic field in the meridional planes, equivalent to an azimuthal vector

Discovery of the Magnetosphere
History of Geophysics Volume 7

potential. A little smoothing through resistive diffusion and the mean axisymmetric dipole field was the result, thereby completing Elsasser's conjecture. The mean rate of generation of the dipole field (azimuthal vector potential) is given by the product of the azimuthal B_φ already established by Elsasser and a coefficient Γ characterizing the cyclonic velocity of the convective cells multiplied by their filling factor. If the azimuthal vector potential is denoted by A, then, for axial symmetry ($\partial/\partial\varphi = 0$), the mean field equation can be written

$$[\partial/\partial t - \eta(\nabla^2 - 1/\varpi^2)]A = \Gamma B_\varphi$$

in cylindrical polar coordinates (ϖ,φ,z) where η is the resistive diffusion coefficient and $\Gamma = \Gamma(\varpi,z)$, now conventionally denoted by $\alpha(\varpi,z)$. The generation of B_φ is, of course, given by the usual magnetohydrodynamic induction equation

$$[\partial/\partial t - \eta(\nabla^2 - 1/\varpi^2)]B_\varphi = \partial(\varpi A,\Omega)/\partial(\varpi,z),$$

where $\Omega = \Omega(\varpi,z)$ is the angular velocity in the convecting core and the right-hand side represents the Jacobian or Poisson bracket. In cartesian coordinates in which y represents the equivalent azimuthal direction of the mean field ($\partial/\partial y = 0$), the equations can be written

$$[\partial/\partial t - \eta(\partial^2/\partial x^2+\partial^2/\partial z^2)]A = \Gamma B_y, \qquad (1)$$

$$[\partial/\partial t - \eta(\partial^2/\partial x^2+\partial^2/\partial z^2)]B_y = \partial(A,v)/\partial(x,z), \qquad (2)$$

where the velocity $v = v(x,z)$ in the y-direction represents the nonuniform rotation. In the simplest case of uniform shear let $G_x = \partial v/\partial x$, $G_z = \partial v/\partial z$ with the result

$$[\partial/\partial t-\eta(\partial^2/\partial x^2+\partial^2/\partial z^2)]B_y = G_z\partial A/\partial x-G_x\partial A/\partial z. \qquad (3)$$

Then if the spatial dependence of Γ is ignored, the basic solutions of equations (1) and (3) have the exponential form $\exp(t/\tau + ik_x x + ik_z z)$, with the dispersion relation

$$1/\tau = \pm[\Gamma(k_xG_z-k_zG_x)/2]^{1/2}(1+i)-\eta(k_x^2+k_z^2).$$

It is obvious by inspection that there are growing oscillating solutions for real k_x and k_z while there are nonoscillatory growing solutions when k_x and k_z are complex, as when the region is bounded. The nonoscillatory solutions are appropriate for Earth (see the review in Elsasser, 1951, 1955, 1956a,b). The oscillatory solutions suggested that the periodic magnetic field of the Sun is generated by the same combination of nonuniform rotation and cyclonic convection that provides the quasi-stationary geomagnetic field. The oscillatory character of the magnetic field of the Sun is a consequence of the relatively thin geometry of the convective zone and the relatively slow rotation (of the order of one rotation per convective turnover) (Parker, 1955b, 1957c). The azimuthal field of the Sun is brought to the surface by its own intrinsic magnetic buoyancy (Parker, 1955a), where it forms the bipolar magnetic regions and the associated activity. It is that activity that produces the outbursts in the solar wind responsible for much of the magnetospheric activity at Earth.

□

II. THE VIEW FROM CHICAGO

In 1955 I became a research associate with Professor John A. Simpson at the University of Chicago. Simpson was observing and analyzing the energy dependence of the cosmic ray variations associated with solar activity. By the time I arrived Simpson (1954, 1955; Simpson et al, 1955; Fan et al, 1960a,b) had established that the energy spectrum of the cosmic ray variations could not be produced by changes of the geomagnetic cutoff energies, or by an electrostatic potential in space. Hence the variations could be attributed only to changing interplanetary magnetic fields. This implied that interplanetary space is filled with magnetic fields and the plasmas to manipulate them. Years earlier Simpson had invented the cosmic ray neutron monitor to study the variation of the lower energy (1-15 Gev) cosmic ray particles supplementing the traditional ion chambers and Geiger counters, that respond principally to the mu meson component produced in the atmosphere by cosmic rays of 15 Gev and above. He distributed five neutron monitors over geomagnetic latitude from the equator (Huancayo, Peru) to Chicago, so that he could determine the energy spectrum of the cosmic ray variations using the geomagnetic field as a mass spectrometer. The effective geomagnetic cutoff at the equator is approximately 15 Gev, falling to about 1.5 Gev at the geomagnetic latitude of Chicago. Thus the geomagnetic dipole field was part of the overall instrumental system, which included the interplanetary "solar corpuscular radiation" and whatever variable interplanetary magnetic fields might be there. It paved the way for recognizing the concept of the solar wind. It must be understood how little was known of conditions above the terrestrial ionosphere at that time. No radio beam could probe beyond the F-layer of the ionosphere, and instruments could be carried no higher than high flying balloons or the V-2 rocket[1]. So knowledge of magnetospheric conditions was limited to inference from the time variation of ground based magnetometers, ground and balloon borne cosmic ray detectors, ionospheric (i.e. radio propagation) conditions, auroral activity, etc. It is interesting and even amusing, now some forty years later, to look back at the random assortment of ideas that many of us constructed to account for the activity of the magnetosphere, the aurora, the cosmic ray variations, etc. For instance, for a brief period I advocated the absurd idea that the Forbush cosmic ray decrease was caused by magnetic interplanetary gas clouds settling down onto the outer geomagnetic field. Can anyone top that for a dubious theory? Generally, each of us started from some glimpse of the truth which was then extrapolated to provide a more complete, if retrospectively naive, scenario. It is a good thing that there were many of us, each pursuing his or her own concepts and misconcepts, so that the random components of the ideas largely canceled out in time.

One of the interesting exceptions to the severe observational limitations was provided by the great solar flare of 23 February, 1956, which produced an enormous outburst of relativistic protons and helium nuclei. The prompt arrival of these energetic particles at Earth came from the direction of the Sun, with isotropic arrival developing within about five minutes. The net effect was a direct probe of interplanetary magnetic conditions. The worldwide observations of this enormous cosmic ray event collectively showed clear passage from the Sun to Earth and strong scattering and impediment beginning at about the orbit of Mars (Meyer et al, 1956). In rereading our old paper I am struck by the mixture and confusion of the old and the new concepts. On the one hand, we thought in terms of the galactic magnetic field penetrating through interplanetary space, with the prompt arrival of the flare particles placing an upper limit at 10^{-7} gauss on any transverse component of that field, otherwise estimated at 10^{-5} gauss (Chandrasekhar and Fermi, 1955). On the other hand, we noted Davis's point (Davis, 1955) that the general solar corpuscular radiation pointed out by Biermann (1951, 1952) must sweep back any galactic field to some great distance ~200 a.u. to where there is a balance of pressure with the outgoing corpuscular radiation. But we hardly knew what to make of these mutually exclusive ideas. Then, we quoted the 500 electrons/cm^3 in the interplanetary space at the orbit of Earth based on the inference by Behr and Siedentopf (1953) that the zodiacal light is primarily sunlight scattered by free electrons rather than dust. We also alluded to Chapman's (1957) concept of a solar corona extending beyond the orbit of Earth, thinking that the 500 electrons/cm^3 represents Chapman's static corona. We wrote the curious paragraph "There is strong evidence that this ionized gas, in the process of escaping from the sun, will occasionally transport away from the sun small amounts of the general solar magnetic field (mapped by Babcock and Babcock, 1955). Additional gas in the form of clouds and so-called solar streams may escape field-free". But we went no further. In fact, all the essential facts to realize the solar wind were there, but I did not understand the dynamics of tenuous plasmas well enough at the time to make the necessary connections. That had to wait a couple of years while I explored the dynamical theory of the large-scale motion of the collisionless plasma. Only then was I able to push forward the dynamical theory of the magnetosphere and only then did the facts of Chapman's and Biermann's work fall into place in my mind to see the implications for coronal expansion, the solar wind, and the associated magnetospheric activity. Nonetheless our stumbling analysis of 23 February event emphasized in our minds the central importance of the dynamical state of interplanetary plasmas and fields, whatever that might be.

III. THE ORIGINS OF MAGNETOSPHERIC PHYSICS

It may be fairly stated that the first step in the physics of the magnetosphere was the work of Gilbert around 1600, who was inspired to recognize that Earth has an external magnetic field in the dipole form of the magnetic field around a lodestone. That is to say, Gilbert asserted that the magnetosphere exists. He supposed that Earth itself is a spherical lodestone. That Earth is not a ferromagnetic body (i.e. not a lodestone) was recognized only with the discovery of the Curie temperature in the late 19th century. But the existence of a three dimensional sphere of influence surrounding Earth stands on its own, regardless of the origin within Earth. In Gilbert's day the geomagnetic field was detected and measured through its mechanical effect on magnetic needles. It is interesting that the small geomagnetic fluctuations were detected mechanically and eventually recognized in the 19th century as following solar activity.

The corpuscular nature of the effects of solar activity were championed by Birkeland (1896). The auroral ray structure suggested to Birkeland the cathode ray streamers produced by the energetic electrons in a Crookes tube. His idea was simply that the Sun emitted electrons of such high energy (~ 200 Mev) as to penetrate through the geomagnetic field to

[1] Van Allen (1957, Meredith, et al, 1955) reached 100 km with a small rocket launched from a balloon at 60,000 ft, achieving the first direct detection of auroral particles.

produce the auroral streamers in the upper atmosphere of Earth. With this inspiration he constructed his ingenious and spectacular terrela experiment, using a small magnetized sphere representing Earth as the anode of a large cathode ray tube. The ultimate irrelevance of the experiment in no way detracts from the important conceptual step to the corpuscular nature of the aurora and corpuscular emission from the Sun that drives the aurora. We know now that the aurora is only an indirect consequence of the particles of kilovolt energies (the solar wind) from the Sun impacting the outer boundary of the magnetosphere. But Birkeland pointed the way with his idea that the aurorae are a direct manifestation of the impinging solar particles themselves. Birkeland observed the deflections of the geomagnetic field associated with auroral activity and deduced from Ampere's law that there are strong horizontal electric currents in the ionosphere flowing perpendicular to the auroral curtains. He recognized that the beginning and end points of the ionospheric currents must be supplied by electric currents flowing down along the magnetic field from above, now known as the Birkeland currents.

In 1918 Lindeman pointed out that the necessary large quantity of corpuscles implies near equal numbers of electrons and ion charges i.e. near electrical neutrality. Chapman and Ferraro (Chapman:1918, 1935; Chapman and Ferraro, 1931, 1932, 1940) picked up on these ideas, addressing the geomagnetic storm phenomenon. They postulated an electrically neutral plasma cloud ejected by the Sun at $\sim 10^3$ km/sec and impacting the geomagnetic field. They recognized that the impacting plasma would not penetrate immediately into the geomagnetic field but would compress the field on the sunward boundary behind a relatively sharp interface (magnetopause). The sudden commencement of the storm was evidently a consequence of the impact of a well defined front surface of the plasma cloud. The sustained increase of the horizontal component (the initial phase of the storm) was a consequence of compression of the geomagnetic field by the continuing pressure of the impacting plasma cloud. Chapman and Ferraro noted that the subsequent world wide decrease of the horizontal component (the main phase of the storm) is equivalent to adding a southward magnetic field in the space around Earth. According to Ampere's law a southward field at Earth would be caused by a westward ring current around Earth somewhere out in the magnetosphere. They suggested, then, that after a few hours the plasma impacting the geomagnetic field penetrated into the magnetosphere and somehow set up a westward ring current. Then the horizontal component of the geomagnetic field gradually relaxed back to normal as the ring current succumbed to electrical resistivity (i.e. collisions with the ambient atmosphere). Ampere's law allows no simple alternative. But in reading the later literature one had the impression that the ring current was the ultimate explanation of the main phase, i.e. there was no more physics involved than a westward current through, say, a hypothetical copper wire circling Earth at a distance of several Earth's radii, R_E . In fact, to have applied an emf to drive a current around a ring of wire would have created a magnetic field that springs outward from the wire, displacing the ambient field from the region around the wire, and compressing the ambient geomagnetic field to provide a world wide increase in the horizontal component (Parker, 1958b,f). Rapid reconnection (Parker, 1957b, 1958b) would soon merge the field of the ring current and the geomagnetic field, of course, providing the expected decrease of horizontal component for the main phase. Thus, one might have argued that the driven current provides both the initial and main phases of the geomagnetic storm, but these points were not generally appreciated.

IV. DYNAMICS OF A TENUOUS PLASMA

It should be kept in mind that the theory of the dynamical behavior of tenuous plasma in a magnetic field was only beginning to be understood in the early and middle fifties. Magnetohydrodynamics was understood to apply to dense gases and liquids, e.g. the liquid core of Earth, or the dense partially and fully ionized gases below the surface of the Sun, with only a slight resistive dissipation of the electric currents. However, the very tenuous ionized gases in space, where the collision rate is small compared to the cyclotron frequency, posed a different problem. For it was well known that the application of an electric field perpendicular to the ambient magnetic field produced no lasting electric current flow. The net effect was merely to set the plasma in motion with the electric drift velocity $\mathbf{u} = c\mathrm{E}\times\mathbf{B}/B^2$, so that the plasma moves on the local frame of reference in which there is no electric field. In that respect the tenuous plasma is the perfect insulator, as distinct from the dense collisional plasma which is an excellent conductor, with $\sigma \sim 2\times10^7\,T^{3/2}$ sec $^{-1}$. On the other hand, there is no dissipation of the current in a collisionless plasma so the effective resistivity is zero. The convention was to work with the conductivity tensor, including the Hall and Pedersen conductivities. Unfortunately the equations relating the current $\mathbf{j}$ with $\mathbf{E}$, $\mathbf{B}$, and the bulk plasma motions $\mathbf{v}$ were too complicated to provide any clear general concepts. Cowling (1957) gave a concise outline of the dynamical theory of the bulk motion of a plasma as it was understood at the time. In fact the equation of motion for $d\mathbf{v}/dt$ and the induction equation for $d\boldsymbol{B}/dt$ were magnetohydrodynamic in form (as emphasized by Schlueter(1950, 1952)) but the additional terms involving the various resistivities and thermal effects gave the impression of a much more complicated state. In fact it is widely believed up to the present day that magnetohydrodynamics does not provide a proper description for the large-scale dynamical behavior of the active magnetosphere (cf. Parker, 1996), so that one must deal directly with electric fields, electric currents, and a tensor conductivity. I was as puzzled as anyone in those early times, when it occurred to me that the extreme limit of an entirely collisionless plasma could be handled quite easily. The basic problem is the dynamics of the bulk motion $\mathbf{v}$ of the plasma and magnetic field $\mathbf{B}$ on a scale l large compared to the cyclotron radius of the ions and electrons. In the limit of large scale l the motion of the individual electron or ion can be written down quite generally using the guiding center approximation (Alfven, 1950). The particle motion consists, then, of the electric drift velocity $\mathbf{u}=c\mathbf{E}\times\mathbf{B}/B^2$ plus the so called gradient and curvature drifts, both of the order of the thermal velocity multiplied by R/l where R is the cyclotron radius. So in the limit of $l >> R$, the bulk motion of both electrons and ions is just $\mathbf{u}$. Summing over the difference in the motions of ions and electrons gives the current density $\boldsymbol{j}$. Writing Maxwell's equation as

$$\partial \boldsymbol{E}/\partial t = c\nabla\times\boldsymbol{B} - 4\pi\boldsymbol{j} \qquad (4)$$

and noting that $\boldsymbol{E} = -\mathbf{u}\times\boldsymbol{B}/c$ and $\partial/\partial t \sim u/l$, it is clear that the left hand side is small $O(u^2/c^2)$ compared to the first term $O(cB/l)$ on the right hand side, so that the left hand side can be neglected, leaving just Ampere's law to the order in u/c considered. Replacing $\boldsymbol{j}$ on the right hand side by the current carried by the electrons and ions yielded Newton's equations of motion from Ampere's law, viz.

$$NM\ d\mathbf{u}/dt = -\nabla_\perp(p_\perp + B^2/8\pi) + \{[(\boldsymbol{B}\cdot\nabla)\boldsymbol{B}]_\perp/4\pi\}[1-4\pi(p_{//}-p_\perp)/B^2] \qquad (5)$$

(Parker, 1957a). It is supposed that there are N singly charged ions per unit volume, each ion and its associated electron having a total mass M, so that the density of the plasma is NM. The subscripts $\perp$ and $//$ refer to the perpendicular and parallel directions relative to the magnetic field. Thus if $w_\perp$ and $w_{//}$ represent the perpendicular and parallel components of the individual particle thermal velocities, it follows that

$$p_\perp = \tfrac{1}{2}\Sigma M w_\perp^2 \qquad \text{and} \qquad p_{//} = \Sigma M w_{//}^2$$

where the sum is over all particles in a unit volume. The term $(p_{//} - p_\perp)$ represents the net centrifugal force exerted by the electrons and ions streaming along a curved magnetic field. It has the effect of opposing the tension $B^2/4\pi$ in the field around the radius of curvature $[(\boldsymbol{B}\cdot\nabla)\boldsymbol{B}]_\perp/B^2$ of the field lines. If $p_{//} - p_\perp$ is as large as $B^2/4\pi$, the centrifugal force cancels the effect of the tension $B^2/4\pi$.

At about the same time it was becoming clear from studies of the dynamical instabilities of anisotropic thermal velocity distributions (cf. Parker, 1958a, 1959a) that any strong anisotropy for which $p_{//} - p_\perp \sim B^2/8\pi$, is unstable on times comparable to the ion and electron plasma frequencies and cyclotron frequencies. Hence a strong anisotropy is rapidly dissipated, so that for the long term bulk behavior of large-scale plasma motions, one expects no important effects of thermal anisotropy. By the same token one expects no long term impediment to motion along a magnetic field. This follows from the fact that an isotropic thermal velocity distribution yields a uniform density distribution along an inhomogenious magnetic field. To a good approximation, then, the bulk motion of tenuous large-scale plasmas is described by the familiar momentum equation

$$\rho d\mathbf{v}/dt = -\nabla(p+B^2/8\pi)+(\mathbf{B}\cdot\nabla)\mathbf{B}/4\pi, \qquad (6)$$

and if we think there may be significant thermal anisotropy, it is obvious how to go about including the effects, writing $p_\perp$ and $p_{//}$, etc. The point is that anisotropy is a gas kinetic effect which in no way alters the Maxwell stress tensor

$$M_{ij} = -\delta_{ij}B^2/8\pi + B_iB_j/4\pi. \qquad (7)$$

The general statement of the momentum equation can be written

$$\partial\rho v_i/\partial t = \partial R_{ij}/_i\partial x_j + \partial M_{ij}/\partial x_j - \partial p_{ij}/\partial x_j - \rho\partial\Phi/\partial x_i \qquad (8)$$

where Φ is the gravitational potential, $R_{ij} = -\rho v_i v_j$ is the Reynolds stress tensor, and p_{ij} is the pressure tensor. This momentum equation is applicable to any plasma with a number density N sufficiently large that Nl^3 is very large compared to unity. In that case the equation is nothing more than the statement of conservation of momentum of the particles in a fixed cube of intermediate dimension $\lambda(Nl^3>>N\lambda^3>>1)$. The mean velocity of all particles in the cube is v_i while the thermal velocity w_i of the individual particle is measured relative to v_i. The acceleration of the center of mass of the particles is represented by $\partial v_i/\partial t$ and the calculation includes the transport of momentum by the streaming of particles in and out through the faces of the cube, giving the contributions R_{ij} and p_{ij} for the mean bulk motion and thermal motions, respectively. Noting that $\boldsymbol{E} = -\mathbf{u}\times\boldsymbol{B}/c$, it follows from Faraday's induction equation that

$$\partial\mathbf{B}/\partial t = \nabla\times(\mathbf{u}\times\mathbf{B}), \tag{9}$$

so that we have the familiar magnetohydrodynamic induction equation. Including collisions would, of course, introduce dissipation on the right hand side of the induction equation. However, dissipation in the tenuous plasmas in the magnetosphere is a small effect that can be neglected for most purposes. Only when the dense ionosphere is involved does resistive dissipation become significant. The reader may find it interesting to consult the development of the momentum equation by Chew et al (1956), based on the transverse and longitudinal invariants of the individual particle motions.

The conclusion (Parker, 1957a, 1960b) was that the large-scale behavior of the active terrestrial magnetosphere is described by the familiar magnetohydrodynamic equations. The field and plasma move together, except for the small gradient and curvature drifts. This result was both a revelation and a disappointment to me at the time. It was a revelation because at last I felt I understood the basis of magnetohydrodynamics of the tenuous plasma. The electric currents and various tensor Ohm's laws with the Hall and Pedersen conductivities were secondary. The basic dynamics concerns the direct interaction of the plasma inertia and pressure with the elastic magnetic field whose stresses are described precisely by the Maxwell stress tensor. It was a disappointment because I had hoped that my efforts would discover some exotic new effect arising in a truly collisionless plasma that was not to be found in the magnetohydrodynamics of an electrically conducting fluid. In retrospect, of course, it was clear that there was nothing missing from the magnetohydrodynamic equations that was anything more than the kinetic effects of an anisotropic thermal velocity distribution: The bottom line is simply that the magnetic field is carried bodily in the loval frame of reference in which the electric field vanishes. For the collisionless plasma that is the frame moving with the electric drift velocity $\mathbf{u}$.

We should not lose sight of the fact that the magnetohydrodynamic formulation does not apply to scales comparable to the cyclotron radii of the ions and electrons, nor to regions such as the spontaneous auroral current sheets.(Parker, 1972, 1994) wherein the number of electrons is too small to carry the enormous current densities required by Ampere's law without acceleration of ions and electrons to suprathermal velocities. In that case there is a significant electric field component $E_{//}$ along the magnetic field. The internal dynamical structure of the collisionless shock front is another example and one of my next efforts was to explore the parallel shock (Parker, 1959a) while others pursued the more difficult case of the perpendicular and the oblique shock. The collisionless shock introduced rapid acceleration of ions and electrons (Parker and Tidman, 1958a,b; Parker, 1959e) with implications for the bow shock of the magnetosphere, interplanetary shock fronts (Parker, 1961a,b,d, 1965a) and solar flares. Finally, the large scale magnetohydrodynamic deformation of the geomagnetic field, including internal Alfven waves, provides diffusion and acceleration of the trapped energetic particles (the Van Allen radiation belts) (Parker, 1960a, 1961b,c).

With the magnetohydrodynamic concepts in mind, it became clear that the decrease of the horizontal magnetic field at the surface of Earth, characterizing the main phase of the magnetic storm, is a consequence of some outward force that expands the elastic magnetic field outward from Earth. It goes without saying that the outward force exerted by the plasma on the field induces a westward ring current, but the essential physics is the outward force on the field. One of my early ideas (Parker, 1956) was that the ionosphere is heated so as to expand outward, lifting the magnetic field and thereby decreasing the field below the ionosphere. However further inquiry showed no such ionospheric effect.

V. SOLAR CORPUSCULAR RADIATION

Turning to the dynamics of the solar corpuscular radiation responsible for geomagnetic activity, I had the good fortune to learn from Sydney Chapman (1957, 1959) at the High Altitude Observatory that the high thermal conductivity and negligible radiative cooling provides a tenuous solar corona that extends well beyond the orbit of Earth. A discussion with Ludwig Biermann, who was visiting John Simpson at Chicago, emphasized that the anti-solar pointing of gaseous comet tails indicates that the Sun emits solar corpuscular radiation in all directions at all times (Biermann, 1951, 1952, 1957). After thinking about these ideas of Chapman and Biermann, it occurred to me that they were mutually exclusive. The idea of interplanetary space filled with a static tenuous coronal plasma and also with corpuscular radiation was not possible because of the powerful plasma instabilities that arise in any but a closely isotropic thermal velocity distribution. That is to say, the two-stream instability prevents one collisionless plasma from passing through another. Yet Biermann and Chapman had solid arguments for their conclusions. After some thought it became clear that the reconciliation was that the static corona near the Sun somehow became the solar corpuscular radiation at large distance from the Sun. With magnetohydrodynamics as the basis for the dynamics of both collisional and collisionless plasmas, it was only a matter of writing down the hydrodynamic equation for radial acceleration of the corona to see that Chapman's extended temperature produced a gentle outward acceleration of the nearly static corona close to the Sun, ultimately reaching supersonic velocity at large distance and fulfilling Biermann's inference from the behavior of comet tails that the Sun emits corpuscular radiation in all directions at all times at speeds of the general order of 500 km/sec (Parker, 1958d). To emphasize the purely hydrodynamical character of the phenomenon I suggested that it be called the "solar wind".

From the basic principles of magnetohydrodynamics it followed immediately that this more or less steady solar wind extends the weak fields of the Sun out through the solar system in an Archimedean spiral from the rotating Sun (Parker, 1958d). Variations in the wind speed complicate the basic spiral form of the field, with occasional blast waves (from flares, etc.) with well defined shock fronts (Parker,1961e,1963). It was immediately evident that the outward sweep of the irregularities provides the modulation of the cosmic rays studied by Forbush and Simpson, and others (Parker, 1958g, 1961e, 1963, 1965b, 1966c, 1967c; Jokipii and Parker, 1967). Thus the modern concept of the heliosphere sprang full blown from the hydrodynamic state of the solar corona. In particular, it is the solar wind streaming past Earth that confines and agitates the geomagnetic field in the manner described by Chapman and Ferraro, while stretching out the outer regions of the geomagnetic field in the anti-solar direction (Parker, 1958c).

The community reacted to the whole proposition with coordinated disbelief. This sociological phenomenon is more or less universal and every scientist should be aware of it. The basic paper (Parker, 1958d) was submitted to the Astrophysical Journal, edited by Prof. S. Chandrasekhar. Chandra came to my office one day and noted that the referees, all authorities in the field, assured him that the work was wrong. In view of that fact, did I really want to publish the paper? I replied that none of the referees had any explicit criticism of the physical arguments and mathematics in the paper, and how else did the referees propose to reconcile Chapman's inescapable extended corona with Biermann's inescapable continuous corpuscular radiation? Yes, I wanted to publish the paper. After a moments thought Chandra said, "Alright. I will publish it." And he did, undoubtedly incurring the annoyance, if not the wrath, of the referees.

A couple of years later Joseph Chamberlain gave a review lecture at the Spring meeting of the National Academy of Sciences in which he assured the audience that the proposed supersonic expansion of the solar corona at large distance was the result of an incorrect choice of the constant of integration. He went on to show that a very dense corona (10^{10} atoms/cm^3) with a limited heat supply provides only a slow subsonic expansion of some 20 km/sec at the orbit of Earth (Chamberlain, 1960,1961). I was not present at that Spring meeting, but several acquaintances later offered their sympathies. The solar wind was an interesting idea, but clearly not acceptable to the expert community. Only two or three colleagues seemed to understand the implications of the work of Chapman and Biermann and the elementary integration of the radial component of the momentum equation for an atmosphere strongly bound by gravity but with an extended temperature. Curiously, Chapman was not one of them.

Fortunately detection and then quantitative study of the solar wind was not many years away. Gringauz, et al (1961) found a continuing streaming of ions in interplanetary space with velocity in excess of 60 km/sec (> 15 ev/unit charge). Then Bridge, et al (1961) measured an intermittent quiet day solar wind from the direction of the Sun at 240 - 400 km/sec and 7 - 20 atoms /cm^3 as Explorer 10 skimmed along the flapping surface of the tailward magnetosphere. Finally the JPL plasma instrument of Mariner 2, bound for Venus, showed the continuous supersonic solar wind with velocities varying between 300 and 800 km/sec but always present with densities in the range 2 - 20 atoms/cm^3 (Snyder and Neugebauer, 1964). Ness, Scearce, and Seek (1964) found the interplanetary magnetic field to have the expected average inclination of about 45° to the radial direction. Meanwhile I collected my theoretical investigations of the solar wind and its effects all the way out to the interstellar wind into a monograph (Parker, 1963). Over the next several years the community came to accept the solar wind as coronal expansion. I do not recall who coined the term "heliosphere".

In retrospect this is a typical example of the universal sociology of scientific progress. It was repeated several times in later years. For instance, when I first pointed out the theoretical limits to the magnetic field of the Galaxy and noted the general dynamical instability (Parker, 1966d), the referee's report began with "I had always thought that Parker was competent, but..." The ideas are now generally accepted and applied to a variety of stellar and galactic circumstances. More recently I have shown that the nature of the Maxwell stress in a magnetic field is such that almost all magnetic field topologies develop internal tangential discontinuities (current sheets) as the field relaxes to equilibrium in an infinitely conducting fluid. It appears that this phenomenon may be the principal source of magnetic dissipation heating the X-ray corona of the Sun and other stars. The ideas are currently passing through the same negative phase, with six papers in the literature declaring, with "proofs", that the spontaneous formation of discontinuities is not possible, without noting the simple formal mathematical examples of the formation of discontinuities and without ever addressing the contradictions that arise were it not so. These papers are now widely quoted and , in fact, have been useful in writing my monograph (Parker, 1994) on the spontaneous formation of tangential discontinuities, because they exhibit the ambient confusion to which the writing is addressed. On the whole it is my impression that the negative reflex reaction of the community may have a positive effect, based on P. T. Barnum's dictum that bad publicity is better than no publicity. Certainly in the case of the solar wind it worked that way.

It should be emphasized that this is all in the normal course of events. It was interesting to read Baldwin's (1995) recent account of the flat rejection of his point, made some forty years ago, that the craters on the moon have all the detailed characteristics of impact craters as distinct from volcanic calderas. Then we recall the early reaction of the community to Walter Alvarez's association of the demise of the dinosaurs with the worldwide iridium layer and the implied impact of a massive boloid.

This is not at all a recent development. For instance, geomagnetic activity as a consequence of corpuscular radiation from activity on the Sun was proposed by Fitzgerald, Lodge, Birkeland, and others in the last decade of the 19th century. On the other hand Kelvin's "proof" in 1892 (based on a hard vacuum with no charged particles in space) was accepted by the community as overwhelming proof that no solar-terrestrial connection exists beyond the steady sunlight. It was 25 years before the pursuit of solar corpuscular radiation was revived by Sydney Chapman.

In 1912 Alfred Wegener pointed out the extensive evidence for the relative drift of the continents. The idea was explored and debated until 1924 when Harold Jeffries "proved" that the crust and mantle are too rigid to allow relative motion, based on his estimates of the rigidity necessary to support the Himalayas and completely overlooking the fact that there are forces strong enough to buckle the crust and raise the Himalayas in the first place.

At the present time the inescapable conclusion that the well documented variations of the brightness of the Sun and similar stars (by several tenths of a percent over periods of years and decades) may have some noticeable effect on terrestrial climate is undergoing similar "disproof". The interested reader is referred to such articles as (Thomson, 1995) offered as the definitive proof that the variations of the Sun have no discernible effect on mean temperatures at Earth. The "disproof" in this case is founded on the curious assumption that a change in the brightness of the Sun would have greater effect on the mean temperature in the Summer than in the Winter because the Sun shines for more hours per day in the Summer. Shining for more hours contributes to the higher Summer temperatures compared with Winter, but Thomson failed to realize that the hours of sunshine are unaffected by the brightness of the Sun. In fact terrestrial climate is a sufficiently complicated phenomenon, with heat transported over latitude by wind and ocean currents and with sunlight blocked by cloud cover, that we are in no position to state the relative contributions of a changing solar brightness to Summer and Winter temperatures, any more than it is possible to give a reliable statement of the precise warming brought about by the accumulating anthropogenic greenhouse gases.

There seems to be an overweening reluctance in the scientific community to admit that nature presents us with puzzles whose solution requires new ideas to be introduced into the repertoire. A reasoned caution with new concepts is part of the scientific process, but eager rejection tells a sad tale of the common scientific mind. All too often "truth" is defined as the consensus of our peer group. So we should have no illusions about ourselves, nor should we be surprised at the normal course of scientific events.

I next undertook a study of the dynamics of the outer boundary of the magnetosphere in response to the impact of a beam of protons and electrons, representing an idealization of the incident supersonic solar wind (Parker, 1958c). With the assumption of specular reflection of the incident beam, it was easy to show that the boundary and the confining wind are subject to the familiar Kelvin-Helmholtz instability. The instability suggested an effective viscosity across the magnetopause, converting the incident kinetic energy into more chaotic thermal motions and accelerating some particles to high energy by the Fermi mechanism. The individual particle motions within the idealized shear layer at the boundary showed elliptical orbits and gave an effective kinematic viscosity of the order of the wind velocity multiplied by the ion cyclotron radius. However, there was a basic difficulty arising from the fact that the aurora and the main (expansion) phase of the magnetic storm occurred deep within the magnetosphere, on field lines that cross the equatorial plane at r sometimes as samll as 2-3 R_E where R_E is the radius of Earth. Using the idealized Chapman-Ferraro plane boundary for the sunward magnetopause it was easy to show that compression to $r = 2R_E$ would require an interplanetary blast wave of immense density (N~10^5 atoms/cm^3) and velocity (v~10^3 km/sec), as well as providing an increase $\Delta B/B$~0.02 in the vicinity of Earth. Neither so massive a wind nor so large an increase ΔB has ever been observed. So I turned to the unstable character of the magnetopause and conceived the idea that tongues of solar wind plasma are injected through the magnetopause and deep into the magnetosphere. I wondered if the particle acceleration at the magnetopause provided the aurora, with the fast particles coming in with the intrusive tongues of plasma. In retrospect it seems so naive. What pressure besides the impact pressure of the wind could force the tongues of plasma deep into the magnetosphere in opposition to the increasing magnetic pressure? And such impact pressures were beyond anything to be expected in the solar wind. Then I conjectured that the injected solar wind gas might be trapped in the geomagnetic field; accumulating to sufficient mass as to give a net outward centrifugal force (for gas at $r > 6.2\ R_E$) that expands the magnetosphere to provide the main phase of the geomagnetic storm. It would have been more astute to recognize that there was some missing physics.

Returning, then, to the unstable magnetopause I noted that "the solar wind tends to carry away the lines of force of the outer geomagnetic field, just as a high wind blows smoke away from a chimney"(Parker, 1958c). Gold (1959a,b) and Johnson (1960) described the general comet or tear drop shape of the magnetosphere as a whole, while others then went ahead and formally calculated the general comet shape of the sunward magnetopause and the extended geomagnetic tail.

Then the missing physics began to appear. Dungey (1961, Parker, 1966a) emphasized the importance of active reconnection of a southward interplanetary magnetic field component with the geomagnetic field at the sunward magnetopause, stretching magnetic field lines into the geotail. This becomes the magnetic substorm in the present understanding of geomagnetic activity. Axford and Hines (1961) worked out the important consequences of the overall convection of the magnetosphere driven by viscous forces across the unstable magnetopause. It has been realized since that the reconnected flux bundles stretched into the geotail are the principal driver of the enhanced convection during active times. But whatever the driver, Axford and Hines showed that particles from the solar wind, or particles accelerated by the solar wind, are convected deep into the magnetosphere, thereby sparing us the unworkable hypothesis that fingers of solar wind are directly forced in to the auroral field lines and ring current field lines deep in the magnetosphere.

VI. THE GEOMAGNETIC STORM

To go back a little in time, I had a continuing interest in the main phase of the geomagnetic storm, in which the field is displaced outward from Earth to decrease the horizontal component at the surface of Earth. Singer (1957) had pointed out that the gradient drift of the energetic particles (e.g. the Van Allen radiation) trapped in the geomagnetic field provides a westward directed electric current, with the effect of decreasing the horizontal component in the vicinity of Earth. This idea fitted naturally and simply into the general picture. Indeed, what else could be the cause of the main phase of the magnetic storm? It was a simple matter for me to work out the small perturbation ΔB in the neighborhood of a magnetic dipole as a consequence of an equatorial concentric ring of radius r of trapped charged particles. The particles with velocity $w_\perp$, mass M, and charge q drift westward with the gradient drift velocity $3w_\perp R/2r$ where R is the cyclotron radius $Mw_\perp c/qB(r)$ and the geomagnetic field $B(r)$ is $B_E(R_E/r)^3$ where B_E(~ 0.3 gauss) is the field intensity on the magnetic equator of Earth ($r = R_E$). Besides the westward drift, each trapped particle has a diamagnetic moment $\mu = Mw_\perp^2/2B(r)$. The net effect of all this can be written

$$\Delta B/B_E = 2E/3E_m \tag{10}$$

where E is the total kinetic energy of all the particles and $E_m = B_E^2R_E^3/3$ is the total magnetic energy (0.8×10^{25}ergs) of the dipole field outside the surface $r = R_E$. It is interesting to note that ΔB, which is parallel to the dipole moment, depends only on the total kinetic energy of the trapped particles and not at all on the distance at which the particles are trapped, subject to the restriction that the particles only slightly perturb the field.

The centrifugal force of any particle motion $w_{//}$ parallel to the field provides an outward force on the field of $3Mw_{//}^2/r$ per particle. This force is exerted on the field by a westward current I, producing a perturbation field

$$\Delta B/B_E = 2E_{//}/E_m \tag{11}$$

in the vicinity of the dipole by particles near the equatorial plane, where $E_{//}$ is the total kinetic energy of the parallel motion. The general relation for particles distributed along the field with both $w_\perp$ and $w_{//}$ was worked out by Sckopke (1966; Parker and Stewart 1967; Olbert et al, 1968)

During this same period of time Alexander Dessler (at Lockheed, Palo Alto, California) had also been thinking about the magnetic storm. In particular he realized that the abrupt onset of the compression of the geomagnetic field represents the arrival at the ground based magnetometer of the signal from the progressive points of impact on the magnetopause as the interplanetary blast wave sweeps over the magnetopause (Dessler, 1958) He also learned that laboratory measurements of charge exchange between protons of a kev energy and a neutral hydrogen atmosphere show an effective cross section ten or more times larger than the geometrical cross section (10^{-16}cm^2) of the individual atoms and very much larger than the Coulomb collision cross section between two protons. So it appeared that the trapped particles were largely protons and that charge exchange was the principal dissipation channel for the trapped energetic protons that produce the main phase of the magnetic storm. Conversations with Dessler led to combining our results into a paper on the geomagnetic storm (Dessler and Parker, 1959, 1968; Parker, 1962a,b, 1969a). We pointed out that the time of sudden commencement of a geomagnetic storm differs at various locations around Earth as a consequence of the differing Alfven transit times as the interplanetary shock front sweeps over the sunward magnetopause (Dessler, 1958; Frances et al, 1959). We followed up the next year (Dessler et al, 1960) with detailed ray tracing, showing how it takes about 10^2 sec for the shock front to engulf the magnetosphere and 10^2 sec more in differing Alfven transit times to the surface of Earth, with the night side of Earth experiencing the impulse later than the day side. We suggested that solar wind ions (protons) are injected into the geomagnetic field to radii of 3-5 R_E to produce the main phase of the geomagnetic storm, as described by equations (10) and (11) above. We emphasized the magneto-hydrodynamic nature of the storm phenomenon, beginning with equations (5), (6), and (9).

As already noted, Dessler had come across the laboratory results of Fite et al (1958) showing that the charge exchange cross section for proton-hydrogen atom collisions varied inversely with the relative velocity over the energy range 10^2 to 2×10^4 ev. It followed that the life τ of a proton in an ambient neutral hydrogen gas of number density N H atoms/cm^3 is independent of the proton velocity in the kev energy range. The laboratory measurements gave the simple result that

$$\tau \approx 10^7/N \text{ sec.}$$

The large charge exchange cross section dominates other losses for a uniform current of protons circling around Earth, so that the rate of decay of the main phase is given approximately by the charge-exchange life τ (Parker, 1966b). The characteristic relaxation time of the axisymmetric main phase of a large storm is about 1 day, say 10^5 sec. Therefore the ambient density must be about 10^2 H atoms/cm^3, arising at $r \approx 3R_E$.

On the basis of charge exchange it follows that the rate of decay of the main phase should be largest for the great storms and much less for weak storms, assuming that the vigor of the solar wind implants the particles deeper in the magnetosphere to create the stronger storm. The initial decay of a storm proceeds more rapidly than later on as the innermost ions are the first to charge exchange, leaving the more distant ions in the tenuous outer neutral hydrogen to charge exchange more slowly. Both these theoretical characteristics conform to the observed facts.

It also followed that the decay of magnetic storms should proceed more slowly during the years of minimum solar activity when the ambient neutral hydrogen atmosphere is about one third as dense as during activity maximum. This is in good agreement with the recorded decay of magnetic storms. We estimated that the typical storm particles repose at $r = 4R_E$, based on the observed decrease of the vertical field component of the surface at Earth at ±60° latitude and on the characteristic 3-day decay time of the typical storm. These simple facts collectively suggested that the main phase of the geomagnetic storm is primarily a consequence of protons with energies of the general order of a kev, i.e. velocities of the order of 500 -1000 km/sec, trapped in a broad band at a few Earth radii. A couple of years later we explored the possibility that the protons are accelerated *in situ* by magnetohydrodynamic waves and shocks during the initial phase of the storm, thereby creating the subsequent main phase as the wind pressure subsides (Dessler et al 1961). We estimated that most of the acceleration occurs on the night side where the magnetic field is somewhat less, accelerating the ambient ions to velocities ~10^3 km/sec or 6 kev, in order of magnitude. It has been gratifying to find that the information on the geomagnetic storm subsequently accumulated from particle detectors, magnetometers, and electric field detectors on orbiting spacecraft has largely confirmed this basic picture over the decades since we first wrote about the geomagnetic storm. The acceleration of particles is more varied and complex than we imagined, picking up terrestrial atmospheric ions as well as protons, and is still not quantified in spite of the best efforts of many investigators. The charge exchange with the ambient outer terrestrial ionosphere has been forgotten and reinvented in recent years so that Dessler and I feel pretty good about our old paper of 1959.

VII. THE MAGNETO PAUSE

My final foray into magnetospheric physics in those early years was an investigation of the submagnetohydrodynamic microstructures, i.e. the plasma physics, of the magnetopause. The problem is interesting at the most elementary level because the geomagnetic field lines at the magnetopause extend down into the ionosphere, so that under steady conditions there are no significant electric fields perpendicular to the magnetic field lines in the stationary magnetopause. Any such electric fields would be short circuited by a static conducting ionosphere. It follows, then, that the ions and electrons of the impacting solar wind each penetrate a distance comparable to their respective cyclotron radius. So the ions and electrons part company, with the ions penetrating much deeper ($\sim 10^2$ km) than the electrons.

Now the solar wind has a velocity component $v_{//}$ parallel to the geomagnetic field at the magnetopause, forming a current layer about 10^2 km thick. Ampere's law stipulates that there is a magnetic field generated by the parallel current, and that magnetic field lies in the plane of the magnetopause and perpendicular to the current, i.e. perpendicular to the initial magnetic field. That is to say, the magnetopause represents a rotational layer of field, with the field direction turning from the normal field direction on the inner side to essentially perpendicular to the incident solar wind at the outer surface. The curious feature is that the perpendicular field component produced by the penetrating ions deflects the ions inward rather than providing the usual outward reflection. Thus the perpendicular field component is not confined by the impact of the ions. Indeed, the contrary. The magnetic flux in the magnetopause is continually expelled into the wind sliding along the magnetopause. The wind then carries the expelled field into the geotail, of course. The perpendicular component increases outward, according to Ampere's law. The scenario suggests that there is no steady equilibrium (stable or otherwise) of the magnetopause when we take account of the solar wind velocity component parallel to the geomagnetic field at the magnetopause. We have suggested that the effect plays an important role in the "viscous" coupling of the solar wind to the magnetopause.

One may argue that the parallel electric current of the impacting ions induces an equal and opposite current in the ambient magnetospheric plasma. But any such induced current closes through the ionosphere, where the current is subject to resistivity. Hence it soon decays, and we are left with our parallel current and related field in the magnetopause. So there is no evident stationary state for the magnetopause. Indeed, this condition would seem to apply to all circumstances in which there is a rapid plasma flow along a magnetic surface from which the field lines connect into a resistive layer. The effect is not included in my own earlier development (Parker, 1957a) of the large-scale ($L >> R$) dynamics of plasmas, nor are we aware that it has been developed beyond our own simple demonstration (Parker, 1967a,b, 1968, 1969b; Lerche, 1967, 1968; Ferraro and Davies, 1968; Hurley, 1968). Lerche and Parker (1967, 1970) suggested the occurrence of the effect in surges of gas upward along the magnetic fields in the Sun, where the photosphere is the effective resistive layer. Eviatar and Wolf (1968) suggested that a sharp interface may be so unstable that the question of the stationary equilibrium is academic, and if an equilibrium does not exist, it has little or no effect on the chaotic instability. Whether one adopts that point of view or whether the parallel current causes the surface field to be expelled outward into the solar wind, it appears that there is an effective steady state erosion of magnetic flux from the magnetopause by a solar wind that is otherwise steady. The eroded field lines are carried back into the geotail. It would be interesting to know the rate at which flux is eroded under quiet conditions in the wind. The problem has been taken up again in recent years in the simple case that the magnitude of the magnetic field is the same on both sides of the velocity shear layer (Roth et al, 1996). This avoids the erosion and expulsion of magnetic field into free space, of which the authors seem to be unaware, although the calculations nicely demonstrate the rotation of the field direction.

VIII. EPILOGUE

My adventures with the geomagnetic field have provided a rewarding experience over the years. Direct quantitative observational results were forthcoming as the space age got underway, so it was possible to do hard science, with the expectation that a theoretical conjecture could soon be tested and affirmed or discarded. Some questions are sufficiently subtle that they remain unanswered for extended periods, e.g. the precise internal dynamics of the magnetopause. Many of the lingering theoretical questions are academic, but no less interesting, if their consequences are so subtle as to be undetected up to the present time. But I still wonder about the dynamical state of the magnetopause and the ejection of flux bundles into the passing solar wind. So the magnetosphere has a continuing fascination for me and I recently reactivated my old acquaintance with the magnetosphere with a critical review of the present state of the theory of magnetospheric activity (Parker, 1996). It is my impression that the substantial progress of the subject has been impeded by a reluctance to

geomagnetic storm main phase ring current, *J. Geophys. Res.*. *66*, 3631, 1961.

Dessler, A. J. and E. N. Parker, Hydromagnetic theory of geomagnetic storms, *J. Geophys. Res.*, *64*, 2239, 1959.

Dessler, A. J. and E. N. Parker, corrections to paper by A. J. Dessler and E. N. Parker, Hydromagnetic theory of geomagnetic storms, . *J. Geophys. Res.* *73*, 3091, 1968.

Dungey, J. W., Interplanetary magnetic fields and the auroral zones, *Phys. Rev. Letters*, *6*, 47, 1961.

Elsasser, W. M., Induction effects in terrestrial magnetism, *Phys. Rev.*, *69*, 106, 1945.

Elsasser, W. M., Induction effects in terrestrial magnetism, Part II. The secular variation, *Phys. Rev.* ,*70*, 202, 1946.

Elsasser, W. M., The earth's interior and geomagnetism, *Rev. Mod. Phys.* *22*, 1, 1950.

Elsasser, W. M., Dimensional relations in magneto-hydrodynamics, *Phys. Rev.*, *95*, 1, 1954.

Elsasser, W. M., Hydromagnetics I , A review., *Amer. J. Phys.*, *23*, 590, 1955.

Elsasser, W. M., Hydromagnetic dynamo theory, *Rev. Mod. Phys.*, *28*, 135, 1956a.

Elsasser, W. M., Hydromagnetism II. A review, *Amer. J. Phys.*, *24*, 85, 1956b.

Eviatar, A. and R. A. Wolf, Transfer processes in the magnetopause, *J. Geophys. Res.*, *73*, 5561, 1968.

Fan, C. Y., P. Meyer, and J. A. Simpson, Rapid reduction of cosmic radiation intensity measured in interplanetary space, *Phys. Rev. Letters*, *5*, 269, 1960a.

Fan, C. Y., P. Meyer,and J. A. Simpson, Experiments on the eleven year changes of cosmic ray intensity using a space probe, *Phys. Rev. Letters*, *5*, 272, 1960b.

Ferraro, V. C. A. and C. M. Davies, Discussion of paper by E. N. Parker, Confinement of a magnetic field by a beam of ions, *J. Geophys. Res.*, *73*, 3605, 1968.

Fite, W. L., T. R. Brackman, and W. R. Snow, Charge exchange in proton-hydrogen atom collisions, *Phys. Rev.*, *112*, 1161, 1958.

Francis, W. E., M I. Green, and A. J. Dessler, Hydromagnetic propagation of sudden commencements of geomagnetic storms, *J. Geophys. Res.*,*64*, 1643, 1959.

Gold, T., Magnetic field in the solar system, *Nuovo Cimento Suppl. X*, *13*, 1959a.

Gold, T., Motions in the magnetosphere of earth, *J. Geophys. Res.*, *64*, *1219*, 1959b.

Gringauz, K. I., Bezrukikh, V. V., Ozerov, V. D., and Ribchinsky, R. E., Study of the interplanetary ionized gas, high energy electrons, and solar corpuscular radiation by means of three electrostatic traps for charged particles on the second Soviet cosmic rocket, *Soviet Phys. Doklady*,*5*, 361, 1960.

Hurley, J. P., Discussion of paper by I. Lerche, On the boundary layer between a warm streaming plasma and a confined magnetic field, *J. Geophys. Res.*, *73*, 3602, 1968.

Johnson, F. S., The gross character of the geomagnetic field in the solar wind, *J. Geophys. Res.* *65*, *3049*, 1960.

Jokipii,J. R. and E. N. Parker, Energy changes of cosmic rays in the solar system, *Planet. Space Sci.*, *15*, 1375, 1967.

Lerche, I., On the boundary layer between a warm streaming plasma and a confined magnetic field, *J. Geophys. Res.*, *72*, 5295, 1967.

Lerche, I., Reply to Hurley, 1968, *J. Geophys. Res.*, *73*, 3602, 1968.

Lerche, I. and E. N. Parker, Nonequilibrium and enhanced mixing at a plasma field interface, *Astrophys. J.* *150*, 731, 1967.

Lerche, I. and E. N. Parker, Comments on "Steady state charge neutral models of the magnetosphere, *Astrophys. Space Sci.*, *8*, 140, 1970.

Lundquist, S., Studies in magnetohydrodynamics, *Ark. Fys.*, *5* (15) , 297, 1952.

Meredith, L. H., M. B. Gottlieb, and .J. A. Van Allen, Direct detection of soft radiation above 50 kilometers in the auroral zone, *Phys. Rev.*, *97*, 201, 1955.

Meyer, P., E. N. Parker, and J. A, Simpson, Solar cosmic rays of February, 1956 and their propagation through interplanetary space, *Phys. Rev.*, *104*, 768, 1956.

Ness, N. F., Scearce, C. S., and Seek, J. B., Initial results of the IMP 1 magnetic field experiments, *J. Geophys. Res.*, *69*, 3531, 1964.

Olbert, S., G. L. Siscoe, and V. M. Vasyliunas, A simple derivation of the Dessler-Parker-Sckopke relation, *J. Geophys. Res.*, *73*, 1115, 1968.

Parker, E. N., The formation of sunspots from the solar toroidal field, *Astrophys. J.*,*121*, 491, 1955a.

Parker, E. N., Hydromagnetic dynamo models, *Astrophys. J.*, *293*, 1955b.

Parker, E. N., On the geomagnetic storm effect, *J. Geophys. Res.* *61*, 625, 1956.

Parker, E. N., Newtonian development of the dynamical properties of ionized gases of low density, *Phys. Rev.*, *107*, 924, 1957a.

Parker, E. N., Sweet's mechanism for merging magnetic fields in conducting fluids, *J. Geophys. Res.*, *62*, 509, 1957b.

Parker, E. N., The solar hydromagnetic dynamo, *Proc. Nat. Acad. Sci.*, *43*, 8, 1957c.

Parker, E. N., Dynamical instability in an anisotropic ionized gas of low density, *Phys. Rev.*, *109*, 1874, 1958a.

Parker, E. N., Electrical conductivity in the geomagnetic storm effect, *J. Geophys. Res.*, *63*, 437, 1958b

Parker, E. N., Interaction of the solar wind with the geomagnetic field, *Phys. Fluids*, *1*, 171, 1958c.

Parker, E. N., Dynamics of the interplanetary gas and magnetic fields, *Astrophys. J.*, *128*, 644, 1958d.

Parker, E. N., Suprathermal particles II: Electrons, *Phys. Rev.*, *112*, 1429, 1958e.

Parker, E. N., Inadequacy of ring current theory for the main phase of a geomagnetic storm, *J. Geophys. Res.*, *63*, 683, 1958f.

Parker, E. N., Cosmic ray modulation by the solar wind, *Phys. Rev.*, *110*, 1445, 1958g.

Parker, E. N., Plasma dynamical determination of shock thickness, *Astrophys. J.*, *129*, 217, 1959a.

Parker, E. N., Auroral phenomena, *Proc, IRE*, *47*, No. 2, 239, 1959b.

Parker, E. N., Geomagnetic fluctuations and the form of the outer zone of the Van Allen radiation belt, *J. Geophys. Res., 65,* 3117, 1960a.

Parker, E. N., Solar, planetary, and interplanetary magnetohydrodynamics, *Proc. Woods Hole Symp. on Plasma Phys.,* 10-14 June 1958, Addison Wesley, Reading, Mass., ed. F. H. Clauser, 1960b.

Parker, E. N., Quasi-linear model of plasma shock structure in a longitudinal magnetic field, *Nucl. Energy, Part C, Plasma Phys., 2,* 146, 1961a.

Parker, E. N., Effect of hydromagnetic waves in a dipole field on the longitudinal invariant, *J. Geophys. Res., 66,* 693, 1961b.

Parker, E. N., The distribution of trapped particles in a changing magnetic field, *J. Geophys. Res., 66,*2641, 1961c.

Parker, E. N., Transresonant electron acceleration, *J. Geophys. Res. 66,* 2673, 1961d.

Parker, E. N., Sudden expansion of the corona following a large solar flare and the attendant magnetic field and cosmic ray effects., *Astrophys. J., 133,* 1014, 1961e.

Parker, E. N., Theory of magnetic storms, *J, Phys. Soc. Japan, Suppl. A-1, 17,* 199, 1962a.

Parker, E. N., Dynamics of the geomagnetic storm, *Space Sci. Rev., 1,* 62 1962b.

Parker, E. N., *Interplanetary Dynamical Processes,* Interscience Div. John Wiley, New York, 1963.

Parker, E. N., Interplanetary origin of energetic particles, *Phys. Rev. Letters, 14,* 55, 1965a.

Parker, E. N. The passage of energetic particles through interplanetary space, *Planet. Space Sci., 13,* 9, 1965b.

Parker, E. N., Particle effects in the geomagnetic field, in *Radiation Trapped in the Earth's Magnetic Field,* D. Reidel, Dordrecht, ed. B. M. McCormac, p 302, 1966a.

Parker, E. N., Nonsymmetric inflation of a magnetic dipole, *J. Geophys. Res., 71,* 4485, 1966b.

Parker, E. N., The effect of adiabatic deceleration on the cosmic ray spectrum in the solar system, *Planet. Space Sci., 14,* 371 1966c.

Parker, E. N., The dynamical state of the interstellar gas and field, *Astrophys. J. 145,* 1966d.

Parker, E. N., Confinement of a magnetic field by a beam of ions, *J. Geophys. Res., 72,* 2315, 1967a.

Parker, E. N., Small-scale nonequilibrium of the magnetopause and its consequences, *J. Geophys. Res., 72,* 4365, 1967b.

Parker, E. N., Reply to Ferraro and Davies, *J. Geophys. Res., 73,* 3607, 1968.

Parker, E. N., Solar wind interaction with the geomagnetic field, *Rev. Geophys., 7,* 3, 1969a.

Parker, E. N., Dynamical properties of the magnetosphere, in *Physics of the Magnetosphere,* D. Reidel, Dordrecht, eds. R. L. Carovillano, J. F. McClay, and H. D. Radoski, p 3, 1969b.

Parker, E. N., Topological dissipation and the small-scale fields in turbulent gases, *Astrophys. J., 174,* 499, 1972.

Parker, E. N., *Spontaneous Current Sheets in Magnetic Fields,* Oxford University Press, New York, 1994.

Parker, E. N., The alternative paradigm for magnetospheric physics, *J. Geophys. Res.,101,* 10587, 1996.

Parker, E. N. and H. A. Stewart, Nonlinear inflation of a magnetic dipole, *J. Geophys. Res., 72,* 5287, 1967.

Parker, E. N. and D. A. Tidman, Suprathermal particles I., *Phys, Rev., 111,* 1206, 1958a.

Parker, E. N. and D. A. Tidman, Suprathermal particles II., *Phys. Rev. 112,* 1048, 1958b.

Roth, M., J. De Jager, and M. M. Kuznetsova, Vlasov theory of the equilibrium structure of tangential discontinuities in space plasmas, *Space Sci. Rev., 76,* 251, 1996.

Schlueter, A. Dynamik des plasma. 1 Grundleichungen, plasma in gekreuzten feldern, 72, 1950.

Schlueter, A., Plasma im Magneticfeld, *Ann. Physik, 10,* 422, 1952.

Sckopke, N., A general relation between the energy of trapped particles and the disturbance field near Earth, *J. Geophys. Res., 71,* 3125, 1966.

Simpson, J. A., Cosmic radiation intensity time variations and their origin. II The origin of 27- day variations., *Phys. Rev., 94,* 426, 1954.

Simpson, J. A., The cosmic radiation and solar terrestrial relationships, *Ann. de Geophys., 11,* 305, 1955.

Simpson, J. A., H. W. Babcock, and H. D. Babcock, Association of a "unipolar" magnetic region on the Sun with changes of cosmic ray intensity, *Phys. Rev., 98,* 1402, 1955.

Singer, S. F., A new model of magnetic storms and aurora, *Trans. Amer. Geophys. Union, 38,* 175, 1957.

Snyder, C. W. and Neugebauer, M., Interplanetary solar wind measurements by Mariner 2, *Space Research,4,* 89, 1964.

Thomson, D. J., The seasons, global temperature, and precession, *Science, 269,* 59, 1995.

Van Allen, J. A., Direct detection of auroral radiation with rocket equipment, *Proc. Natl. Acad. Sci., 43,* 57, 1957.

E. N. Parker, Enrico Fermi Institute, University of Chicago, 933 East 56th Street, Chicago, Il 60637

The Role of the DISCOVERER Program in Early Studies of the Magnetosphere

Joseph B. Reagan

Retired
Advanced Technology Center
Lockheed Martin Missiles & Space
Palo Alto, California 94304

The DISCOVERER/CORONA program was the world's first operational photo satellite reconnaissance system. Over the August 12, 1960 to May 25, 1972 operational lifetime of the program, there were 130 DISCOVERER/CORONA satellite flights. In addition to the primary mission the DISCOVERER/ CORONA satellites hosted a wide variety of piggy-back scientific payloads. The frequent space flight opportunities in these early years of magnetosphere exploration led to many scientific discoveries and provided the foundation for several successful space scientists and institutions that have persisted to the present.

1.THE DISCOVERER/CORONA PROGRAM

In February 1958 President Dwight D. Eisenhower formally endorsed the creation of a satellite imaging reconnaisance system that would take pictures from space as it passed over the Sino-Soviet bloc. The code name given to the highly secret program was CORONA. The satellite periodically would deorbit a capsule with film which would be sent to the National Photographic Interpretation Center for imagery analysis. Eisenhower's bold decision was an attempt to counter the effects of the Iron Curtain which had closed the West's view into the communist world. Nikita Khruschchev of the Soviet Union had rejected Eisenhower's 1955 "Open Skies" proposal that was to be an essential basis for mutual arms control. In addition there was a growing US public concern over a perceived "missile gap" with the Soviet Union. US policymakers were under pressure to obtain timely, accurate and comprehensive information about world events, especially those occurring in the Soviet Union. High altitude aircraft and balloons provided limited information. The objective of the CORONA program was to use a space platform to acquire photographic intelligence to satisfy what was viewed as a critical requirement.

Eisenhower's decision was also bold from the perspective that such a system in February 1958 was based purely on theoretical concepts that were yet to be demonstrated using technologies that were not proven nor for which hardware existed. Issues that are taken for granted today had yet to be proven. The Soviet Union had successfully orbited the Sputnik satellite a few months earlier on October 4, 1957 and the Explorer I scientific satellite was successfully launched by the US on January 31, 1958; but even if you could launch a camera into orbit, would it work in the space environment about which relatively little was known? Remember that film is notoriously sensitive to radiation and the radiation belts were not discovered by James Van Allen until May 1,1958. If you took pictures from a satellite could you see through the atmosphere? Would the space platform be steady enough to obtain decent resolution? Could you control the satellite in orbit and take pictures when desired? Could you de-orbit through the atmosphere a

Discovery of the Magnetosphere
History of Geophysics Volume 7

capsule containing sensitive film? Could you capture the capsule in the air or at sea without degrading the film quality? These were formidable technological challenges to the team that developed the CORONA program.

The CORONA program was managed jointly by the Central Intelligence Agency (CIA) and the US Air Force with the funding source for development being the Advanced Research Projects Agency (ARPA). The Air Force Ballistic Missile Division (BMD) was responsible for the development of the space vehicle and for launch, tracking and, in conjunction with the US Navy, for capsule recovery. The CIA was responsible for the development of the photo-reconnaissance equipment. The industry team that supported the US Government consisted of Lockheed Missiles and Space Company, Itek Corporation, Fairchild Camera and Instrument Corporation, Eastman Kodak, General Electric, and Douglas Aircraft Company. Lockheed had broad responsibilities as the technical director and integrator of all equipment other than the Thor booster which was the responsibility of Douglas Aircraft. Lockheed also developed the Agena orbiting upper stage and integrated and led the test, launching and on-orbit control operations. Itek and Fairchild developed the camera system and General Electric was the contractor for the recovery capsule. Kodak supplied the film and assisted the government in the development of film processing .

The government/industry team worked feverishly to develop an operational system because of the critical national need. The first attempted but unsuccessful launch of the system, announced as DISCOVERER I, was on February 28, 1959, only one year after go-ahead. The next 12 launches were also unsuccessful for a variety of reasons but the team learned from each failure. Finally on August 12, 1960 the gold-plated capsule from DISCOVERER XIII, without any film, was presented to President Eisenhower in a White House public ceremony. The capsule had been released from the three-axis stabilized Agena spacecraft to effect atmospheric reentry near Hawaii. The gold plating enabled the capsule to survive the high temperatures of reentry through the atmosphere. A deployed parachute slowed the capsule descent to enable a specially equipped USAF C-119 cargo aircraft to snag the parachute in mid-air. The mid-air capture did not succeed on this mission and the capsule, which splashed down in the Pacific Ocean some 330 nautical miles northwest of Hawaii, was retrieved by a Navy helicopter and deposited on the deck of the surface recovery ship Haiti Victory. This capsule, the first object ever to be returned from space, currently resides in the Smithsonian National Air and Space Museum. During the next flight of DISCOVERER XIV on August 18, 1960, the first imagery of the earth from space was obtained. The era of space reconnaissance was born. Over the next year launches were as frequent as monthly but with mixed success. By early 1962, however, the launch rate was up to 20 per year and the success rate was up to 85 percent. From that time forward until the final mission of CORONA on May 25, 1972 the flights were highly successful. Over 130 satellites were placed in orbit over the 12 year lifetime of the program. Image resolution on the ground improved from 7.5 meter in the early missions to 1.8 meter or better in later flights. The CORONA program provided essential information for US policymakers in 1) exposing the "missile gap" myth 2) monitoring arms control and 3) detecting nuclear proliferation.

On February 22, 1995, 35 years after initiation, President William J.Clinton declassified the CORONA program and made available approximately 2 million linear feet of reconnaissance film acquired by the program for scientific and academic purposes [*McDonald, 1995a*]. This imagery has the potential to contribute significantly to the analysis and understanding of global environmental processes.

CORONA not only played a major role in answering key national security questions and revolutionized the way the US collects intelligence, but it also contributed to advances in the overall US space program. For example, Itek's camera technology evolved into imaging capabilities for the Apollo lunar mapping program and the Mars Viking Lander. Lockheed's exacting rocket steering that was necessary for precision imaging of ground targets evolved into a capability for accurate space maneuvering and docking. For a greater understanding of the CORONA program and its contributions to national security the reader is referred to [*McDonald, 1995b*].

A fundamental decision was made early in the CORONA program by Col. Frederic C.E. Oder, USAF Program Manager. Dr. Oder realized that the space environment might play an important role in the success of the program but he was concerned about the limited knowledge of this environment. He directed that piggy-back payloads be carried under the DISCOVERER program name to investigate the space environment and that the information be used to support the CORONA program [*Personal Communication*]. This was a monumental decision for the early space scientists eager to obtain flight opportunities on the new satellite platforms. Because of the frequent flights of the DISCOVERER satellites, the generous piggy-back payload capability, and the long 12 year duration of the program, numerous opportunities were created for individual scientists and institutions to make discoveries and to build their reputations in the new field of magnetospheric physics as well as in other scientific areas. Much of what was learned about the space environment was applied back into CORONA and contributed to the program's later outstanding performance record. This was especially true regarding the nature of the radiation belts and their effects on the sensitive imagery film. The purpose of this paper is to highlight some of the personal scientific measurements made aboard DISCOVERER satellites and how they contributed to our early knowledge of the magnetosphere.

1.1 *The First Soviet Space Imagery Program*

It is of historical interest to compare the Soviet space imagery program with the US CORONA program. Even before the launch of Sputnik on October 4, 1957 Russia's chief space designer, Sergei Korolev, recently revealed that he launched a satellite imagery reconnaissance program [*Oberg, 1995*]. The Soviets adapted the spherical Vostok capsules that carried the first cosmonauts into space into the ZENIT space reconnaissance system. The initial ZENIT design consisted of a cylindrical camera module and a conical film return capsule. Film was wound on cassettes in the small recovery capsule, while the camera system was not recovered. Similarities to the CORONA program are striking. The first launch attempt of the ZENIT was on December 11, 1961 but failed due to the booster's third stage not igniting. About the same time, the Soviets began to successfully launch small satellites called KOSMOS for the scientific exploration of space. The first successful ZENIT space reconnaissance satellite was launched on March 16, 1962 and because of secrecy was disguised as KOSMOS-4 [*Oberg, 1995*]. Usable photographs were returned after a four-day flight of KOSMOS-7 in July 1962, some 23 months after the successful imagery obtained with DISCOVERER XIV. Over the next eighteen months, the Soviets carried out ten flights. Initial resolution on the ground was 10 to 15 meters but improved such that the Soviets reported that they could "identify the type of automobiles in the Pentagon parking lot" [*Oberg, 1995*]. Korolev's design team eventually evolved into Energiya NPO and through privatization into Energiya Rocket and Space Company, which is now actively engaged with the US in the development of the International Space Station.

2.0 DISCOVERER RESEARCH PAYLOADS

The Lockheed Missiles and Space Company in Sunnyvale, California developed the AGENA spacecraft as the host vehicle for the DISCOVERER/ CORONA program. The AGENA was a three-axis stabilized vehicle, 1.5 m in diameter by approximately 7.6 m long, that could fly in the horizontal or vertical orientation to the earth. For CORONA missions the AGENA flew in the horizontal mode with the camera system mounted on the forward bulkhead. The recovery capsule containing the film was at the foremost end. An artist's concept of the KH-4B camera system in flight is shown in *McDonald* [*1995b*].

The aft bulkhead of the spacecraft housed a restartable engine that was used to adjust orbit parameters. Surrounding the engine were four panels each approximately 60 cm by 100 cm to which "piggy-back"scientific payloads and new technology components and systems under test could be mounted. The available weight on each flight was determined by the film load and the boost capability of the launcher. Typical "piggy-back" payloads were in the 50 kg range. Electrical power, telemetry, thermal control, and on many flights magnetic tape recorders, were provided to support the experiments.

Experiments involving biological samples, film pack, cosmic ray film stacks, film and solar cell degradation, as well as radiation shielding tests that needed to be recovered for analysis, were mounted inside the recovery capsule.

The Research Payloads program office at Lockheed Missiles and Space was responsible for integrating the "piggy-back" payloads with the AGENA spacecraft and the CORONA recovery capsules. The requirements were kept simple and payloads could be interfaced and tested in a matter of days to weeks. The author recalls installing and testing one payload that was launched within 48 hours. Many of the experimental payloads bore no security classification and most experimenters had no knowledge of the highly secret primary mission. With launches occurring every few weeks in the early days and with most carrying experiments on board, the DISCOVERER program was the forerunner of the "smaller, faster, cheaper" philosophy of conducting space science that is so in vogue today.

The Research Payloads office accommodated a wide variety of space experiments from a large number of government and industry laboratories as well as universities across the US. This is illustrated in Table 1 which shows an actual Research Payloads manifest from the early days of the DISCOVERER program.

The Air Force Cambridge Research Labs (AFCRL), the Lockheed Palo Alto Research Labs (LPARL), and the Los Alamos National Labs (LANL) got their start and/or early growth in scientific space research through these initial opportunities aboard DISCOVERER. Based on the knowledge and reputation gained these institutions went on to become major contributors to the new field of magnetospheric physics. In addition to these and the others institutions shown in Table 1, the Space Sciences Laboratory of the Aerospace Corporation under the leadership of Dr. George Paulikas conducted many of their early space physics experiments aboard DISCOVERER in support of the U.S. Air Force. Dr. Paulikas is currently the Executive Vice President of the Aerospace Corporation. As another significant example of the scientific opportunities

Table 1. DISCOVERER RESEARCH EXPERIMENTS

<u>TRACKING</u>	
Precision Doppler	Applied Physics Lab, Johns Hopkins Univ.
Precision Optical	Smithsonian Astrophysical Observatory
Secor	Army Map Service
<u>SPACE SYSTEMS DEVELOPMENT</u>	
Radiometers	Air Force Space Systems Division
AGENA Restart	Air Force Space Systems Division
Horizon Sensor	Air Force Space Systems Division
<u>RECOVERY</u>	
Biological	Air Force Aerospace Medicine
	Air Force Special Weapons Center
	Air Force Aeronautical Systems Division
	Brookhaven National Lab,
	Air Force Space Systems Division
Dosimetry	Air Force Aerospace Medicine,
	Air Force Cambridge Research Lab,
	Lockheed Palo Alto Research Lab,
	Naval Research Lab
Cosmic Ray Studies	Air Force Cambridge Research Lab,
	Air Force Aerospace Medicine,
	Lockheed Palo Alto Research Lab
Film Deterioration	Air Force Cambridge Research Lab
	Air Force Aeronautical Systems Division
Solar Cell Deterioration	Air Force Space Systems Division
Shielding Studies	Air Force Aerospace Medicine
	Air Force Special Weapons Center
<u>GEOPHYSICAL ENVIRONMENT</u>	
Cosmic Ray Monitors	Air Force Cambridge Research Lab
Atmospheric Density Gauges	
Atmospheric Impedance Probes	
Micrometeorite Detectors	
Erosion Detectors	
Radiometers	
Galactic Radio Frequencies	
Charged ParticleEnergy Analyzers	Defense Atomic Support
Auroral Electrons	Agency/Lockheed Palo Alto Research Lab
Proton Spectrometers	
Optical Luminosity Distribution	
Radio Frequency Propagation	Univ. of Illinois
X-Ray Detectors	Los Alamos National Laboratory
Neutron-Gamma Detectors	
Electron Spectrometers	
Proton Spectrometers	

presented by the DISCOVERER flights, Dr. Edward Stone, the current President of the Jet Propulsion Laboratory, conducted the first cosmic ray composition measurements aboard DISCOVERER XXXVI in partial completion of his doctoral degree from the California Institute of Technology.

3.0 THE AURORA EXPERIMENTS

The LPARL conducted a series of experiments aboard DISCOVERER satellites between 1962 and 1965 for the Defense Atomic Support Agency (DASA) and the ARPA through a contract with the Office of Naval Research. The objective of the experiments was to use the aurorae as natural disturbance phenomena in the D- and E- ionosphere and to correlate the "input" parameters (the particle types, intensity and energy spectra that caused the energy deposition and resulting ionization profiles) with the "output" parameters (radio propagation attenuation and optical luminosity). From such correlation the recombination characteristics of the ionosphere following a major man-made disturbance could be derived. DASA needed such information to understand the duration and extent of blackouts to critical military communication systems in the event of a nuclear explosion in the atmosphere.

The "Aurora" experiments, as they were called, were complex for this time period and involved simultaneous coordination of satellite, aircraft and ground observations. The key LPARL scientists involved in the design and execution of the program were Drs. Roland Meyerott, Richard Johnson, John Evans, Richard Sharp, Martin Walt, Raymond Smith and William Imhof. The author was a young member of the LPARL team and actively engaged in the instrumentation development. The satellite payloads consisted of large arrays of electron and proton detectors oriented along and at various angles to the zenith to observe the downward spiraling particles producing the aurora as well as a more limited set oriented to the nadir to observe the backscattered particles from the atmosphere. Photometers sensitive to several auroral wavelength bands were oriented in the nadir direction to correlate the "input" energy from the particles with the "output" luminosity of the auroral forms.

The particle detectors and the photometers were designed and developed at the LPARL. The primary auroral particle detectors were plastic scintillators coupled to ruggedized photomultipliers. The thin scintillators were coated with evaporated aluminum to render them opaque to moon light but sensitive to the keV electrons and protons that create the aurora. The thickness of the aluminum layer determined the lower energy threshold of this total energy detector. Several detectors with different threshold settings between 1 and 30 keV were used to obtain energy spectral data on the auroral electrons. By the addition of a ceramic magnet system in front of the scintillator, electrons could be swept away and low energy protons could be detected. Both the electron and proton detectors were placed at different orientations to the zenith and nadir to obtain angular distribution information.

The photomultipliers were operated in a unique, constant anode current mode. By maintaining the anode current constant the high voltage across the photomultiplier necessary to produced that current would decrease in an exponential manner as the scintillation light intensity increased. By sampling the high voltage as a 0-5 volt analog output it was possible to obtain over five decades of dynamic range. The large dynamic range was needed to measure the wide intensity variations observed in the different classes of aurorae. The temporal response of the detector system was in the few millisecond range, allowing the tracking of the rapidly changing auroral forms. The miniature high voltage supply that powered the photomultiplier was designed for this purpose and later patented [*Smith et al., 1962*]. It became the backbone high voltage power supply for future LPARL instrumentation for many years into the future. The photometers which were filtered to observe the 3.914 and 4.278 nm bands of ionized nitrogen were operated in a similar manner to the scintillation detectors.

From an historical perspective it should be mentioned that in the early 1960's there were a very limited number of detectors available for the measurement of low energy electrons and protons. Geiger-Muller tubes with thin windows, bare electron multipliers (the multiplying dynodes of a photomultiplier) and scintillation detectors were the only available systems. In addition to the mainstay scintillation detectors, electrostatic analyzer spectrometers were also designed and flown by LPARL on certain flights. Channel multipliers did not exist in the early 1960's and indeed their first use in space was in later Aurora flights aboard DISCOVERER.

A radio beacon operating at 20, 40, and 120 MHz was provided by Prof. George Swenson of the University of Illinois. The radio beacon would propagate from the satellite through the undisturbed and aurora-disturbed ionosphere and be detected by receivers on aircraft flying below the aurora and on the ground. Measurements of the rate of change between the signals at different frequencies, i.e. the dispersive doppler, would provide the total electron density integrated along the transmitted path. Signal strength measurements indicated the D- and E- region absorption due to the auroral particle ionization. The aircraft, operated by the AFCRL and flown from airfields in Alaska, also carried zenith viewing photometers to observe the auroral luminosity for correlation with the satellite measurements. Ground-based radio receivers and optical instruments were also located and operated by personnel of the University of Alaska at Fairbanks. Coordinated measurements in the nighttime ionosphere above the University of Alaska were planned and conducted using the predicted ephemeris of the satellite.

In the early years of the DISCOVERER program the satellite lifetime was typically 2 to 3 days. The short lifetime resulted from a combination of 1) the low perigee altitudes of typically 200-300 km in the northern latitudes, chosen to obtain maximum imagery resolution but resulting in considerable atmospheric drag that reduced lifetime 2) the limited film load and 3) the desire to return the capsule with the precious information as soon as possible. In later years the satellite lifetime was extended to 7-10 days as improved film and success rates increased [*McDonald,* 1995b].

The low altitude, polar-orbiting satellite basically placed the detectors in an ideal location crossing the magnetic field lines just above the auroral forms in the northern hemisphere. The short satellite lifetime provided about 12 opportunities for overpass coordinations in Alaska, far more than could be obtained in a typical rocket campaign. The satellite particle and luminosity measurements were most often obtained on every crossing of the northern and southern auroral zones using a combination of on-board tape recorders and frequent ground station contacts provided by the Air Force world-wide tracking network.

An all-electronic 64 channel analog multiplexer operating at 64 samples per second was designed especially for these experiments. The high sampling rate of the particle and photometer sensors provided a few millisecond temporal resolution in the auroral forms. Typically, this translated to obtaining auroral spectra every 63 meters but some detectors were sampled as often as every 13 meters. This is believed to be the first use of an electronic multiplexer in space since slow and noisy mechanical commutators were still the standard spacecraft multiplexer of the early 1960's. These high sampling rates on the auroral forms were not to be repeated by other experiments for many, many years to come.

Seven of these Aurora payloads were built and flown aboard DISCOVERER vehicles between 1962 and 1965. Table 2 shows a list of these experiments by their payload designator, the DISCOVERER and/or CORONA mission designator, the experiment payload complement and the performance of the mission. For more detail on these coordinated auroral experiments the reader is referred to the article by J. E. Evans [*Walt, 1964*] and for additional information on the instrumentation to *Reagan et al.* [*1964*].

The first flight on January 13, 1962 aboard DISCOVERER XXXVII was unsuccessful due to the AGENA failing to orbit. Within a month a second large payload was built and launched on February 27, 1962 aboard DISCOVERER XXXVIII. Such a rapid turn around of a complex payload was typical of this time period but nonetheless a tremendous tribute to the dedication and talents of the early space scientists, engineers and technicians. The main telemetry on DISCOVERER XXXVIII failed but 12 passes of limited data on the aurora were obtained over Alaska via a back-up telemetry link. The scientific data were analyzed and reported in the literature [*Sharp et al. 1963, 1964*]. This was

Table 2 AURORA EXPERIMENTS CONDUCTED BY LPARL ON DISCOVERER SATELLITES

Payload	Mission	Launch	Payload	Performance
1962a	XXXVII -9030	1-13-62	10 particle detectors 1 scanning photometer 1 radio beacon 1 multiplexer	AGENA failed to orbit
1962b	XXXVIII -9031	2-27-62	11 particle detectors 1 scanning photometer 1 radio beacon 1 multiplexer	Partial success, 12 passes over Alaska
1963a	9052	2-28-63	11 particle detectors 1 electrostatic analyzer 1 scanning photometer 1 nadir photometer 1 radio beacon 1 multiplexer	AGENA/ THOR failure
1963b	8002	5-18-63	8 particle detectors 1 radio beacon	Complete success
1963c	9059a	10-29-63	6 particle detectors 1 ion probe	Complete success
1964	1003	3-24-64	12 particle detectors 1 electrostatic analyzer 1 photometer 1 radio beacon 1 multiplexer	AGENA failed to achieve orbit
1965	1026	10-28-65	21 particle detectors 1 7-channel CEM 1 electrostatic analyzer 1 photometer 1 UV photometer 1 radio beacon 1 multiplexer	Complete success

the last flight using the unclassified DISCOVERER designation. All of the subsequent 92 flights were classified secret under the CORONA program name. Unclassified scientific payloads thereafter referred only to the international designation for the satellite launch, e.g. 1963-40.

The 1963a flight produced no data due to failure of the AGENA satellite to separate from the THOR booster. The 1963b flight consisted of a small payload without photometers since Alaska at launch time was in constant sunlight. This mission was a complete success and contributed much to our early understanding of the aurora phenomena [*Sharp et al. 1965*]. Another small payload, 1963c, containing five electron detectors, one proton detector and an ion trap to measure atmospheric composition was also flown successfully aboard CORONA mission 9059a in October 1963. The purpose of this payload was to obtain detailed low- energy electron spectra and angular distribution data in the northern and southern auroral regions. Particle data on 141 crossings of the auroral zones were obtained on this flight. The exciting new results prompted an International Symposium on the Aurora held at the LPARL in January 1964 [*Walt, 1964*].

A large payload was flown in 1964 that included an electrostatic analyzer but the AGENA vehicle failed to achieve orbit and no data were obtained. A photograph of this payload mounted on the aft rack of the AGENA as the spacecraft was being mated to the Thor launch vehicle at Vandenberg Air Force base is shown in Figure 1. Finally, the largest and most comprehensive Aurora payload flew in 1965 and was a complete success. Twenty one particle detectors and two photometers as well as the radio beacon successfully operated. Three days of world-wide data were collected. Included in the payload was a Channel Electron Multiplier spectrometer with seven channels, the first use of this new class of low energy detectors in space.

These early DISCOVERER flights led to a much better understanding of the aurorae phenomena. The distributions of the electrons and the protons around the auroral oval were mapped as a function of latitude, local time, and magnetic disturbance and it was discovered that 1) the peak electron and proton precipitations did not occur at the same latitudes 2) the proton energy deposition was significantly lower in latitude than the electron energy deposition 3) a region of "soft" electron precipitation distinct from the "hard" electron precipitation existed of the dayside auroral oval and 4) the electrons were more energetic on the night side of the auroral oval than on the dayside. These were all early clues to the complex magnetospheric phenomena that would be unraveled over subsequent years. The coordinated measurements between the "input" parameters and the "output" parameters led to a better measurement and understanding of the effective recombination rates in the upper D- and E- regions of the ionosphere.

4.0 THE STARFISH EVENT

On July 9, 1962 the U.S. detonated a 1.4 megaton nuclear device in the upper atmosphere over the Pacific ocean as part of the Cold war nuclear test readiness program. The detonation, code name Starfish, resulted in an intense artificial radiation belt that did not decay back to the natural levels until 1965. The fission-spectrum electrons trapped in this radiation belt were very intense and penetrated spacecraft skins causing considerable radiation damage to solar cells and other sensitive components. A more detailed scientific account of the Starfish event is included elsewhere in this monograph [*Walt, 1997*]. CORONA vehicle 9039 was launched on July 21, 1962, a few weeks after the Starfish event. Upon recovery and analysis it was discovered that some of the film in the supply canister was

My Adventures in the Magnetosphere

S. Fred Singer

Science & Environmental Policy Project (SEPP)
4084 University Drive, Suite 101, Fairfax, VA 22030-6812
e-mail: ssinger1@gmu.edu

1. SPACE RESEARCH AND COSMIC RAYS AT APL (1946-1950)

In August 1946, fresh out of the U.S. Navy as a mineman (gunner's mate) 2nd class, I started work as a junior physicist at the Applied Physics Laboratory of Johns Hopkins University (APL) in Silver Spring, Maryland. I had been recruited by James Van Allen to join the High Altitude Research Group, which was planning a series of experiments using captured German V-2 rockets.

Before volunteering for Navy service in 1944, I had already received a bachelor's degree in electrical engineering from Ohio State University at the tender age of 18 and had just taken my Ph.D. prelims in physics at Princeton. I knew next to nothing about the upper atmosphere or "space," and I had only a vague idea about cosmic rays. But I was willing to learn and the research sounded exciting, working in a newly established group full of enthusiasm under the triumvirate direction of Van Allen, Howard Tatel, and Robert Peterson--all of them basically nuclear physicists. My great companion and friend at APL was Jim Jenkins, a little older and more experienced than I. He and his wife Lu took me under their wing and told me about the wonders of California and the West, places I had never seen but would soon become familiar with. Tatel, a superb scientist, who had trained at the University of Michigan, became my mentor on nuclear instrumentation and physics generally; I was truly sorry when he left APL after a couple of years and later died quite young. I owe him a great deal.

My main assignment was to work with them on studying the intensity of the primary cosmic radiation (CR). My secondary job was to work with John Hopfield and Harold Clearman on the construction of an ultraviolet spectrometer to measure the altitude distribution of ozone. I was also put to work designing sun-based (and later also magnetic) systems to establish the attitude (angular coordinates) of the rocket so that we would know which way the CR "telescopes" (Geiger-counter coincidence-arrays) were pointing.

It was a very active job, involving long hours designing and testing circuits and preparing instrumentation that could withstand the accelerations of rocket launch, protecting high-voltage terminals from the near-vacuum conditions of space, and designing the experiment so that the data could be telemetered to ground-receiving stations. Generally, the rocket and all of the instrumentation would be destroyed on impact in the White Sands Proving Ground (New Mexico)--that is, if we could locate the debris at all. The actual recovery involved mainly armored film cassettes from rocket-borne cameras, whose information about the position of the horizon I also used to analyze the spin and attitude of the rocket.

The initial cosmic-ray experiments used Geiger counters that were heavily shielded with lead, in an effort to distinguish between the high-energy CR primaries (having energies of at least 7 BeV) and lower-energy secondaries. It was soon discovered, however, that the lead shielding created additional secondaries that could not be distinguished from primaries; it made the experiment infeasible. After about a year of experimentation and several rocket flights, Van Allen and Tatel dropped the idea of lead shielding or of using ion chambers, and went instead to "clean" CR-telescopes, consisting of a threefold-coincidence configuration, with additional "guard" counters to detect secondary events from

Discovery of the Magnetosphere
History of Geophysics Volume 7

CR showers. While this scheme eliminated most of the secondaries produced within the rocket and instrumentation, it could not eliminate "splash albedo," i.e., secondaries coming out of the atmosphere, energetic enough to penetrate the telescope and cause a coincidence. [Fig. 1]

By that time, I had become a full-fledged collaborator of Van Allen's and was trying to devise some means of distinguishing between primaries and secondaries. Time-of-flight would have been the obvious solution, but was beyond our technical means at the time. Nor would the method distinguish between a primary CR protons and splash "albedo" protons that had been turned around by the geomagnetic field [Fig. 1]--although I calculated what this contribution might be (and later used this information in developing the "neutron-albedo" theory of trapped radiation). I did have another experimental approach: placing a "low-efficiency" Geiger counter (filled with pure hydrogen) within the CR telescope, and thereby measure the specific ionization of the particle. (A primary CR proton with energy of several Bev, having minimum specific ionization, would generally not set off such a counter, while a secondary, e.g., a 100-MeV proton, would ionize heavily enough to set off the counter. Of course, primary CR alpha-particles, having 4 times minimum ionization, would also set off the counter; but their flux was known to be small.)

But our most basic approach was to measure the CR flux at different latitudes, pointing the telescope not only upward but also into other directions as the rocket spun and pitched--using the Earth's magnetic field as an energy spectrometer. (More specifically, the relevant parameter was the magnetic rigidity, defined as momentum divided by electric charge.) We therefore had to become experts on geomagnetic theory, and in particular, the formulation developed by the Norwegian geo-mathematician Carl Størmer. Fortunately, one of our team members was Ralph Alpher (of later fame for the Big-Bang theory of element formation, with George Gamow). He developed geomagnetic theory in enough detail to make it useful to experimenters like myself, who wanted to interpret not only the observed latitude variation, but also the so-called east-west asymmetry arising from the positive charge of CR primaries. (For example, one should observe, at the equator, a strong excess counting-rate for primaries from a telescope pointing towards the west, while albedo particles were thought to be isotropically distributed, thus diluting the expected asymmetry.)

In 1950, Van Allen and I finally published the results of our latitude survey of primary cosmic rays. In 1949 we had launched CR instrumentation in Aerobee sounding rockets from the USS Norton Sound (a small aircraft carrier) off the coast of Peru at the geomagnetic equator, and continued the work the following year with launches in the Gulf of Alaska. Even with some uncertainty about CR secondaries, we could

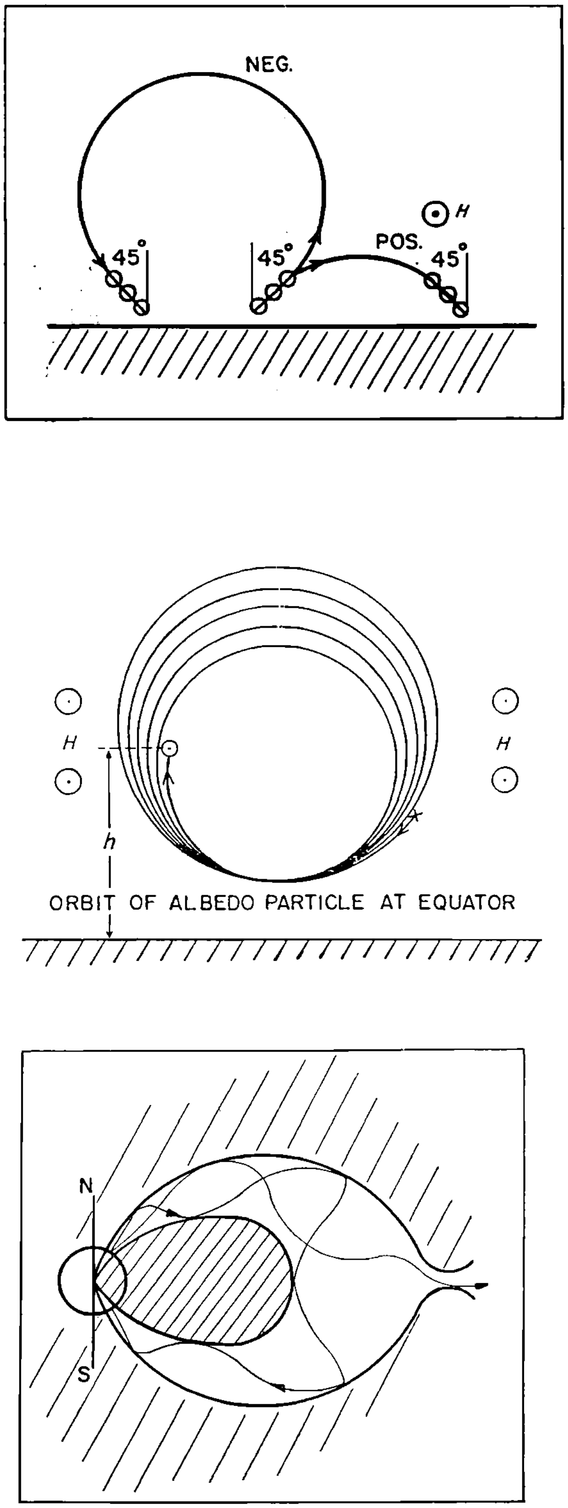


Figure 1. "Splash albedo"; motion of secondary protons originating from primary cosmic rays bombarding the atmosphere (Progress in Elementary Particle and Cosmic Ray Physics, Vol. 4, p.264).

calculate an energy spectrum for the predominant component of primary CR--high-energy (multi-BeV) protons.

2. FIRST ENCOUNTERS WITH GEOMAGNETICALLY TRAPPED PARTICLES (1950)

I had some misgivings about the results, however, because we never fully eliminated secondaries, as judged from the fact that they should dilute the expected west-to-east asymmetry of primary CR protons. It was essential to show that the splash albedo was isotropic. I wrestled with this problem for some time and published some suggestions that involved the idea of albedo protons rising from the atmosphere and spiraling in the Earth's magnetic field. (This work was published in an American Physical Society abstract and in a later review article; see Figure 1) But it was only after the discovery of the radiation belts that I calculated that geomagnetically trapped protons (from neutron albedo) were not isotropic but should show a west-east asymmetry--because of an intensity gradient with altitude.

This was not the only time that I missed out on trapped radiation. In the summer of 1950, as my stay at APL was drawing to a close, I joined a Navy Arctic resupply expedition to Thule, Greenland, aboard the icebreaker USS Edisto. Riding an icebreaker is a great experience; without a keel, they roll and pitch even in calm seas. With some difficulty, I launched balloons carrying lightweight cosmic-ray telescopes to measure primary radiation all the way up to the geomagnetic pole to investigate the CR energy spectrum down to a Bev or less. (In those days we believed that a solar magnetic dipole would produce a low-energy cutoff ("knee") in the energy spectrum.) Everything looked fine as we left Boston and Halifax and worked our way up the coast of Labrador and Greenland. But then the counting rates at top balloon altitudes, around 80,000 feet, started to go wild. Were these genuine high-energy particles, or was the high voltage feeding the thin-walled counters discharging? In further flights I applied extra wax protection to all high-voltage terminals and still the "problem" persisted. My co-worker, Russ Ostrander, a very capable APL engineer, decided to pack up after we finished our balloon launch series, a little disappointed by the absence of solid data.

We now know that I was seeing trapped particles leaking into the auroral zone. In 1954, at the University of Maryland, my graduate student Ray Rhodes and I finally figured out what was happening, and a little later, I began to realize that high-energy trapped particles must exist in the magnetosphere after all, whereas according to Størmer theory they could not have entered from infinity and should therefore not be present.

3. THE EQUATORIAL ELECTROJET (1948-1952)

Around 1948, Dr. E.H. Vestine of the Department of Terrestrial Magnetism of the Carnegie Institute of Washington suggested to APL that we try to measure the actual distribution of electric currents flowing in the ionosphere: their presence was known through the well-known magnetic variations observed at the Earth's surface; but these could not tell us about the location and vertical extent of the currents. Specifically, he suggested that we focus on the equatorial electrojet (EJ), a current-system generated by atmospheric tidal motions, with a highly concentrated maximum at local noon at the magnetic equator. The project was assigned to me and I began to try to understand the theory behind this effect by reading a book on the upper atmosphere by S. K. Mitra. The book became a sort of bible for me since it covered our knowledge--or rather lack of knowledge--of the ionosphere and exosphere (the region from which atmospheric atoms escape from the gravitational field into space). Parenthetically, much of my later research was devoted to correcting errors in the book relating to ionospheric conductivity, density distribution in the exosphere, ozone photochemistry, etc.

The task of measuring the electrojet seemed daunting, if not impossible. As one penetrates through the current layer, one should see a gradual weakening and then reversal of its magnetic field effect. But unless the layer were very limited in vertical extent, the available Aerobee rockets would not surmount the layer. Furthermore, the small magnetic perturbation would be horizontal and north-south, and therefore require an attitude-stabilized magnetometer for optimum detection sensitivity. But I discovered a saving possibility: since the main geomagnetic field at the equator is also horizontal and north-south oriented--and therefore parallel to the perturbation--a total-field magnetometer could do the job without requiring any stabilization. At any other latitude, however, the effect of the perturbation would be reduced and make it even less detectable.

With the help of Harry Vestine and Victor Vacquier I found such a total-field magnetometer at the Naval Ordnance Laboratory in White Oak, Maryland, and two scientists who were willing to work with me: Elwood Maple and William Bowen. Together we modified an existing, highly classified flux-gate magnetometer used in naval torpedoes and adapted it for use in a rocket. Lots of problems, in addition to security; I still remember having to construct a wooden internal structure and replacing the standard Aerobee aluminum cone so as to avoid induced current effects. Our first test came in 1948 at White Sands where we established that the magnetometer worked and that the Earth's field decreased with altitude in about the right manner. We saw no magnetic effects that could be attributed to an ionosphere current, but were not discouraged. The current is very small away from the equator; and when its horizontal magnetic field is added vectorially to the inclined Earth's field, it would produce an even smaller change.

In 1949, I was ready for the equatorial flights to be conducted from the deck of the USS Norton Sound, off the west coast of Peru. My main concern was whether we would be able to surmount an appreciable fraction of the current layer so as to establish its distribution. According to the book by Mitra, based on the work of Sydney Chapman, the current would extend from the bottom of the E-layer, at around 90 km, to the very top of the ionosphere, some 450 km--and even then there might not be enough conductivity to support all of the current. Yet the ground observations (by Alberto Giesecke at Huancayo, Peru) convinced me that the current must be there; and the rapid disappearance of the magnetic signal north and south of the equator seemed to indicate a current concentration at low altitude.

We fired two rockets, one at noon during current maximum, and one away from noon. A comparison of the two records should establish the existence of the current--and it did. What's more: the rocket seemed to have surmounted all of the current layer within about 12 km; unfortunately the rocket did not go any higher so that we could not see a change in slope of the field decline. We were jubilant and published our results in the Journal of Geophysical Research, but without explaining why the current was so highly concentrated in the vertical dimension. My jubilation didn't last very long. Sydney Chapman visited APL soon after, saw the results, and shook his head sadly. He obviously did not believe that our observations were valid, and I felt quite discouraged.

I found the answer some two years later after reading carefully the book on "Cosmical Electrodynamics" by Hannes Alfven. The reason for the assumed low west-east electrical conductivity of the ionosphere is the interfering effect (Hall effect) of the north-south magnetic field. This was the conductivity used by Chapman, Mitra--and everyone else. But if there should be an electric polarization set up in the vertical direction, then the Hall effect is canceled and the full conductivity restored as if there were no magnetic field. I published these conclusions explaining our observations in Nature in 1952. The very same issue carried a more elaborate theoretical paper by David F. Martyn giving exactly the same result.

4. THE MOUSE AND TRAPPED-PARTICLE RING CURRENTS (1952-1957)

In 1950, I was faced with decisions: Jim Van Allen had decided to leave APL and accept a position at the University of Iowa. He wanted me to join him there and I did visit. I also was interviewed by Edward Teller at Los Alamos who wanted me to work on instrumentation that would survive for even less time than it did in rocket experiments. Finally, APL wanted me to continue as a co-director of the High-Altitude Research Program. I was very flattered, but eventually decided to go to London instead.

My tenure as Scientific Liaison Officer for the Office of Naval Research (ONR) in Europe (1950-1953) gave me a chance to think more deeply about the rocket experiments of 1946 to 1950. That period had been too hectic, with too many deadlines, too many night sessions, to leave time for fundamental thinking. Visiting European laboratories gave me a chance to learn. At Prof. Powell's lab in Bristol I learned about cosmic-ray interactions with nuclei in photographic emulsions. In Durham, Prof. Paneth showed me work on uranium-helium dating of iron meteorites that gave impossibly large ages. Putting two and two together, I developed a theory of CR interactions with meteorites that would create helium in nuclear fragmentations--then a new idea. I figured that about 30% of the helium would be in the form of the rare He-3 isotope--and so it turned out. I should have become a chemist right then and there; instead, I landed in space--in a different way.

Partly stimulated by lectures I gave to the British Interplanetary Society in London in 1951, I developed ideas for an instrumented Earth satellite to carry on the kinds of measurements we had been doing in rockets. (Years later, Professor Joseph Kaplan who was head of the U.S. Committee for the 1957 International Geophysical Year (IGY), would refer to satellites as LPR--Long Playing Rockets). It was quite a radical idea at the time, which offended those who poo-poohed any notion about working in space as well as those who had already set their aim on manned exploration of the solar system. What I brought to the discussion, mainly, was the notion that instrumentation could be miniaturized and that useful research could be done with a satellite weighing only a few kilograms--even if it survived only for days or weeks. With rocket propulsion power limited, I figured that "lifetime" would be limited by aerodynamic drag; I soon published calculations of lifetime under various conditions. So was born the MOUSE--the Minimum Orbiting Unmanned Satellite of the Earth--with the help of futurist Arthur C. Clarke and rocket engineer Val Cleaver and some alcoholic conviviality at the Players' Club near Trafalgar Square. For the next few years, I would try to think of all kinds of experiments that could be done by such a satellite: meteorological observations, including worldwide measurements of ozone; ultraviolet measurements of the Sun and other stars; measurements of incoming interplanetary dust as well as the zodiacal light/solar dust corona; magnetic measurements of ionospheric currents; the use of the satellite lifetime to measure the density of the upper atmosphere; primary cosmic rays, and finally, geomagnetically trapped particles. All these ideas were duly worked out and published in some detail.

The idea of a particle population trapped in the geomagnetic field developed slowly, but was certainly helped along by a three-months' stay in Stockholm in the Spring of 1953. As a guest worker in the institute of Hannes Alfven at the Royal Institute of Technology (KTH), I was put in touch with many interesting investigations. For example, I thoroughly absorbed Alfven's criticism of the magnetic-storm theory of Chapman-Ferraro. They had visualized an equatorial ring current at a distance of several Earth radii, in which protons and electrons moved along the equator in the same direction but with different velocities; the current thus generated was supposed to account for the observed magnetic field decrease during the main phase of a magnetic storm. Their model was highly artificial--and also physically unstable. On the other hand, Alfven's theory of magnetic storms, based on an electrically polarized particle beam coming from the Sun, seemed equally improbable.

All theories at the time were dominated by the geomagnetic theory of Størmer, where a fixed magnetic dipole of the Earth would prevent low-energy particles from entering the field so that the region out to many Earth radii was termed "forbidden." I noticed, however, that the original terrella experiments by Birkeland, in which he shot electron beams at a magnetized sphere in a vacuum chamber, showed luminosity coming from these forbidden regions. I concluded that the imperfect vacuum allowed electrons to be scattered, changing their (Størmer) invariant integral of motion (the vertical component of angular momentum). This observation stimulated me to think about the consequences of what would happen if particles were really to exist in the magnetosphere--even though we had no clear idea of how they would get to these inaccessible regions in the first place. But once there, they would not only spiral about lines of force from pole to pole, but would also drift in longitude--electrons one way, and protons the other--thus creating a completely stable ring current. I completed my calculations in 1955, but was disappointed when the Journal of Geophysical Research returned my paper with unfavorable comments from referees. (Something about being "too fantastic.") Fortunately, just then, Helmut Landsberg asked me to prepare a review article on "geophysical research with Earth satellites." I had already published several chapters in a book edited by Van Allen based on a conference at the University of Michigan. My "review" for the Landsberg volume, which appeared in 1956, contained a more complete discussion of research possibilities, including the results of the trapped-particle/ring-current theory. These particles were supposed to have entered into the geomagnetic field by perturbing it and being scattered in at the same time; they occupied the region of what was later referred to as the "outer" radiation belt. My research paper giving the details of the theory was finally published in 1957 in the Transactions of the American Geophysical Union. [Fig. 2]

5. TRAPPED PARTICLES AND FARSIDE AT CASP (1953-1958)

I had joined the physics faculty of the University of Maryland in 1953, and, with the enthusiastic help from department chairman John Toll, gradually built up a research operation called the Center of Atmospheric and Space Physics (CASP). It involved some dozen research associates and graduate students, and cooperating faculty members Ernst Öpik, Howard Laster, and William MacDonald. Ray Rhodes, starting in 1954, worked on trapped auroral particles, and together with Ken-Ichi Maeda, studied their energy loss and penetration into the upper atmosphere. Hans Griem, who joined me as a post-doc from the University of Kiel in 1954, studied the lifetime (looping factor) of trapped particles in the exosphere [Fig. 1]. Our aim then was to put some limits on the importance of albedo when measuring primary cosmic radiation. Robert Wentworth became my most important research assistant, working on everything from impact points of ballistic rockets on a rotating Earth to ozone observations from satellites and, of course, trapped particles. He and I started a thorough analysis of the trapped-particle ring current, later continued by John Apel in his master's thesis. It was also a time for intensive experiments in cosmic rays (David Stern, George Homa, Martin Swetnick, John Corrigan), and development of small research rockets for a variety of experiments.

Our first rocket was the Terrapin, a simple 2-stage solid-propellant vehicle, designed by us (Yuri Kork, Dick Bettinger) and built by Republic Aviation. We launched from Wallops Island and off Puerto Rico, from aboard a Navy vessel so small that it could barely be called a ship. We flew a sodium vapor experiment to observe ionospheric winds (Einar Hinnov); what with thermite bombs to evaporate the solid sodium, it's a wonder we didn't kill ourselves or sink the vessel. It's a good thing the ship's captain never knew what we were doing--except when we disposed of the sodium in a spectacular fashion by tossing it overboard.

From the Terrapin we went to an even simpler and cheaper vehicle: our Oriole design was a 1½-stage vehicle, a 1-inch diameter "pencil" of extremely low drag, without fins but with a tungsten tip (for stability), boosted by a standard 3-inch LOKI rocket. It worked, but we had a hard time cramming

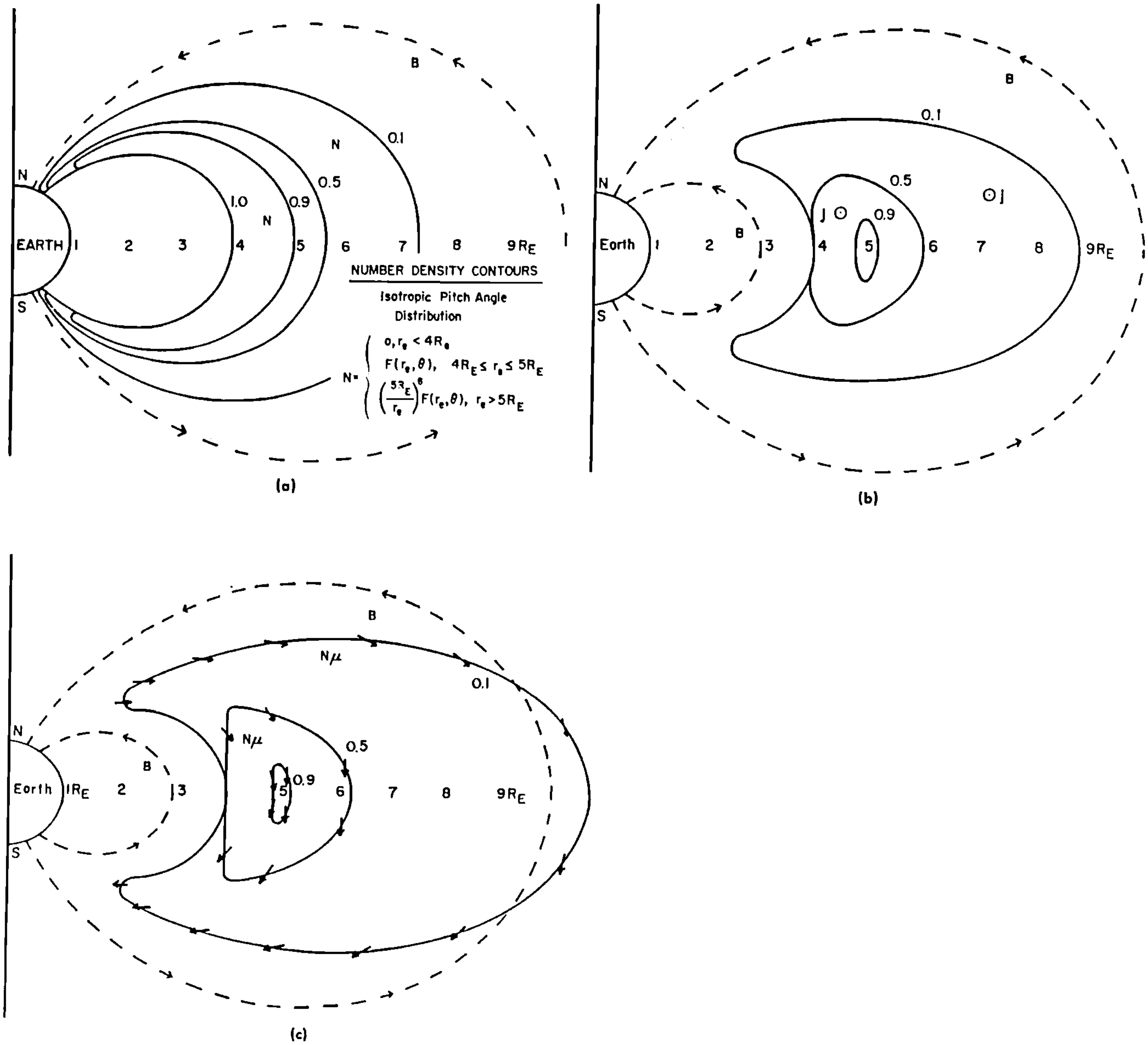


Figure 2. (a) Contours of equal concentration of trapped particles. (b) Contours of equal drift current density. (c) Contours of equal magnetic moment density (Apel et al, 1962).

instrumentation and telemetry into the pencil; it would be much easier now with solid- state miniaturization.

We also experimented with an aircraft-launched rocket--to overcome the drag of the lower atmosphere. We instrumented a 2.75" FFAR rocket and persuaded Patuxent Naval Air Station to try a loop maneuver in which the rocket was fired when the plane was moving straight up. In spite of technical success, funding for all these small-rocket programs dried up as NASA built its own research centers. After 1957, I soon switched to theoretical work.

Beginning in 1955, much of my activity revolved around schemes for observing the geomagnetically trapped particles I had been publishing on. I had met Colonel William Davis of the Air Force Office of Scientific Research and Morton Alperin, head of the AFOSR Pasadena office and a prize-student of the legendary Theodore von Karman (who kept referring to me alternately as "the Mouse-Man" or as "Herr Kollege," a title of which I was rather proud). I proposed to them building a four-stage vehicle, using well-known solid rockets (Loki and Recruit), to be launched from a balloon platform. I calculated that the payload would reach an altitude of about one Earth radius (6370 km), enabling us to see not only an expected increase in cosmic radiation (calculated by Bob Wentworth) but also look for the existence of trapped particles (as stated in my proposal). CASP at the University of Maryland received the funding to build the instrumentation, consisting of a simple geiger counter. Ford-Aeronutronics received the contract to build the vehicle, which the Air Force named FARSIDE. (The general in charge of AFOSR had become convinced by super-salesmen Davis/Alperin that we would pass the farside of the Moon; I couldn't talk him out of it.) In 1957, and in a later 4-part series, in the magazine "Missiles and Rockets," I published a sequence of articles discussing the prospective results, together with an analysis of how atomic explosions in the upper atmosphere could generate artificially trapped particles.

After the Soviets launched Sputnik in October 1957, the Air Force became serious about FARSIDE and rushed into launchings from a Pacific atoll. I never had any written confirmation, but was told that the balloons failed to perform. In any case, I never received data from my experiment. Since I had already been turned down for satellite experiments in the IGY, I had to give up the idea of looking for trapped radiation.

6. THE PRE-SATELLITE YEARS (1953-1957)

In 1953, after my move from ONR London to the University of Maryland, I became much involved in propagating the cause of small instrumented Earth Satellites--a novel idea at that time. On the one hand I was a popularizer, speaking at the Hayden Planetarium in New York a few months later; it involved such notables as Arthur Clarke and Willy Ley, but also Harry Wexler, chief scientist of the US Weather Bureau--who only a few years later, in 1962, would recruit me to take over as the first director of the US Weather Satellite Service.

On the other hand, I had started to publish serious scientific papers pointing out the kind of geophysical research that could be done with very simple instrumentation by remote sensing or in situ observations. Inevitably, this brought me into contact with government agencies that had the capability of launching such payloads. My friend and American Rocket Society colleague, Frederick C. Durant, introduced me to Navy Commander George Hoover, who was actively promoting satellites in the Office of Naval Research (and everywhere else) and in 1954 arranged a meeting in his Washington office with Wernher von Braun. Project Orbiter was born at that meeting, and from then on I was in close touch with the Army Redstone Arsenal in Huntsville, Alabama, where the von Braun rocket group developed a satellite launch system based on the Redstone ballistic missile. (I remember visiting von Braun, General John Medaris, and Ernst Stuhlinger there around 1956-57.)

At the same time, this brought me into conflict with the Navy effort to develop a satellite launcher, the Vanguard project, based on an extrapolation of the Viking research rocket. The Viking, in turn, was developed to be a replacement for the V-2 and became a rather expensive rival to our APL-developed Aerobee. President Eisenhower decided that an American satellite, to be launched during the International Geophysical Year, 1957-1958, should not be based on a military rocket. And so the Army project was put aside. Ironically, once Sputnik went up in October 1957 and after the Vanguard launching failed, the Army project was permitted to launch Explorer-1, the first U.S. satellite.

The groundwork for an IGY satellite was laid by two international resolutions, both drafted by me. The first one was adopted at the URSI (International Scientific Radio Union) meeting in The Hague in the summer of 1954; the second one a few weeks later at the IUGG (International Union of Geodesy & Geophysics) meeting in Rome. In both cases it was Lloyd Berkner and Athelstan Spilhaus who, by sheer force of their personalities, pushed through these rather radical resolutions. When a representative of Project Vanguard objected, trying to slow down a rush towards satellites that might have favored the far-advanced Army project, and complained that batteries might not work in a vacuum, Spilhaus slammed the table: "Dammit! We'll get batteries that don't bubble!" And so satellites--Long-Playing Rockets--became part of the IGY program. Nobody, of course, suspected that the Soviets would get there first. I got an award from President Eisenhower, but missed out on getting an experiment on satellites during the International Geophysical Year.

7. RADIATION BELT THEORY (1958-1962)

The situation changed in a quite unexpected manner, when Explorer-1, launched in 1958, reported the existence of trapped radiation at low altitudes. Actually, the radiation had first been observed by Vernov and Chudakov in Sputnik-2, but since they did not receive the data from apogee (over Australia), they did not see the rapid rise in intensity with altitude until much later. The rapid rise in counting rate with altitude of the Explorer-1 Geiger-counter indicated to me that the particle flux rate was controlled by the atmospheric density and, therefore, by lifetime in the atmosphere. In turn, this suggested a very low injection rate that had nothing to do with solar eruptions or magnetic storms. When Van Allen's results appeared in the newspapers, I speculated that we must be dealing here with cosmic ray "albedo" whose lifetime with altitude we had already studied [see Fig. 1]. (The albedo particles would stem from the disintegration of a nucleus of atmospheric oxygen or nitrogen hit by a primary high-energy proton.) But how could an "albedo" proton travel from the atmosphere (from, say, 30 km) to an altitude of several thousand km at the equator in the presence of a horizontal magnetic field? It suddenly occurred to me that the injection into the magnetic field must have occurred from albedo neutrons emanating from the atmosphere and decaying within the geomagnetic field. From my earlier studies, I already knew that the energy spectrum of these neutrons would extend into the range of several hundred MeV, and that their decay would produce protons of that energy--highly penetrating and of long lifetime. I quickly published my neutron albedo theory in Physical Review Letters, having calculated the energy spectrum of the trapped protons and their detailed spatial distribution up to about two Earth radii. I also predicted that trapping would break down at that altitude as the adiabatic invariance broke down and that a second radiation belt would occur at higher altitudes [Fig. 3].

From then on I focused my attention entirely on the trapped protons in the lower magnetosphere. Independently, Paul Kellogg had published on the albedo theory, focusing his attention on electrons. But because of their short range and short lifetime, their role would not be very important. Much later I learned that Vernov and Chudakov in Moscow had independently thought of neutron albedo; as far as I know, they never developed the theory in as much detail as we did.

We now had a big theoretical effort going at the University of Maryland. Bill MacDonald joined Wentworth and me in fully developing the electron component of the neutron-albedo theory. Because of their scattering, the full Fokker-Planck formalism had to be applied. Graduate student Alan Lenchek joined me in detailing the proton component. We expanded this work to calculate the to-be-expected west-east asymmetry and were gratified when Harry Heckman and George Nakano (UCAL, Berkeley) confirmed the theory with their observations. We also expanded the albedo theory to cover neutron albedo generated from solar-flare-produced cosmic rays, which introduced an interesting time-dependent feature. So while I missed on actual experiments of trapped radiation, we were able to cover the theoretical explanation fairly completely.

8. OTHER MAGNETOSPHERIC PROBLEMS (1953-1962)

There were three other problem areas that occupied me during my years at the University of Maryland (1953--1961) and my year as a guest scientist at the Jet Propulsion Lab of Cal Tech (1961-1962). One was the problem of magnetospheric plasma composition, the others were magnetospheric effects on cosmic rays and on the motion of electrically charged bodies.

8.1. *Exosphere and Magnetospheric Plasma*

Beginning in 1959, Öpik and I published a complete theory of the Earth's exosphere, for the first time considering its non-equilibrium, non-Maxwellian character, manifesting itself in

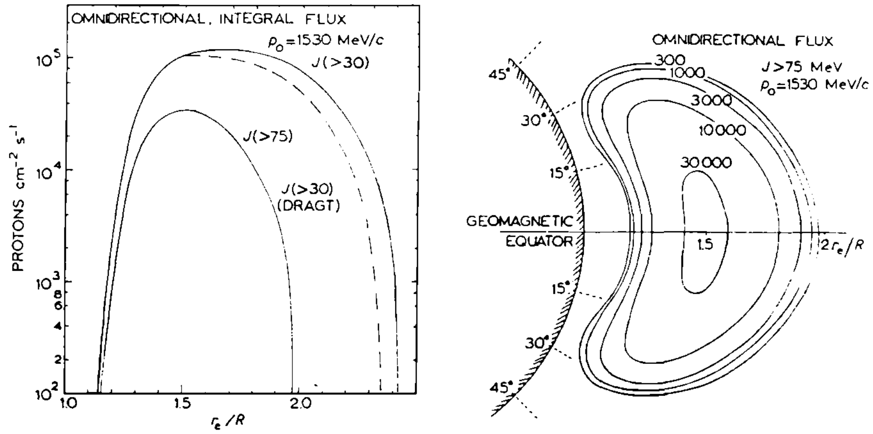


Figure 3. Geomagnetically trapped protons (forming a radiation belt) derived from albedo neutrons. Calculated omnidirectional fluxes are shown; the directional intensities and energy spectra, calculated from theory and basic cosmic-ray data, without any further adjustments, compared well with subsequent detailed observations (Progress in Elementary Particle and Cosmic Ray Physics, Vol. 6, p.314).

the absence of bound orbits of exospheric atoms. (As a result, unlike the existing theory described in Mitra's book which gave a finite density at infinity, our theory quite properly gave a zero density). Independently, Francis Johnson and Joe Chamberlain published similar theories, but from different points of view. We all agreed, however, that the solar wind could be thought as the evaporation of the hot solar corona in the Sun's gravitational field. Graduate student Mort Liwshitz studied the difficult problem of whether, in view of the escape of fast particles, a full Maxwellian distribution existed at the base of the Earth's exosphere, as generally assumed. Aharon Eviatar studied the density distribution of ions in the combined Earth's gravitational and magnetic field and discovered many interesting new facts in this theoretical analysis which was stimulated by discussions with Bob Helliwell and Don Carpenter, who were studying "whistlers." Paul Nakada and I published on the existence of a vertical electric field in the ionosphere due to the higher velocity of electrons (and their tendency to escape); one consequence of such a field is an increased scale-height for positive ions, and even a negative scale-height for multiply charged ions of helium and oxygen.

Richard Bettinger constructed Langmuir probes for rocket experiments involving the emission of electrons from hot cathodes, in order to measure the electric potential, Debye length, and other plasma parameters of rockets in space. Einar Hinnov developed a detailed proposal, complete with calculations, for emitting barium clouds from rockets and observing their motion from the ground, with the aim of studying magnetospheric motions. Unfortunately, neither Bettinger nor Hinnov's experiments ever received funding. Graduate student Evan Harris Walker was chasing after "ghosts," calculating whether the motion of a charged spacecraft could generate plasma condensations in the geomagnetic field that might explain the strange radar echoes observed by John Krauss at Ohio State University.

8.2. *Forbush Decreases: Magnetosphere or Interplanetary?*

It had been known for many years, largely through the observations of Scott Forbush, that the intensity of cosmic rays at the Earth's surface decreased during the main phase of magnetic storms (when the observed geomagnetic field also decreased). It had been commonly assumed that the Forbush decrease was a geomagnetic effect and could be explained somehow in terms of a modification of the Størmer theory. This proved not to be the case.

Quite by accident, I had become a participant in the debate. It happened in the summer of 1947 when as part of my Ph.D. thesis I made observations on extensive cosmic-ray showers (Auger showers) and decided to take advantage of a Navy Arctic resupply expedition to also investigate (as a subsidiary objective) the possible existence of a latitude effect. (There was no latitude effect observed, but I did obtain good data on the density distribution of showers and thereby on the energy spectrum of cosmic-ray primaries in the range of 10^{14} to 10^{16} eV.)

Forbush, whom I knew in Washington, asked me to install in Thule, Greenland, essentially at the geomagnetic pole, one of his lead-shielded ion chambers, a rather old-fashioned instrument for measuring cosmic rays. I was happy to oblige and later analyzed the data for him. I noticed that a typical Forbush decrease occurred during a magnetic storm in 1951. Upon reflection, I realized that there was no way an equatorial ring current could affect primary cosmic rays, typically of energy 10 BeV, at the pole. Tracing back their path, they would have to be arriving from beyond the magnetosphere.

There now developed an interesting debate about the cause of Forbush decreases, in which Phil Morrison and Eugene Parker played leading roles. They favored mechanisms that would delay the arrival of cosmic rays at the Earth; they visualized a turbulent magnetic field that would increase the path length. But it became clear to me that according to Liouville's theorem (from statistical mechanics) the directional intensity (and flux) of cosmic rays reaching the Earth would not be affected. Using phase-space arguments, the decrease had to be caused by an energy change along the cosmic ray trajectory, due to a retarding electric field or some other means. Using a thermodynamic argument, I finally settled on a mechanism whereby cosmic rays caught in an expanding turbulent magnetic field would lose energy to the field and thereby be reduced in flux. I remember proposing this mechanism at an American Physical Society meeting and later, at a cosmic-ray conference on Lake Como, Italy (Nuovo Cimento). But the theory was only taken seriously after I had renamed it as an "inverse Fermi effect." Interesting! Howard Laster and Alan Lenchek joined me in a more detailed publication.

8.3. *Coulomb Drag: Needles, Interplanetary Dust Particles, and Orbital Debris*

The story of "charged drag" also has a long history. Around 1954, in preparing ideas for different satellite experiments, it occurred to me that interplanetary dust particles (zodiacal light particles) would become electrically charged while moving through interplanetary space and through the outer ionosphere (as the magnetosphere was then known); would their interaction with the geomagnetic field lead to a latitude dependence of the incident flux, similar to that for cosmic rays? Lyman Spitzer and Fred Whipple had already published on the negative charge of particles in a plasma; I added the idea that the photoelectric effect would produce a positive charge for particles in interplanetary space and that there might be a narrow region in the outer ionosphere where the two effects could balance, producing a near-zero charge. (Years later, using this phenomenon I developed a trapping mechanism for particles ejected from lunar impacts.)

All this speculative material was duly published, first in the Van Allen volume, together with what I called the "coulomb drag," the retarding force experienced by a charged body moving through a plasma (I didn't know at the time that this had already been discussed in the literature and tried laboriously to reinvent an expression for the drag force).

Much to my disappointment, both geomagnetic effects and coulomb drag did not seem to be very important in practice. For satellites, the ordinary aerodynamic drag was far more important because the dimension of the satellite was large in relation to the Debye length. I did invent a special satellite made out of chicken wire, which had essentially no aerodynamic drag and only charged drag, in order to demonstrate experimentally the importance of coulomb drag. Even for particles in the micron range, it could be important only in special cases: For example, Harris Walker and I published on the migration of charged dust on the lunar surface and found that there would be no deep accumulation--no dust oceans, an important issue before the first manned landings. In a 1969 review article, Lothar Bandermann and I discussed the problem extensively for zodiacal particles in interplanetary space.

But for particles in the magnetosphere, the effect would only be important if the particle were in a bound orbit rather than a hyperbolic trajectory. Three opportunities arose to apply the theory.

(i) Around 1960, there was much excitement and a flurry of papers when it seemed that high levels of observed dust impact rates pointed to a "dust belt" around the Earth. I was able to show, using the extended Liouville theorem, that the observed impact rate for a detector should increase several hundred-fold near the Earth, just because of focusing and acceleration by the gravitational field--without any bound orbits. There should be additional interesting features: a maximum impact rate well above the Earth's surface--because of the Earth's shadow effect; a morning-to-evening asymmetry and seasonal variation--because of the Earth's orbit eccentricity.

(ii) In 1961, the Air Force launched Project Westford. It was designed to release into the upper atmosphere millions of copper needles, cut to the length of a dipole, to form an artificial ionosphere for communication purposes. To meet objections that this project would interfere with radio astronomy, Irwin Shapiro, at Lincoln Lab, tried to demonstrate that a radiation-pressure resonance would limit the lifetime of these particles so that the needle belt would only be temporary. I noticed that he had neglected to consider coulomb drag and showed in a paper in Nature that charged drag would interfere with the resonance required to make the radiation-pressure perturbation effective. We engaged in a hot debate over these points, which was never quite settled. The first Westford project launch failed because the experimenters had forgotten about some basic points of classical mechanics so that the needles were never released. (Raymond Lyttleton and I pointed out to them that the stable rotation of the payload, a nonrigid body, would be about the axis of <u>maximum</u> moment of inertia, not about that of minimum as designed.) The second launch attempt did not settle the controversy either.

(iii) A quarter-century later, in 1986, my one and only space experiment went into orbit on the LDEF (Long Duration Exposure Facility) satellite, launched by space shuttle. With my University of Virginia graduate student John Stanley (now deceased) we wanted to measure the flux of interplanetary dust particles with an array of solid-state detectors--capacitors that discharged when impacted. LDEF was supposed to be recovered a year later, but the Challenger accident changed all that. LDEF was finally brought back six years later, just days before it would have re-entered the atmosphere and burned up. Unfortunately, we used a tape recorder that only had a capacity of one year; there was no telemetry, and so we missed seeing Halley's comet. We did, however, observe quite unexpected results: clouds of micron-sized orbiting particles, almost certainly manmade debris. When we wrote up the results in 1990-93, coulomb drag finally came back in its own as determining the lifetime of the debris particles.

9. POSTSCRIPT: FROM MAGNETOSPHERE INTO THE ATMOSPHERE (1962-1996)

With the exception of the LDEF analysis of orbital debris in 1992, I left active magnetospheric research in 1962 when I was asked by the Department of Commerce to head the activities that started the U.S. Weather Satellite Service--certainly my most worthwhile government assignment. My adventures in promoting the simple spin-stabilized TIROS satellite, a re-incarnation of the MOUSE, over the elaborate and costly Nimbus satellite are recorded in a thesis and book on public administration by political scientist Richard Chapman of Syracuse University, NY.

From 1964 to 1967, as Dean of the new School of Environmental and Planetary Sciences at the University of

on the value of nuclear energy–I learned a great deal from his book "Cosmical Electrodynamics" and his general approach to problems that proved to be useful later.

Ernst Öpik became somewhat of a role model for me. He taught me how important it was to work things out in great numerical detail. I invited him to the University of Maryland in 1956, sight unseen, based on our common interests in the origin of meteorites. I am proud to say that I may be the only one with whom he has ever co-authored scientific articles. We started with the dynamics of meteorite production in collisions and my technique of "cosmic-ray ages" of meteorites; we continued with the theory of the exosphere and ionospheric electric fields, and investigated the exosphere and electrosphere of the Moon, leading to the transport of lunar dust. It was Öpik who got me interested in planetary physics and in the origin of Phobos and Deimos, the Martian satellites.

Among my contemporaries, I certainly learned a great deal from the papers and publications of the many scientists mentioned in this review. But in addition, I was much stimulated by the writings and ideas of Tommy Gold and Bob Helliwell.

Bob Wentworth is the only one of my research students who continued with magnetosphere problems, specializing first on micropulsations and later on important topological questions of the geomagnetic tail. He joined the Lockheed Research Labs in Palo Alto to work with Alex Dessler, Francis Johnson, Bill Hanson, and Martin Walt–certainly the most outstanding group in this field; I have learned much from their writings.

REFERENCES AND NOTES

1. Reference Books

Akasofu, S.I. and S. Chapman, "Solar-Terrestrial Physics," Oxford University Press, 1972.

Alfven, H., "Cosmical Electrodynamics," Oxford University Press, New York, 1955.

Chapman, S. and J. Bartels, "Geomagnetism," Clarendon Press, Oxford, 1940.

Singer, S.F. ed., "Interactions of Space Vehicles with an Ionized Atmosphere," Pergamon Press, Oxford, 1965.

Størmer, C., "The Polar Aurora," Oxford University Press, New York, 1955.

Van Allen, J.A., ed. "Scientific Uses of Earth Satellites," University of Michigan Press, Ann Arbor, 1956.

2. Primary Cosmic Radiation & Time Variations

Corrigan, J.J., S.F. Singer and M.J. Swetnick. "Cosmic Ray Increases Produced by Small Solar Flares." *Phys. Rev. Letters*, **1**, 104—105, 1958.

Laster, H., A.M. Lenchek and S.F. Singer. "Interplanetary Gas Cloud Modulation of Cosmic Rays." *Bull. Am. Phys. Soc.*, **5**, 259, 1960.

Lenchek, A.M., H. Laster, and S.F. Singer, *J. Phys. Soc. Japan*, **17**, Suppl. AII, 583, 1962.

Maeda, K., V.L. Patel and S.F. Singer. "Solar Flare Cosmic Ray Event of May 4, 1960." *J. Geophys. Res.*, **66**, 1569—72, 1961.

Singer, S.F., *Phys. Rev.*, **76**, 701, 1949.

Singer, S.F., *Phys. Rev.*, 77, 729, 1950.

Singer, S.F., *Phys. Rev.*, **80**, 47, 1950.

Singer, S.F., *Nature* **170**, 63, 1952.

Singer, S.F., Duke University Symposium on Cosmic Rays, Durham, North Carolina, 1953.

Singer, S.F., *Phys. Rev.*, **95**, 647 (A), 1954.

Singer, S.F., *Phys. Rev.*, **98**, 1163 (A), 1955.

Singer, S.F., *Phys. Rev.*, **98**, 1547 (A), 1955.

Singer, S.F. IUPAP Cosmic Ray Conference, Varenna, *Nuovo Cimento* **5**, Suppl. II, 1957.

Singer, S.F. "Observations of Cosmic Ray Decreases at the Pole." *Nuovo Cim.*, **8**, 326—333, 1958.

Singer, S.F. "On the Low Energy Cutoff of the Cosmic Radiation." *Nuovo Cim.*, **8**, 342—348, 1958.

Singer, S.F. "The Primary Cosmic Radiation and its Time Variations." *Progress in Elementary Particle and Cosmic Ray Physics*, (ed. J. G. Wilson and S. A. Wouthuysen), Vol. **4**, pp. 203-335, No.-Holland Publ. Co., Amsterdam, 1958.

Singer, S.F. "Cosmic Ray Time Variations Produced by Deceleration in Interplanetary Space." *Nuovo Cimento* **8**, Suppl. II, 334–341, 1958.

Singer, S.F. and R.C. Wentworth. "Cosmic Ray Measurements in the Vicinity of Planets and Some Applications: Part I. Primary Cosmic Radiation." *J. Geophys. Res.*, **64**, 1807—1813, 1959.

Van Allen, J.A. and S.F. Singer, *Phys. Rev.* **78**, 819; **80**, 116, 1950.

Van Allen, J.A. and S.F. Singer, *Nature* **170**, 62, 1952.

3. Meteorites and Cosmic Rays

Öpik, E.J. and S.F. Singer. "Reinterpretation of the Uranium-Helium Ages of Iron Meteorites." *Trans. Am. Geophys. Union*, **38**, 566—568, 1957.

Singer, S.F. "Meteorites and Cosmic Rays." *Nature*, **170**, 728—729, 1952.

Singer, S.F. "Meteorites and Cosmic-Ray Meters." *Phys. Rev. Letters*, **90**, 168, 1953.

Singer, S.F. "Origin of Meteorites." *Scientific American*, **191**, 36—41, 1954.

Singer, S.F. "The Origin and Age of Meteorites." *Irish Astronomical J.*, **4**, 165, 1957.

Singer, S.F. "Cosmic Ray Evidence on the Origin of Meteorites." *Nuovo Cim.*, **8**, 539—548, 1958.

4. Scientific Uses of Satellites

Elton, R.C., S.F. Singer, A.E. Tripp, and R.C. Wentworth. "Scientific Instrumentation for Interplanetary Vehicles." *National Telemetering Conference Reports*, 312—316, 1958.

Goldman, D.T. and S.F. Singer. "Studies of a Minimum Unmanned

Satellite of the Earth (MOUSE), Part III-Radiation Equilibrium and Temperature." *Astronaut. Acta*, **3**, 110—129, 1957.

Masterson, J.E. and M.D. Ross. "Rock-Air, a Promising New Tool for Specialized High Altitude Research." *Jet Prop.*, **2**, 276—278, 1957.

Singer, S.F. "A Minimum Orbiting Unmanned Satellite of the Earth (MOUSE)." *J. Brit. Interplan. Soc.*, **11**, 61, 1952.

Singer, S.F. "Rocket Exploration of the Upper Atmosphere." (Pergamon Press, London) p. 368, 1954.

Singer, S.F., *Astronautica Acta*, **1**, 171, 1955.

Singer, S.F. "The Effect of Meteoric Particles on a Satellite." *Jet Propulsion*, **26**, 1071—1087, 1956.

Singer, S.F. "Geophysical Research with Artificial Earth Satellites" in *Advances in Geophysics*, Vol. 3, H. E. Landsberg, Ed., Academic Press, New York, 1956.

Singer, S.F. "Studies of Minimum Orbital Unmanned Satellite of the Earth (MOUSE), Part II-Orbits and Lifetimes on Minimum Satellites." *Am. Rocket Soc.* Preprint 160-54; *Astronautica Acta*, **2**, 125—144, 1956.

Singer, S.F. "Project Far Side." *Missiles and Rockets*, **2**, 120—128, Oct. 1957.

Singer, S.F. "Effects of interplanetary dust and radiation environment." Chap. IV in *Intl. Symp. on Phys. and Med. of Atm. and Space*, San Antonio, Nov. 1958 (J. Wiley, New York, 1960).

Singer, S.F. and A.L. Lawrence. "Terrapin - An Upper Atmosphere Vehicle." *Jet Propulsion*, **2**, 281—288, 1957.

Singer, S.F. and R.C. Wentworth. "A Method for the Determination of the Vertical Ozone Distribution from a Satellite." *J. Geophys. Res.*, **62**, 299—308, 1957.

Singer, S.F. and R.C. Wentworth. "A Method for Calculating Impact Points of Ballistic Rockets." *Jet Prop.*, **27**, 407–409 (1957); *Jet Prop.*, **28**, 684—687, 1958.

5. Geomagnetically Trapped Radiation

Griem, H. and S.F. Singer. "Geomagnetic Albedo at Rocket Altitudes at the Equator." *Phys. Rev.*, **99**, 608 (A), 1955.

Lenchek, A.M., S.F. Singer and R.C. Wentworth. "Geomagnetically Trapped Electrons from Cosmic Ray Albedo Neutrons." *J. Geophys. Res.*, **66**, 4027—4046, 1961.

Lenchek, A.M. and S.F. Singer. "Geomagnetically Trapped Protons from Cosmic Ray Albedo Neutrons." *J. Geophys. Res.*, **67**, 1263, 1962.

Lenchek, A.M. and S.F. Singer. "Effects of the Finite Gyroradii of Geomagnetically Trapped Protons." *J. Geophys. Res.*, **67**, 4073, 1962.

Lenchek, A.M. and S.F. Singer. "The albedo neutron theory of geomagnetically trapped protons." *Planet. Space Sci.* **11**, 1151—1208, 1963.

Maeda, K. and S.F. Singer. "Energy Dissipation of Spiralling Particles in the Polar Atmosphere." *Ark. f. Geophys.*, **3**, 531—538, 1961.

Rhodes, R.M. "Study of Auroral Particles." M.S. Thesis, Univ. of Md. Phys. Dept. (1955). *Bull. Am. Phys. Soc.*, **3**, 81, 1959.

Singer, S.F. "Trapped Orbits in the Earth's Dipole Field." *Bull. Am. Phys. Soc.* Series II, **1**, 229 (A), 1956.

Singer, S.F. "Radiation Belt and Trapped Cosmic Ray Albedo." *Phys. Rev. Letters*, **1**, 171—173, 1958.

Singer, S.F., "Trapped Albedo Theory of the Radiation Belt." *Phys. Rev. Letters*, **1**, 181—183, 1958.

Singer, S.F. "New Acceleration Mechanism for Auroral Particles." *Bull. Am. Phys. Soc.* Series II, **3**, 40 (A), 1958. *Ann. Geophys.*, **14**, 433, 1958.

Singer, S.F. "Artificial Modification of the Earth's Radiation Belt." *Advances in Astronautical Science*, Vol. **4** (Plenum Press, New York) p. 335–354, 1959; *J. Astronaut. Sci.*, **6**, 1—10, 1959.

Singer, S.F. "Cause of the Minimum in the Earth's Radiation Belt." *Phys. Rev. Letters*, **3**, 188—190, 1959.

Singer, S.F. "Latitude and Altitude Distribution of Geomagnetically Trapped Protons." *Phys. Rev. Letters*, **5**, 300—303, 1960.

Singer, S.F., *J. Geophys. Research* **65**, 2577, 1960.

Singer, S.F. "Properties of the Upper Atmosphere and Their Relation to the Radiation Belts of the Earth." *Planet. Space Science*, **2**, 165—173, 1960.

Singer, S.F. and A.M. Lenchek, "Geomagnetically Trapped Radiation" in *Progress in Elementary Particle and Cosmic Ray Physics*, Vol. **6**, (ed. by J. G. Wilson and S. A. Wouthuysen), No.-Holland Publ. Co., Amsterdam, 1962.

Wentworth, R.C. "Lifetimes of Geomagnetically Trapped Particles Determined by Coulomb Scattering." Ph. D. Thesis, Univ. of Md. Phys. Dept., 1960.

Wentworth, R.C. and W.M. MacDonald "Pitch Angle Diffusion in a Magnetic Mirror Geometry." *Bull. Am. Phys. Soc.*, **6**, 53, 1961.

Wentworth, R.C. and S.F. Singer. "The Albedo Contribution in the Measurement of Cosmic Ray Primaries." *Phys. Rev.*, **98**, 1546 (A), 1955.

Wentworth, R.C., W.M. MacDonald and S.F. Singer. "Lifetime of Trapped Radiation Belt Particles Determined by Coulomb Scattering." *Phys. Fluids*, **2**, 499—509, 1959.

Van Allen, J.A. and L.A. Frank, *Nature*, **183**, 430, 1959.

Vernov, S.N., A.E. Chudakov, P.V. Vakulov, and Yu. L. Logachev, *Doklady Akad. Nauk S.S.S.R.*, **125**, 304—307, 1959.

6. Magnetic Storm Ring Current

Apel, J.R., S.F. Singer and R.C. Wentworth. "Effects of Trapped Particles on the Geomagnetic Field." *Advances in Geophysics*, Vol **9**, pp. 131–189, (ed. by H. E. Landsberg) Academic Press, New York, 1962.

Singer, S.F. "A New Model of Magnetic Storms and Aurorae" *Trans. Am. Geophys. Union*, **38**, 175—190, 1957.

Singer, S.F. "Role of Ring Current in Magnetic Storms." *Trans. Am. Geophys. Union*, **39**, 532, 1958.

Singer, S.F. "Geophysical Effects of Solar Corpuscular Radiation." *Annales Géophys.*, **14**, 173–177 (1958). *J. Atmos. Terrest. Phys.*, **15**, 48–50, 1959.

7. Interplanetary Dust and Orbital Debris

Bandermann, L.W. and S.F. Singer. "Interplanetary dust measurements near the earth." *Rev. Geophys.* **7**, 759—97, 1969.

Oertel, G.K. and S.F. Singer. "Some Aspects of a Three-Body Problem." *Astronaut. Acta*, **5**, 356–366, 1959.

Oliver, J.P. et al. "Estimation of Debris Cloud Temporal Characteristics and Orbital Elements." *Adv. Space Res.*, **13**, 103—106, 1993.

Singer, S.F. "Interaction of West Ford Needles with Earth's Magnetosphere and Their Lifetime." *Nature*, **192**, 303—306 (1961) and **192**, 1061 (1961).

Singer, S.F. "Interplanetary Dust Near the Earth." *Nature*, **192**, 321—323, 1961.

Singer, S.F. "Dust and Needles in the Earth's Magnetosphere." *Trans. Am. Geophys. Union*, April 1963.

Singer, S.F. and Bandermann, L.W. "Cosmic dust: inter-comparison of observations." *Space Research VIII* (eds. A. P. Mitra, L. G. Jacchia, W. S. Newman), pp. 475–88. North Holland, Amsterdam, 1968.

Singer, S.F. et al. "First Spatio-Temporal Results from the LDEF Interplanetary Dust Experiment." *Adv. Space Res.*, **11**, 115—122, 1991.

8. Exospheres and Magnetospheric Plasma

Hinnov, E. "Optical Cross-sections from Intensity-density Measurements." *J. Opt. Soc.*, **47**, 156–162, 1957.

Liwshitz, M. and S.F. Singer. "Thermal escape of neutral hydrogen and its distribution in the earth's thermosphere." *Planet. Space Sci.* **14**, 541—61, 1966.

Nakada, M.P. and S.F. Singer. "Multiply Charged Ions in the Magnetosphere." URSI Mtg., Washington, D.C., April 1962.

Öpik, E.J. and S.F. Singer. "Distribution of Density in a Planetary Exosphere." *Phys. Fluids*, **2**, 653–655 (1959); *Phys. Fluids*, **3**, 486–488 (1960); *Phys. Fluids*, **4**, 221—233, 1961.

Öpik, E.J. and S.F. Singer. "Escape of Gases from the Moon." *J. Geophys. Res.*, **65**, 3065—3070, 1960.

Öpik, E.J. and S.F. Singer. "Density of the Lunar Atmosphere." *Science*, **133**, 1419—1420, 1961.

Singer, S.F. and E.H. Walker, "Photoelectric Screening of Bodies in Interplanetary Space." *Icarus*, **1**, 7, 1962.

Singer, S.F. and E.H. Walker, "Electrostatic Dust Transport on the Lunar Surface." *Icarus*, **1**, 112, 1962.

Appendix- A Student's Story

Robert C. Wentworth, Ph.D.

8072 Broadway Terrace, Oakland, CA 94611

1. PERSONAL BACKGROUND- A GEOLOGIST'S MICROSCOPE

It was the gift of my father's professional microscope that led to my career in Space Science. My parents were divorced, and I was raised by my mother in Baltimore, Maryland. I attended a small, private high school, and my high school years were noted principally for Honors in mathematics, and for my winning the Maryland State junior badminton championships. However, my father (in Hawaii, C. K. Wentworth was a very well-known geologist) was unimpressed with the latter accomplishment, and sent me one of his professional microscopes as an alternate enticement. I went on to Swarthmore College in Pennsylvania. In my senior year I took a part-time job at the Bartol Institute (under W. F. G. Swann and Martin Pomerantz) reading high-altitude-balloon cosmic-ray exposed film packets under strong microscopes (inspired by my prior experience with my father's microscope). Following graduation from Swarthmore, I joined the physics department at the University of Maryland in the Fall of 1953.

2. COSMIC RAYS AT MARYLAND

One of my first courses at Maryland was a seminar in cosmic rays given by S. Fred Singer, one of the first faculty members recruited by the new department head, John Toll. I took the course because of my prior experience scanning cosmic-ray exposed film plates at the Bartol Institute.

In addition to his regular lectures, Singer suggested topics of current importance to students interested in working with him, and at some point I had a go at one of his 'problems.' It was the analysis of Morrison's magnetic cloud explanation of Forbush decreases. We were able to show, by invoking Liouville's Theorem, that a dynamic, expanding cloud was required-that a stationary cloud would not do the job. I subsequently joined his group.

I recall three early problems that he had suggested. The first resulted in my first publication. It was entitled, 'A method for the determination of the vertical ozone distribution from a satellite' (by Singer and Wentworth) in June 1957 (this was prior to the 'invention' of satellites in October 1957 with the launching of the Russian Sputnik). The second was to develop calculational techniques to determine impact points for high- altitude rockets (I do not think we had a computer yet). Three publications resulted, 'A method for calculating impact points of ballistic rockets' (by Singer and Wentworth) in 1957, 'A method for calculating impact points of ballistic rockets: Convenient representations' (by Singer and Wentworth) in 1958, and, 'Scientific instrumentation for interplanetary vehicles' (by Elton, Singer, Tripp, and Wentworth) in 1958. The third problem was to study the intensity of cosmic radiation as a function of distance in space from the dipole center. This was a pure Størmer problem, and resulted in the publication, 'Cosmic-ray measurements in the vicinity of planets and some applications: Part I, Primary cosmic radiation' (by Singer and Wentworth) in 1959. It should be pointed out that this straight-forward calculation had the potential for revealing the existence of the radiation belts if it had been applied to early deep-space probes such as Sputnik II, and the unsuccessful Project FARSIDE.

3. BEGINNINGS OF SPACE SCIENCE

Three bodies of theoretical work were available during the first few years at Maryland. First was the Chapman/Ferraro plasma cloud theory for the Sudden Commencement/Initial Phase of geomagnetic storms, which was summed up in the book, "Geomagnetism" by Chapman and Bartels in 1940. Second was the orbit perturbation and magnetohydrodynamic work of Hannes Alfven, which was published in book form as, "Cosmical Electrodynamics." Third was the great body of cosmic ray orbit calculations in the earth's dipole magnetic field by Carl Størmer, which was summarized in "The Polar Aurora."

3.1 *Singer's Prediction of the Radiation Belts.*

It seems worth repeating the introduction to my Ph. D. thesis:

'The existence of trapped particles in the earth's dipole magnetic field seems to have been considered first about three years ago. Even though a single particle could not enter into the trapping region from infinity (as shown by Størmer's theory), Singer (1956) reasoned that if a large number of particles were to arrive from the sun, their collective action could perturb the strict dipole field sufficiently to allow entry into the trapping regions. It must be understood that the assumption of particles in the forbidden (actually, 'allowed,' but 'inaccessible') regions was purely a hypothesis and was developed mainly to explain magnetic storms and aurora. Subsequently, Singer (1957) developed a more detailed theory for the main phase of magnetic storms which makes use of the drift of these trapped particles to explain the formulation of a ring current. He further suggested that some of these trapped particles are accelerated to auroral energies (Singer, 1958). Although the initial assumption of particles in the Størmer forbidden (allowed, but inaccessible) regions was purely a hypothesis, the existence of such particles has been demonstrated lately by the Pioneer rockets of Van Allen and Frank (1959) and by Vernov et al (1959).'

Singer asked me to work on the problem of the magnetic effects to be expected from such a cloud of trapped particles (probably in 1956). He suggested assuming a plasma having an energy density equal to 50% of the energy density of the earth's dipole field in the equatorial plane. The plasma would be isotropic except for the so-called loss cone. The plasma would be injected from 4 to 10 earth radii. We had a computer by then, and were able to calculate particle densities, magnetic moments, and drift currents in three dimensions, starting with conditions at the equator.

Singer published his theory, "A New Model of Magnetic Storm and Aurorae" in Transactions of the Geophysical Union in June, 1957. Subsequently, this work was expanded, and published as, 'Aspects of the magnetic storm belt' (by Singer and Wentworth) in July, 1959, and as, 'Effects of trapped particles on the geomagnetic field' (by Apel, Singer and Wentworth) in 1962 (cf. Fig. 2).

A serious question was asked about this material: what would prevent the ring current from spiraling into the earth? We had several answers. First, each of the individual particles was in a stable orbit, as viewed as an Alfven perturbation, or as a Størmer 'low-energy cosmic ray.' In addition, each of the spiraling particles was the equivalent of a small dipole magnet, parallel to the earth's dipole center. As such, the repulsion of the parallel magnets would provide the force to counter the tendency to spiral in toward the earth.

3.2 *Rockets.*

I am only peripherally cognizant of Singer's involvement with rocket hardware. He was associated with Project FARSIDE, a scheme to launch a 4-stage cluster of rockets suspended below Skyhook balloons. I did some work on a Geiger counter in the payload, and if everything had worked perfectly, the radiation belts might now be called the "Singer Belts." Singer was also involved in the development of an intermediate altitude rocket, called the "Terrapin" (the nickname of the U. of Maryland's football team). At a later date (probably in 1959) Singer was entertaining a delegation of Russian scientists. His graduate student, Dick Bettinger, had written a computer program to calculate the position of a rocket following launch. As the Russians crowded around, Bettinger pushed the button. Out typed the computer:

Time	Altitude
0 sec	0 feet
1 sec	-16 feet
2 sec	-64 feet

After a moment of shocked silence the Russians roared with laughter, and one congratulated Singer on his excellent simulation of an American rocket launch (Bettinger had not fueled the rocket, and it had fallen off the launch pad)!

3.3 *Lifetimes of geomagnetically trapped particles.*

Following Van Allen's discovery of the radiation belts, Singer had his group move full speed to develop a theoretical understanding of the physics of the region. Singer, himself, was particularly involved with his neutron albedo theory for the high-energy, inner proton belt, and I started working with Bill MacDonald on an application of the Fokker-Planck equation on the magnetic mirror geometry of the radiation belts; my Ph.D. thesis was published in Sept. 1959 as, 'Lifetimes of geomagnetically trapped particles determined by Coulomb scattering' (by Wentworth, MacDonald, and Singer). Somewhat later, I published, 'Pitch angle diffusion in a magnetic mirror geometry' in 1963.

4. PC-1 MICROPULSATION PEARLS

Following receipt of my Ph.D. I joined the Lockheed Palo Alto Research Lab in California in December of 1960. There, I completed my Ph.D. topic paper, and following the exodus of Francis Johnson, Alex Dessler, and Bill Hanson to Texas in 1962, I began working with Lee Tepley on PC-1 micropulsation 'pearls' (so named by the Russian physicist V. M. Troitskaya). This work resulted in many joint publications with Tepley, covering the years from 1962 through 1966. In 1968, I wrote a review article which was published in a UCal, Berkeley, conference proceeding edited by Stan Ward. To quote from the abstract,

'It is now known that PC1 pearls propagate as wave packets of finite duration along high-latitude geomagnetic field lines, bouncing between hemispheres in a manner exactly analogous to VLF whistlers'.

This was shown by Tepley's use of the sonogram (cf. Figure 4) to analyze signals from high-latitude stations (run by Heacock), mid-latitude, northern hemisphere stations (Palo

note that I had shown earlier that PC-1 pearl events tended to occur predominantly during the week following sudden-commencement magnetic storms).

5. GEOMAGNETIC TAIL

The initial view of the magnetosphere was formulated in the interval from about 1957 through about 1965, that the Chapman/Ferraro plasma cloud (or the continuous Parker 'solar wind') would compress and limit the earth's geomagnetic field within what was called a 'tear-drop' cavity (as suggested by Francis Johnson in 1960). The direct pressure would create a magnetosphere boundary on the sunward side at some 10 R_E, while the thermal pressure would create an antisolar boundary at some tens of R_E downstream on the back side.

In 1962, I had a conversation with Alex Dessler and Bill Hanson. They asked me what would happen to particles trapped in the distant anti-solar portion of the Johnson 'tear-drop' magnetosphere. There, contours of constant magnetic field strength at the equatorial plane would not continue on around to the sub-solar region, but would instead intersect with the boundary. I pointed out that particles drifting on such contours would reach the boundary and then pass on out of the magnetosphere. Did this mean that there would be no trapped particles on such contours? Yes, but there would be a large amount of external plasma crossing the boundary, drifting across the distant tail, and then escaping again on the other side. And this would not be just an occasional particle. By Liouville's theorem it would be the whole weight of the external plasma, and it would make the distant closed tail untenable. It would blow the tail away! None of us followed up on this point! At that time I was still working on my follow-up Ph.D. research paper.

Two years later IMP-1 discovered the geomagnetic tail (Norman Ness, 1965, although apparently Ed Smith had seen it earlier in 1962), and I published my first paper on this subject, 'Diamagnetic Ring Current Theory of the Neutral Sheet and its Effects on the Topology of the Antisolar Magnetosphere' in 1965 (cf. Figure 5). This suggestion was not received with great favor at Lockheed. I remember an announcement to the Lockheed assembly by Roli Meyerott that,

'Wentworth will be giving a talk on his theory of the geomagnetic tail at NASA, Ames. I talked to Martin Walt, and he said that Wentworth's ideas are so full of holes that they cannot be shot down!'

This process was simultaneously, and independently suggested by Jules Fejer in 1965, and came to be known as the 'gradient-drift' effect.

Subsequently, I won a Guggenheim Fellowship to study this problem, and ended up publishing, 'The Geometry of the Magnetosphere' in 1967. In effect, I suggested that the extended tail was a second, unexpected solution to an 'extended,' field-free, finite-temperature Chapman-Ferraro boundary-value problem. In my introduction, I state,

'Previous suggestions have been based on the implicit assumption that the unperturbed finite-temperature solar wind would compress the geomagnetic field into a closed cavity in the form of a tear-drop. They have postulated the existence of

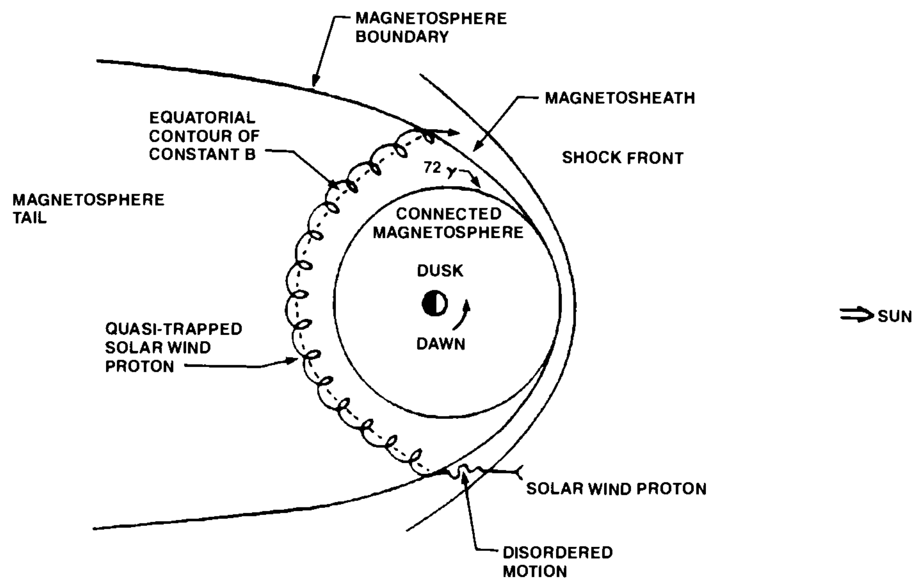


Figure 5. Equatorial plane of the magnetosphere. The injected solar-wind proton is shown following a contour of constant magnetic field strength across the tail (beyond 10 R_E- Wentworth, 1968).

Jacobs, J. A. and T. Watanabe, 'Micropulsation Whistlers', *J. Atmos. and Terr. Phys.*, **26**, 825–829, 1964.
Tepley, L. R., 'Observations of Hydromagnetic Emissions', *J. Geophys. Res.*, **66**, 1651—1658, 1961.
Tepley, L.R., 'Low-Latitude Observations of Fine-Structured Hydromagnetic Emissions', *J. Geophys. Res.*, **69**, 2273—2290, 1964.
Tepley, L.R., 'Regular Oscillations near 1 c/s Observed at Middle and Low Latitudes', *Radio Sci. J. Res.*, NBS 69D, 1089—1105, 1965.
Tepley, L.R., 'Recent Investigations of Hydromagnetic Emissions, Part I. Experimental Observations', *J. Geomag. and Geoelect.*, **18**, 227–256, 1966.
Tepley, L.R. and R.C. Wentworth, Hydromagnetic Emissions, X-Ray Bursts, and Electron Bunches, Part I: Experimental Results', *J. Geophys. Res.*, **67**, 3317–3333, 1962.
Tepley, L.R. and R.C. Wentworth, 'Hydromagnetic Emissions Associated with the Magnetic Storm of September 30, 1961', *J. Geophys. Res.*, **68**, 3733–3737, 1963.
Tepley, L.R., R.R. Heacock and B.J. Fraser, 'Hydromagnetic Emissions (Pc 1) Observed Simultaneously in the Auroral Zones and at Low Latitudes', *J. Geophys. Res.*, **70**, 2720–2725, 1965.
Tepley, L.R. and R.K. Landshoff, 'Waveguide Theory for Ionospheric Propagation of Hydromagnetic Emissions', *J. Geophys. Res.*, **71**, 1499–1505, 1966.
Troitskaya, V., 'Pulsations of the Earth's Electromagnetic Field and their Connection with phenomena in the High Atmosphere, *J. Geophys. Res.*, **66**, 5—18, 1961.
Wentworth, R.C., 'Evidence for Maximum Production of Hydromagnetic Emissions Above the Afternoon Hemisphere of the Earth, Parts I and II', *J. Geophys. Res.*, **69**, 2689–2705, 1964.
Wentworth, R.C., 'Enhancement of Hydromagnetic Emissions Following Geomagnetic Storms', *J. Geophys. Res.*, **69**, 2291–2298, 1964.
Wentworth, R.C., 'Recent Investigations of Hydromagnetic Emissions, Part II. Theoretical Interpretation', *J. Geomag. and Geoelect.*, **18**, 257–273, 1966.
Wentworth, R.C., 'Instabilities and Cyclotron Resonances of Hydromagnetic Pearls', U. C. Berkeley Conference, Stan Ward ed., 1968.
Wentworth, R.C. and L.R. Tepley, 'Attenuation of Hydromagnetic Emissions in the Ionospheric Waveguide', Semiannual Rept., Contract AF 33(657)-15303, Air Force Technical Applications Center, March 25, 1966.
Wentworth, R.C., L.R. Tepley, K.D. Amundsen and R.R. Heacock, 'Intra- and Interhemispheric Differences in Occurrence Times of Hydromagnetic Emissions', *J. Geophys. Res.*, **71**, 1492–1498, 1966.

Magnetosphere Geometry

Dessler, A.J., 'Length of Magnetospheric Tail', *J. Geophys. Res.*, **69**, 3913–3918, 1964.
Fejer, J.A., 'Geometry of the magnetospheric tail and auroral current systems', *J. Geophys. Res.*, **70**, 4972–4975, 1965.
Johnson, F.S., 'The Gross Character of the Geomagnetic Field in the Solar Wind', *J. Geophys. Res.*, **65**, 3049–3051, 1960.
Parker, E.N., 'Interaction of the Solar Wind with the Geomagnetic Field', *Phys. Fluids*, **1**, 171–187, 1958.
Smith, E.J., 'A Comparison of Explorer VI and Explorer X Magnetometer Data', *J. Geophys. Res.*, **67**, 2045, 1962.
Wentworth, R.C., 'Diamagnetic ring current theory of the neutral sheet and its effects on the topology of the antisolar magnetosphere', *Phys. Rev. Letters*, **14**, 1008–1010, 1965.
Wentworth, R.C., 'The geometry of the magnetosphere', *J. Geophys. Res.*, **70**, 4582–4586, 1967.
Wentworth, R.C., 'The geometry of the magnetosphere. II: A weakly connected model of the geomagnetic tail', *J. Geomag. and Geoelect.*, **20**, 299–304, 1968.
Wentworth, R.C., 'A 'geomagnetic' boundary condition at infinity leads to an extended-tail solution to the original field free Chapman/Ferraro boundary value problem', AGU 1993 Fall Meeting, Dec 6-10, San Francisco, California, 522 (SM31A-30, Poster).

An Education in Space Physics

D.J. Southwood

Blackett Laboratory, Imperial College, London SW7 2BZ, United Kingdom

The author was a student at Imperial College in the mid-sixties, a time when much of the initial exploration of the magnetosphere had been done and many ideas that were later to be seen to be correct were around but often not appreciated. The paper reviews the author's experiences and the ideas he picked up then.

BACKGROUND

I make it into this volume by the skin of my teeth. In 1957, I remember, as a schoolboy, greeting the arrival on the scene of Sputnik with amazement and no sense that within a decade I would be involved in space science. It all seemed remote from the west country of England where I was growing up.

By 1966 things had changed. I had come to London where possibilities seemed greater. What set me off on my space science career was what in retrospect was the most enormous piece of luck, although I did not recognise it at the time. I was about to graduate (in Mathematics) from Queen Mary College, London and I was set on doing a doctorate with Vincente Ferraro who had been my undergraduate tutor. Fate intervened. Ferraro had a coronary and was ordered to cut back on his work load. He sent me across town to Imperial College where Jim Dungey had recently arrived. Jim was appointed first as a Reader then rapidly given a chair (Professorship) of which more below.

One says of teachers that they 'taught me all I know'. Of course, everyone says this about influential teachers. Jim did that, but also, by his remarkable prescience or intuition of the magnetosphere, he gave me an enormous headstart in space physics. Happily, as I shall reveal below, in this case I have documentary proof. The effect was that by 1970, I had a way of understanding the magnetosphere that worked but was not generally accepted for another decade or so.

I remember an interview with a tired-looking Ferraro when he explained that he was not going to take me on but that he had recommended me to Dungey. He was a very nice man and I think that he felt badly about sending me off. As if to apologise, he gestured with a sweep of his hand to a shelf of yellow-spined JGR's on his bookshelves and said something like "I cannot keep up with the rate of new material. One of these comes every month!". I shudder to think what he would make of what now appears monthly in just JGR-Space Physics. In those days, JGR covered all geophysics and maybe half a dozen papers on any space-related topic.

I went across town to meet Jim Dungey in either January or February 1966. I immediately liked him and, what was more important, he accepted me. He decided to take me on after an interview where I remember him doing most of the talking. As far as I can recall, he outlined two potential research problems. One was about the generation of magnetohydrodynamic waves on the boundary between solar wind and magnetosphere and the other was the acceleration of charged particles at neutral points. What else we discussed, I do not know. However, I was sold on the subject.

The former topic became central to my PhD thesis and I have worked on magnetospheric magnetohydrodynamic wave problems throughout my career. Here was an area where there was a lot of work to do and another paper is required to do justice to the way the field has developed since then. However in terms of what I remember of that magic time when I was a PhD student the subject of my thesis is an almost incidental fact.

Discovery of the Magnetosphere
History of Geophysics Volume 7

IMPERIAL COLLEGE

The research group I joined at Imperial in 1966 was small. As I remember, there were three other students, one post-doc, two faculty other than Jim and two US visitors, Bill Ross (from Penn State) and Ted Speiser who had been one of Jim's students at Penn State. Later another student, more or less a twin with me, Maha Abdalla, joined the group. Subsequently, in the summer of 1967, an undergraduate student, Stan Cowley, came to work for the post-doc, Roger Etherington (an endearing but fierce anarchist who sadly was to die from Hodgkin's disease within a few years).

The group was small but well connected. Americans at the time seemed to travel more freely than others and there was a regular stream of visitors to the group who came through, no doubt to talk to Jim but also because London was an exciting place to be at that time. David Beard came so often that we named a room after him (also known as the 'magnetospheric cavity'). Norman Ness who had just discovered the magnetic tail of the Earth (of which more below) was a regular visitor. I remember Fred Scarf whose enthusiasm for measuring waves from spacecraft seemed unusual at the time and yet fundamentally so sensible. Non-US visitors included Roger Gendrin and Valeria Troitskaya. Devrie Intriligator, who was later to convince me Los Angeles was a good place to live, spent a summer with a desk in the office I used. In my years as doctoral student I shared the office with, at different times, Ira Bernstein (the plasma physicist), Carl McIlwain, Alfredo Baños, Chuck Sonett and the South African, Desmond Clarence. All this provided a wonderful education.

In fact Jim's group was a subgroup of a cosmic ray group which at some time changed its name to 'Cosmic Rays and Space Physics' and was headed by Harry Elliot. The other part of the group was firmly experimental. People went off to fly balloons in Africa and the like but the new activity was building space instrumentation. There was a large involvement in the British Ariel programme and the new ESRO (European Space Research Organisation) programme - the joke was that the initials HEOS for the ESRO spacecraft launched in 1972 stood for 'Harry Elliot's Own Satellite'! Through this group I met people like André Balogh, Bob Hynds, and Peter Hedgecock who were going to become close colleagues later as well as George Haskell (with whom I worked and who later got me strongly involved in space politics). The division in the group also marked elements in the origin of the field of work. I remember crudely classifying people in the field by their origins - either as radio (ionospheric) physicists or 'sawn-off' cosmic ray men. I guess temperamentally the Dungey group was more the former, Elliot's clearly the latter.

SPACE PHYSICS IN 1966

The education I received in research working with Jim went much further than working on my particular thesis topic. The smallness of the group meant that one knew what everyone was working on. In fact, the first thing I remember intensely thinking about was not associated with my thesis at all. 1966 had seen the publication of a paper by *Kennel and Petschek* [1966] which purported to provide a quantitative means of understanding the trapping of radiation belt particles. This paper, combining as it did microscale processes and global effects, was very influential in my own coming to terms with what space plasma physics was about. It needs to be understood that the notion of how collision-free media behaved was still controversial. A critical element of the Kennel-Petschek story was the non-linear (pitch angle) diffusion of particles which was self-sustaining because particles released energy as they diffused. The notion of diffusion being a saturation process was central in the recently derived theory for the saturation of electrostatic plasma oscillations derived both in the USSR [*Vedenov* et al., 1962] and in the USA [*Drummond and Pines*, 1962] at around the same time. I believe the junior faculty member, Jeff Klozenberg, who had come to Jim from Culham Laboratory, had been to a summer school in France in 1966 and paid the price of his ticket by lecturing to us on quasilinear theory through the autumn of that year.

Jim Dungey dominated activity in the group. Once I started working with Jim, I realised that his mode of thinking was regarded by many as eccentric. This could put people off attending to what he was saying. His presentation was the thing that was really different. He speaks and writes 'telegraphically', i.e. in short sentences, which express only the essence of an argument. I found his brevity immensely appealing (perhaps because I have never developed the knack myself). Jim's response to ideas was always interesting. Assuming you had already grasped the obvious, he would leap to a further implication that you had not yet reached. Once one had the hang of it, one learnt a lot.

It is hard to get back into the mind of the time. So much of what seemed radical and off the wall at that time is now commonplace. In preparing this article I sat and tried to think what I learnt then and what I learnt later. Shockingly, I concluded that just about every useful idea I have exploited in a career in solar terrestrial physics spanning thirty years had its seed in ideas picked up whilst a student at Imperial College. However it did not seem like that at the time. What has happened in the intervening years is that the old orthodoxies have been swept away and the old barriers to understanding have gone with them. In order to patch together the feelings of the time, I would like to explore this a little.

Jim Dungey's greatest contribution to magnetospheric physics must be his open magnetosphere model. It was launched on the world in a famous Phys. Rev. Lett. in 1961 [*Dungey*, 1961], but in 1966 it was clear that Jim's ideas in this sphere were still regarded by many people working in solar terrestrial physics as from the radical fringe. A direct benefit to me was that accepting the basic ideas of the open model in the sixties, gave me a ten year head start in my career. It all seems so straightforward now that magnetic reconnection couples the solar wind and magnetosphere differently when the interplanetary field is northward or southward. It is hard to see why people had trouble coming to terms with the idea.

A couple of examples of the barriers to understanding ideas that then existed might illustrate the dark ages we lived in. I was sent to Jim by Ferraro. Ferraro always took a personal interest in my progress. I felt he wanted to be reassured that he had not disadvantaged me by sending me away. Ferraro's original claim to fame was his work while a student at Imperial College with Sydney Chapman in the early thirties. Chapman and Ferraro had first suggested that the Earth's magnetic field might be contained within a cavity by the passage through space of corpuscular material emitted from the Sun during solar disturbances. The cavity would be bounded by a thin boundary later to be called the magnetopause (which to this day is still said to carry the 'Chapman-Ferraro' current). The formation of such a boundary was still an interest of Ferraro's in the late sixties. A competing idea of magnetic storms was advanced by Alfvén in the 1940's who postulated the penetration of an electric field from interplanetary space into the Earth's field during storm-time. Alfvén's model completely ignores the formation of the Chapman-Ferraro boundary, which of course, to a first approximation excludes the magnetosphere from direct experience of the solar wind electric field. As a model of magnetic storms, both ideas were wrong. However because the Sun continually sends out a stream of charged particles into space, the solar wind, Chapman and Ferraro's idea certainly explained the existence of a terrestrial magnetosphere whilst Alfvén's idea of an electric field, although he had the direction precisely wrong, is arguably at the root of our current understanding of geomagnetic disturbances. For Ferraro the penetration of the interplanetary field remained a major puzzle. At one of the early national MIST (Magnetosphere, Ionosphere and Solar Terrestrial Physics) meetings I can remember him asking after a talk by Stan Cowley about the magnetic neutral sheet, "But where does the electric field come from?". In similar spirit, I recall talking in La Jolla with Hannes Alfvén about his admiration for Jim Dungey. Alfvén, a man not given to acknowledging lapses, admired Jim, he said, because he not only recognised that a magnetopause would form but also that the electric field would penetrate it. Alfvén's real admiration was for the fact that Jim had seen that a magnetopause would form. When I gently suggested that Chapman and Ferraro had got there first, he would have nothing of it "No, their idea was wrong. They did not include the solar magnetic field.".

A second issue which might illustrate the dark ages, concerns the Earth's tail. This was a relatively recent discovery (by IMP I in 1964) when I was a student. The presence of an extended magnetotail behind the Earth proved conclusively that the magnetosphere was not raindrop-shaped but that the magnetic field splits into northern and southern lobes containing respectively field pointing towards and away from the Earth. Its length was not clear and debate started in the pages of JGR [*Dessler*, 1964; *Dungey*, 1965]. In fact, the existence of magnetotail at all is extremely good evidence of the need for the Dungey open model magnetosphere. The field lines have in some way to be dragged out to obtain the observed configuration. One does not even need to believe in the tenets of magnetohydrodynamics to see that the body forces exerted by the field in the near tail are towards the Earth and extreme and there must be an effective mechanism for tugging in the opposite direction. Oddly the discovery of the tail did not seem to be greeted at the time by any large body of people as a triumph for the open magnetosphere model propounded by Dungey. Part of the reason for this must have been the topological nature of the sketches showing the field configuration which bear little resemblance to the actual geometry [cf. Figure 1, which I discuss later.]. However, even today an audience can seem to have trouble making the connection, but against a background of the apparently competing ideas of Chapman and Alfvén and others, it was much harder.

There were a lot of contradictory and confused ideas around in 1966. It is important to understand that fact or one cannot understand why it took so long to make the ultimate progress we did make.

ENLIGHTENMENT

On the issue of how the magnetosphere worked I made my own private resolution. In those days students had a lot more time to read around their work (and there was a lot less to read). Jim did not spoon-feed one and one was left very much to do one's own background reading. As a result, I can remember coming to an independent conviction that Dungey had to be right after a bout of browsing in JGR and the limited number of conference proceedings then available. Epiphany was seeing Don Fairfield's work

[*Fairfield and Cahill*, 1966] which correlated the geomagnetic disturbance DS current system with southward interplanetary field occurrence. Don had been a graduate student of Jim's at Penn State and I think that the work had been part of his thesis. Since then innumerable connections have been established between the sense of the interplanetary field and geomagnetic response. Such effects, whatever they may be, cannot be explained by magnetohydrodynamics alone. MHD effects are insensitive to the sense of the magnetic field. The reconnection process which Dungey postulated to occur at the magnetopause did depend on the sense of the external field. Once you had accepted that reconnection occurred, many related phenomena or epiphenomena were open to explanation. In particular, one was no longer puzzled by why there was a magnetopause and/or by why there was an externally driven magnetospheric electric field.

In fact, Jim Dungey was supremely right about the open magnetosphere but in fact he was right about lots of things, indeed rarely wrong. In 1966, he had thought about many of the basic problems of the magnetosphere in his own way. Only slowly did the rest of the world come round to that way of thinking. My luck was to meet him and pick up the rules at just the right time.

INAUGURATION

'History belongs to the winners.' In planning this article, I was haunted by the fact that so much of what Dungey's group thought was true in 1966 was proven correct that the story would all seem a little unlikely to an audience who did not remember the time in question.

Happily in pursuing evidence of what we did and did not know I came across excellent documentary proof of Jim's overall prescience in matters solar terrestrial. Whilst I was looking for material for this article, my secretary proposed I look at Jim's inaugural lecture. This was a great idea. The inaugural text encapsulates well the knowledge one came to take for granted at Imperial College in the late sixties. It stands well the test of time.

At Imperial College, on being appointed to a chair (the term for making the rank of professor in England) the new professor is expected to give an inaugural lecture. Jim gave his lecture on May 3rd 1966. At that time all inaugural lectures were published, a tradition that has lapsed somewhere in the intervening years. (No doubt because now we are all so busy filling the pages of journals like JGR with our outpourings instead.) Publication took the form of an annual collected volume of inaugural lectures. In addition bound off-prints were made of each lecture (published at a price of two shillings and sixpence by Imperial College).

Jim's lecture title is rather prosaic: 'The Magnetosphere'. After a crack about taking Patrick Blackett's old chair, Jim mentions the long interest he has had in the origin of the aurora. He refers to Chapman and Ferraro (who are clearly both present in the lecture theatre) and their theory that that the Sun might throw out streams of ionized material and that these would then be held off from the Earth by the formation of a magnetic cavity. There is no doubt in Jim's mind about the origin of the idea of the magnetopause here. He then goes on to outline Ludwig Biermann's and Eugene Parker's post-war contributions to the recognition that there could be a continuous solar wind culminating in the Mariner II spacecraft's actual measurement of a supersonic solar wind as Parker had predicted.

Jim then goes into a concise description of the basic unifying concept of MHD, the notion that the magnetic field is frozen into the plasma and field lines move with the plasma. He then introduces another central tenet of MHD, magnetic field tension 'appreciated by Faraday but then somewhat forgotten for a time'. Oddly he does not mention the allied concept of field pressure which is central to the formation of the magnetopause and the Chapman-Ferraro cavity. Tension, however, is the central element in the open magnetosphere, the model he is to describe later.

Next he describes the solar wind and the interplanetary field. The source of heat in the corona which causes the solar wind to flow was a mystery then as it still is now but once given the high temperature of the corona, the outflow is a consequence. The recently discovered interplanetary field sector structure is described as a simple consequence of frozen-in flow. The solar wind established, he goes on to discuss the current (i.e. 1966) state of the Chapman-Ferraro problem of defining the magnetopause shape assuming an unmagnetised solar wind incident on the Earth's dipole field. David Beard (no doubt present) comes in for praise. There is mention of hypersonic gas dynamic flow models for the solar wind flow about the magnetosphere being developed. Almost certainly Jim had in mind the work by John Spreiter and colleagues whose first publications on the solar wind magnetospheric interaction were appearing. These workers included an upstream shock but retained the unmagnetised assumption for the solar wind for dealing with the dynamics of the interaction. Even today, thirty years on, theory has not completely included the magnetic field in our understanding of magnetosheath dynamics and so, as my colleague Margaret Kivelson and I know, one can still get involved in controversy on the topic.

Finally the text gets to what is Dungey's magnum opus, the open magnetosphere. It is not introduced in any way that would lead one to suspect that 30 years on it would be regarded as the most important thing the speaker had done.

Neither is there any indication of the passions that the model could arouse even then and the controversies and arguments that would rage for the next decade and a half until it achieved general acceptance. Rather Dungey writes in the spare economical style mentioned earlier which repays attention to each word.

Figure 1 is reproduced from the lecture. Jim explains that it was his PhD supervisor, Fred Hoyle, who suggested that the interplanetary magnetic field could be important in the magnetospheric interaction (I wished I had known this reference when I talked to Alfvén.). One starts (as Alfvén's comments implied) by ignoring the Chapman-Ferraro effect and simply adds an interplanetary field to the Earth's dipole. The field lines then interconnect from Earth to interplanetary space. If the interplanetary medium is moving, the existence of an electric field both in space and in the Earth's environment is natural. Even without allowance for the Chapman-Ferraro effect, current sheets form between the solar and terrestrial plasma where the field reverses and the plasmas move together. At the current sheet the frozen-in approximation breaks down and field lines break and change partners, the process being called 'reconnection'.

The field topology of Figure 1 implies that the polar cap field lines of Earth extend into the solar wind. Although the sketch is topologically sound it does not show the extended field configuration that naturally results as field lines connecting to earth at one end and to the solar wind at the other are drawn out by the solar wind flow behind the Earth to form the tail. In this sense the tail is a natural result of the open model.

Dungey then goes on to discuss two other consequences of the open model. The first is the ionospheric circulation system predicted by the model. This is the effect that Jim himself credits with inspiring him that the open model would work. The electric field imposed on the polar cap field lines by the solar wind flow must extend down into the ionospheric levels and there it must drive horizontal ohmic currents. The most naive view of the ionosphere is to think of the ions as bound by collisions to the neutrals. In contrast the electrons can still move with the field. The horizontal current induced by the field line flow is thus opposite to the projection of the field line flow on to the ionosphere. The global pattern of current that one derives is that of the DS system. It was this very current system whose strength Fairfield had used as a measure of geomagnetic disturbance and had shown was correlated with the occurrence of southward interplanetary field just as is required for reconnection. Oddly, Jim makes no reference to Fairfield here.

The next part of the lecture is a delight for what is missing. Dungey describes some of the implications of

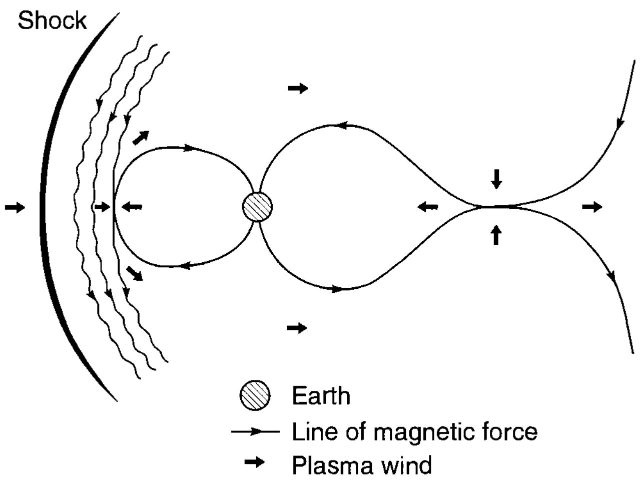


Figure 1. Taken from J.W. Dungey's published inaugural lecture, the original caption of this figure reads "Outline model of the magnetosphere". The diagram's purpose is to show the topology but its simplicity (and the absence of a magnetopause) certainly confused the more literal-minded in the community.

unsteady behaviour in his model. He refers to magnetic bays and to 'pt' pulsations and auroral storms and the auroral oval about the pole. He describes the characteristic sudden brightening and the subsequent explosive-like behaviour of the aurorae and associated phenomenology which occurs in conjunction with a magnetic bay. He also says that at the time of magnetic bays the tail field is seen to suddenly decrease with bay onset and that the aurorae move poleward after breakup. Here is where Jim relates the bay to the DS system and mentions the correlation with southward interplanetary field found by Fairfield and Cahill. He fits all the phenomena into the framework of his reconnection model of the magnetosphere by postulating a sudden onset of magnetic reconnection in the magnetotail current sheet as the seat of activity.

The words that are missing are, of course, 'magnetospheric substorm'. The terms 'bay' and 'pt' have now disappeared. The bay is the trace on the ground magnetogram which looks rather like a bay on a coastal map. The 'pt' described as having 'a pizzicato waveform like that of a plucked string with a period of a few minutes' is now known by the name 'pi2'. Dungey's description of what we know as the magnetospheric substorm would be well-recognised today.

Less well recognised now but more familiar in 1966 were the Van Allen radiation belts. He credits Alfvén with the introduction of the notion of adiabatic invariants to explain their trapped orbits. Dungey was involved in one of the critical studies that showed the external source of the belts [*Nakada* et al., 1965]. This important and perhaps

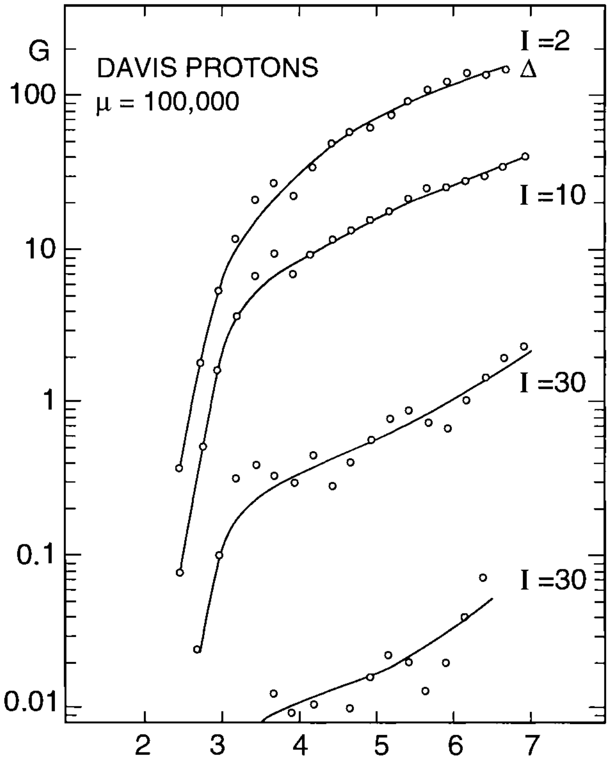


Figure 2. The distribution function of energetic protons plotted against distance from the Earth for fixed values of the two adiabatic invariants, μ, I. The outward gradient is evident and the external source thus is pinpointed. Once again, the figure is taken from Jim Dungey's inaugural lecture. This (informative) version of the figure is interesting to compare to that used in the actual *Nakada* et al. [1965] publication.

under-recognised paper resolved the external nature of the source by looking at the particle distribution as a function of adiabatic invariants. Figure 2 (also reproduced as is from Dungey's inaugural lecture) shows that distribution function (or phase space density) deduced for different values of adiabatic invariants (μ, magnetic moment, J longitudinal invariant) increases toward higher L values. The inaugural mentions the external source and suggests that the particles enter from the tail. Jim then notes that the system pumps the particles up to their final energies by a form of stochastic acceleration. We take this for granted now. These were new and even radical ideas in 1966.

The next section is specialist also. He describes waves. All the waves mentioned are electromagnetic. Kelvin-Helmholtz instability on the magnetopause is mentioned as are whistlers, one of the earliest evidences for the existence of the magnetosphere. The latter waves are very interesting from a plasma physics point of view because of the clearly structured nature of their geophysical emissions. Jim continues with some insights into the non-linear interaction of waves with particles in a collision-free plasma and the phase space 'stirring' that is behind phenomena like VLF-stimulated emissions.

What is missing is any discussion of electrostatic waves or indeed of any of the waves that do not produce a direct electromagnetic response on the ground. This was an area of experimentation which was neglected in the early days. Not until Fred Scarf had flown his electric antenna on OGO 5 in the immediate years to come were these waves detected or their significance appreciated and only in the next decade was the Earth's aurora seen as a source of radio waves. Subsequently, the sophistication of wave instrumentation improved substantially and later spacecraft like GEOS and ISEE had very fine plasma wave measurements. Some of the finest magnetospheric plasma physics has come out of such wave instruments and the belief expressed by Dungey that the magnetosphere could be a laboratory for plasma physics has been well borne out.

CLOSING REMARKS

In respect of using the magnetosphere as a laboratory for cosmic plasma physics, the recent loss of Cluster is particularly tragic and in this respect the closing section of Dungey's inaugural written almost exactly thirty years before the loss of Ariane 501 is extremely ironic. Whole sections of Jim's prescient writing merit quoting. Remember these segments were written thirty years ago.

"Looking to the future I believe that progress requires bunches of satellites, though these are as yet in no published program................."

"When one comes to study waves, bunches of satellites are also needed from several points of view. First one wants to know the geometry of waves and second their direction of propagation. For any magnetic disturbance it would be extremely useful to obtain the curl of the magnetic field because this tells one the current."

"Steady currents are of direct interest and in the case of waves the more reliable part of hydromagnetic theory then enables one to calculate the electric field and the flow of energy which is a very important guide to the location and strength of the source of the waves. Unfortunately, few people yet appreciate the need for satellite bunches and, since satellites are being launched singly, the scientific returns are less than they could be."

In the event, 'Cluster' was a much nicer name than 'Bunch'. It was to be a wonderful mission, one I had waited for throughout my career. I trust that the idea for what would have been the ultimate space physics mission will

Modeling Solar Wind Flow Past the Magnetosphere

John R. Spreiter

Mechanics and Computations Division
Stanford University, Stanford, CA 94305

S. S. Stahara

RMA Aerospace, Inc.
Cupertino, CA 95014

Efforts of the authors and their associates at NASA Ames Research Center and Stanford University to model the interaction of the solar wind and the magnetosphere are described. Primary emphasis is on the discovery period of the sixties, but the environment and prior aerodynamic research at Ames that facilitated this work are also discussed. At the beginning of the discovery period, attention focused on predicting the location of the magnetopause using the corpuscular stream model developed by Chapman and Ferraro over the preceding three decades, but still unsolved. The earliest direct observations of the magnetopause indicated that it was located about where the new solutions indicated. The data also revealed other features that could not be explained by the C-F model. Recognition that certain of these were similar to those associated with a bow shock upstream of a blunt-nosed body in a supersonic stream led to alternative models based on continuum fluid concepts. The earliest of these employed gasdynamic theory, but it was soon replaced by its magnetohydrodynamic MHD counterpart to account for features associated with the magnetic field. The nonlinear partial differential equations of MHD were too difficult to solve with the computers and numerical methods available at that time, however, and it was necessary to introduce simplifications to obtain approximate solutions of useful accuracy. Their nature and limitations are discussed, and an account is provided of how they were gradually refined or removed over the years. This account illustrates how steady progress in modeling was facilitated by the availability of direct observations in space to guide the theory and validate the results, major advances in the capability of computers, numerical algorithms, and MHD theory, and seemingly unrelated advances in supersonic aerodynamics.

PREPARATION IN THE PRE-NASA YEARS

Although this article has two authors, much of it is written in the first person by JRS and is about his activities during the discovery period from the late fifties to the early seventies, and of the preparation that facilitated his rapid transition to magnetospheric studies from prior research in aerodynamics. Association with SSS began in 1964 when

Discovery of the Magnetosphere
History of Geophysics Volume 7

he enrolled, as a graduate student in the Aeronautics Department at Stanford University, in a course in Space Physics taught by JRS, who was then Chief of the Theoretical Studies Branch in the Space Science Division at NASA Ames Research Center and a Lecturer at Stanford. Active collaboration began in early 1969, immediately after SSS earned his Ph.D. at Stanford, and shortly after JRS left Ames to be a Professor at Stanford. JRS had also become a consultant to Nielsen Engineering and Research Inc. in Mountain View, CA almost immediately after joining Stanford to assist Jack Nielsen, a distinguished aerodynamicist formerly employed for many years at Ames, in developing a research program in transonic aerodynamics. Soon thereafter, SSS was employed by NEAR, and JRS was retained to direct the work. Joint work in aerodynamics continued for several years, but magnetospheric modeling began before long and has continued to the present day. In 1984, SSS left NEAR to form his own research company, RMA Aerospace, Inc., and collaboration has continued under its auspices. In 1992, JRS retired from Stanford. So many experiences have been shared in this long collaboration that it seems appropriate to present these recollections under joint authorship.

Of the contributors to this monograph, JRS has the distinction of being the only one who was employed by the NACA (National Advisory Committee for Aeronautics) on October 1, 1958 when it was transformed into the NASA (National Aeronautics and Space Administration). Employment as an Aeronautical Research Engineer at Ames Aeronautical Laboratory at Moffett Field, California began on July 1, 1943, immediately after a B. Aero. Eng. degree was granted by the University of Minnesota and only three years after Ames was established.

As for the years at the University of Minnesota, there is one notable circumstance, beyond getting an excellent engineering education, that relates to later work on the magnetosphere. From 1940 to 1943 I was employed as a general handyman in the laboratory and machine shop of the Aeronautical Engineering Department. There I sometimes made simple instruments and other apparatus for Prof. Jean Piccard, the pioneer stratosphere balloonist, and occasionally prepared inked charts for his wife, also a balloonist, for her research in the social sciences. Although they were quite eccentric and not inclined to casual talk with students, I acquired an early awareness of conditions in the atmosphere, cosmic rays, and related matters of later concern in magnetospheric physics. As described by John Winckler in this monograph, the scientific use of unmanned high-altitude balloons he and his colleagues at the University of Minnesota began in the early fifties was a direct outgrowth of Piccard's work in the preceding decades.

My formative years before entering the university were in Staples, Minnesota, a town of about 2500 people about 140 miles to the northwest of Minneapolis, a little beyond the location of the better known, but fictitious , town of Lake Woebegone. It was a smoky railroad town on the main line of the Northern Pacific Railroad from St. Paul to Seattle with extensive shops to repair steam locomotives and freight cars. All trains stopped at Staples for crews to be changed and locomotives to be serviced or changed. I became well acquainted with the functioning and innards of locomotives, because it was customary, and tolerated by the railroad, for many living in my neighborhood to take a short cut to school and the stores of town by walking through the repair shops. Several times a day for many years I walked, or literally squeezed, past men wielding heavy hammers, crowbars, and wrenches on disassembled locomotives, often still hot and hissing from escaping steam. Others were working at big machines, lathes, grinding wheels, punch presses, metal planers, etc. mostly driven from a single electric motor by an array of whirring belts and pulleys. Many of the workers were the fathers of my friends, and they often stopped their work to show me what they were doing. It was natural in this environment that the local school provided extensive vocational training for the boys in drafting and shop skills in addition to the traditional academic courses, and I took them all. My father, who was a barber, also had a modest metal working shop that filled two rooms in our home where I acquired additional skills that contributed to my later employability at the university.

Mostly though, I liked to make flying model airplanes during the cold and dark winter, and to fly them in the warmer months. I was fascinated with anything that flew. There were no model airplanes kits then, so there was considerable experimentation. I learned much about airplane aerodynamics watching how they flew, sometimes well and sometimes not, when their design was modified in various ways. The thirties was a period of great enthusiasm about airplanes in Minnesota. Lindbergh, from the next town of Little Falls, had flown the Atlantic, was discovering Mayan ruins in tropical Mexico, and pioneering air routes from the USA to the Orient. (In 1939, he also chaired the NACA Special Survey Committee whose recommendation to establish a new aeronautical laboratory at Moffett Field, California led to the breaking of ground for Ames within the year.) Byrd had also flown the Atlantic and over the poles, and with Minnesotan, Lloyd Berkner, was reporting on the radio from their newly established bases in Antarctic. Others were making record-breaking long distance flights, and flying over previously inaccessible regions of the earth, all of which I read about avidly. An indication that others shared this enthusiasm is that there were 625 undergraduates majoring in aeronautical engineering at the University of Minnesota during my years there, i.e., 1939-43. This figure is even more remarkable considering there was no aircraft factory in Minnesota and perhaps only one or two jobs per year in the state for a newly graduated aero engineer.

Although I was the top student in my high school class and had scored in the 99th percentile on the college aptitude test, a problem arose when I applied for admission to study engineering at the state university. I had told everyone for years that I wanted to be an aeronautical engineer, but no one ever told me that engineers used higher mathematics! The only engineers I had ever met were locomotive engineers, and they certainly didn't use much mathematics. Instead of taking math in my last two years of high school, I was advised to take shorthand and typing to facilitate note taking and report writing, and I did for three years. As a result, I was admitted to study engineering on probation. Fortunately for me, there were many others, particularly from small towns, who were similarly deficient and the university offered non-credit courses enabling us to catch up. It is a tribute to this policy and to an intensive six-day per week university schedule that I was able to apply more advanced aerodynamic theory and mathematics to my early research at Ames than most of my contemporaries.

For my first three years at Ames, I was assigned as an Aeronautical Research Engineer to the Flight Research Branch. I was told that I had been assigned to this branch, rather than one of the more numerous positions in the wind tunnels, because of my summer job the year before in the Analytical Branch of the Naval Aircraft Factory in Philadelphia. I was immediately the branch "expert" in airplane design and stress analysis and assigned most of those tasks when aircraft changes, such as a novel horizontal tail or the first triangular dorsal fin of the type now in widespread use, were required for a test. Following design, my responsibilities continued through construction in the Ames shops, installation on the airplane, and finally the flight tests. The pilots were my office mates. Most of the changes were for single-seater fighter planes, but I frequently went along to operate the recorders if the airplane was a multi-seater. All this proceeded without the benefit of plans or design and stress analysis information from the manufacturers, and my analysis was never checked in detail by anyone. Naturally, I checked my calculations and assumptions as thoroughly as I could, but I was well aware that my knowledge was limited in many ways. It is an understatement to say that I worried a lot until some of the flights landed safely.

The Flight Research Branch was headed by Larry Clousing, a soft-spoken and fearless engineer test pilot with advanced degrees in both aero engineering and physics. Fighter planes were going out of control and breaking up in high-speed dives starting at very high altitudes. Clousing was determined that he and his group would learn why, and be able to recommend changes that would reduce or eliminate the hazard. On two occasions, he landed with his airplane so bent and fractured that it never flew again, even though one was a Bell P-39 specially reinforced by the manufacturers. The recorders on the P-39 showed he was pulling on the control stick with a force of 240 pounds to keep the airplane from turning upside down while at maximum speed in a vertical dive from about 42,000 feet, and then reversed this pull to a push of 180 pounds within a fraction of a second to restrain the plane in an abrupt pitch-up as it slowed naturally in response to the increasing drag caused by the greater air density at lower altitudes. Even so, the accelerometer went off-scale at 13.5 g. All this was done before there were g suits, pressurized cabins, or ejection seats for escape! In 1947, the Institute of Aeronautical Sciences awarded Clousing the Octave Chanute Award as the nation's outstanding test pilot, and in 1954 the University of Minnesota, his alma mater, bestowed upon him their Outstanding Achievement Award. It is a marvel he survived. When I asked him at a NACA reunion in 1988 why he was so intent on risking his life in these flights, and not even telling his wife what he was doing, he responded in his usual understated manner, "What risk? I just wanted to be the first to fly faster than the speed of sound.". As everyone knows, he was not in spite of all his efforts.

About this same time, another of our engineer test pilots, Jim Nissen, made two extraordinary high-speed dives from about 28,000 feet in a North American P-51B fighter plane with the propeller removed. On the third flight, he crash landed in a quarry after the cable attachment on the tow plane, a two-engine Northrop P-61A night fighter, broke shortly after takeoff. The P-51B was thoroughly ruined. Nissen escaped serious injury, almost certainly because he had decided before the flights to replace his customary soft leather helmet with one of the new hard plastic football helmets, and it was cut through across the forehead. He was rushed immediately to a nearby hospital for X-rays, but they could not be made because of an electric power outage. These flights were made from a dry lake bed about 12 miles across in the California desert. There were two hazards, a quarry near the center and an electric power line traversing the lake bed. Nissan cut the power line to the hospital on his descent and came to rest in the quarry. I had watched over the installation of iron ballast in the engine compartment of the P-51B to compensate for the weight of the missing propeller, and was relieved to learn that the part that broke was on the P-61A and had been installed by the manufacturer, and not me or my colleagues at Ames.

During this period, I got my an opportunity to learn a little more about the atmosphere. The dives were being made under visual flight rules because everything dispensable, including the radio, was removed from the airplanes to make room for the bulky instruments, and there were no chase planes. To plan the flights, it was necessary to forecast the time at which the frequent morning overcast would clear, hopefully to the quarter hour. The Navy at Moffett Field was already doing this quite successfully by simple analysis using radio-sonde data from weather balloons released each morning. They needed the forecast both for their extensive patrol operations using powered lighter-than-air craft called blimps, and to plan the launching of numerous manned helium-filled free balloons being done as part of a training program for future blimp flight crews. On the "strength" of having had a single undergraduate course in meteorology and a very slight acquaintance with balloons, I was assigned the task of participating in this activity and delivering the forecast for each overcast day before starting my normal aerodynamic work.

Although the propeller airplanes on which most of the research was carried out never exceeded a Mach number of 0.81 (about that of a cruising jet airliner) in an all-out dive from altitudes as high as 42,000 feet, it was evident that a new aerodynamic regime was being encountered in which the compressibility of the air was causing unanticipated changes in the forces on the airplane. We set out to learn about what was called "compressible flow", but had to do much of it ourselves because there were no books and only a handful of articles for guidance. We were not alone, however, because others at Ames were already conducting tests in a new and large (16 foot diameter) high subsonic speed wind tunnel when I arrived. It was, in fact, to answer a nagging question regarding how well the tests of a scale model in this wind tunnel could be used to determine the aerodynamic forces on an actual airplane in high-speed flight that Nissen proposed the dive tests of the propellerless P-51B. Uncertainties arose from two sources. One was the effects of the wind tunnels walls in the transonic speed range for which there was virtually no firm knowledge. The other was that electric motors of sufficient power to simulate the P-51B engine and small enough to fit inside the model were not available. Nissen's solution was to remove the propeller from a P-51B and compare the aerodynamic results from a high-speed dive with those measured in the wind tunnel. Nissen's two successful flights contributed greatly to establishing the validity of the wind tunnel tests.

Even more important, there was also a small elite group that constituted the Theoretical Aerodynamics Branch headed by H. Julian (Harvey) Allen, an imaginative and exuberant aeronautical research engineer raised in California and educated at Stanford who later became the Director of Ames in 1965. Much of their work was applied to the development of compressible flow theory, and its application to the aerodynamics of high-speed flight and the design of wind tunnels, both subsonic and supersonic. Cooperative interaction between those engaged in theoretical, flight, and wind-tunnel research was strong, we were young and tireless, and many of us worked far longer than required by our normal 48 hour work-week. Great progress was made in many directions. It is evident, in retrospect, that Ames had assembled and nurtured an outstanding aerodynamic research staff in its early years

To maintain this young and almost entirely male staff in the face of growing demands of the military draft in late 1943, arrangements were made in Washington that many of us would be inducted into the Navy, returned to the Naval Air Station at Moffett Field to constitute the Ames Detachment, and assigned as Apprentice Seamen to our civilian positions. In addition to carrying on our NACA work, we were given several hours per day of military training for six weeks, and then promoted. Those who had passed the physical exam perfectly became Ensigns, and later Lieutenants, j.g.,. Being near sighted, I became a Chief Petty Officers, a level normally attained by sailors only after many years of service. This arrangement resulted in some bizarre circumstances. For a while, the commanding officer of the Ames Detachment was a Navy Commander assigned to work as a junior engineer in a new Supersonic Wind Tunnel branch headed by Walter Vincenti, who was an Apprentice Seaman at first and then a Chief Petty Officer. If Vincenti wanted to leave Moffett Field for the night or Sunday, he needed permission from the commanding officer. If the commanding officer wanted the day off, he needed permission from his civilian supervisor, Vincenti. In spite of distractions such as Shore Patrol duty every few nights and working our full schedule during the day, the two-year period in the Ames Detachment was one of particularly intensive technical work for most of us. Life in the barracks was not particularly comfortable or stimulating. Many of us virtually lived day and night in our offices at Ames, taking time out from work for little more than sleeping, eating, and an occasional break for tennis or other recreation.

After being discharged from the Navy in 1946, I took leave from Ames for a year, enrolled in Stanford, and obtained a M.S. degree in Engineering Science in 1947. On return to Ames, my goal of being assigned to the Theoretical Aerodynamics Branch was achieved. By then, Harvey Allen had been promoted, and the branch was under the direction of Max Heaslet. Educated as a pure mathematician with a Stanford Ph.D., he had turned his attention to aerodynamic theory a few years before after a number of years on the mathematics faculties at Stanford and San Jose State. He was a very helpful person who by his own quiet example set high standards for others. I continued taking classes at Stanford, carefully using my leave by the hour, and was finally granted a Ph.D. degree in 1954. My research was on transonic flow past airfoils and wings. Although my real scientific mentor was Heaslet, my academic supervisor at Stanford was Irmgard Flügge-Lotz, a distinguished aerodynamicist and automatic control theorist who had just arrived from Germany. I was her first advisee, and she took great interest in my research and other activities. She and her husband Wilhelm Flügge, Professor of Applied Mechanics at Stanford, were very sociable, and I learned much from them, both scientific and otherwise. Several years after her death in 1974, Wilhelm and I wrote a short biography of her life [*Spreiter and Flügge*, 1987] for the book *Women in Mathematics*.

In 1951, I was appointed a Lecturer at Stanford and with Heaslet and Harvard Lomax, my contemporary at Ames and later long-time and highly honored Chief of the Computational Fluid Dynamics Branch, began teaching graduate courses in high-speed aerodynamics. Lomax, from Broken Bow, Nebraska, was another Stanford product with undergraduate and graduate degrees in engineering. Most of the students were research engineers from Ames seeking advanced degrees. Preparation of the lectures provided motivation to learn about topics such as hypersonic and rarefied gas aerodynamics that were outside the range of my personal research at the time, but were later to prove useful when attention turned to modeling solar wind flow past the magnetosphere. About a decade later, I introduced a new course in space and atmospheric physics at Stanford and continued to teach it, in ever enhanced form, until the end of my tenure in 1992. Joe Reagan, an author in this monograph, was a student in one of the early space physics classes, and later obtained his Stanford Ph.D. degree under my supervision. Steve Stahara was another early student.

A number of developments at Ames during the NACA period contributed directly to my ability to make the transition to magnetosphere modeling almost immediately after the beginning of the NASA era. First of all, there was the acquisition in 1951 of an IBM digital Card Program Calculator, assignment of it to the Theoretical Aerodynamics Branch, and the opportunity to be among the first to use it. Although it filled the space previously occupied by several offices, required extensive air conditioning, and received almost constant attention to replace constantly failing vacuum tubes, it's computing power was so feeble that it could barely keep up with one efficient human computer, Ruth Mossman, in calculating the rolling up of the trailing vortex wake of a wing. Experience with this machine sufficed, however, to begin a long tradition of excellence in computational fluid dynamics at Ames that still continues.

Another development in late 1951 that was to find a new use in our magnetosphere modeling a decade later was the recognition by Harvey Allen that the pointed nose of contemporary rockets was unsuitable for atmospheric re-entry after ballistic flight at speeds of 15,000 miles per hour or more. Simple square-cube considerations showed that the pointed tip would get too hot and melt or burn off. Allen's solution was to make the nose blunt. He argued that this would cause a larger fraction of the heat to be generated by a detached bow shock and carried away with the flow, and less to be produced at the surface by viscous skin friction. Little was known about highly supersonic flow past a blunt-nosed object, so he initiated a vigorous experimental program at Ames to provide knowledge and quantitative data needed for engineering design. Allen's ideas were described initially in a classified document, later in NACA TN 4047, and ultimately published [*Allen and Eggers*, 1958] in the last annual report of the NACA to the U. S. Congress. They were so completely accepted by 1958 that it is hard for us today to comprehend the doubt and skepticism with which they were greeted initially.

A couple of years later, my long-standing colleague, Milton Van Dyke, returned to the Theoretical Aerodynamics Branch after earning his Ph.D. degree at the California Institute of Technology and spending post-doctoral years there and at Cambridge University. (We met standing in line at the bank to cash our first pay checks from Ames in July 1943, worked in the same group at Ames and checked and edited each others work for more than a decade, later became professors at Stanford, and now share a university office in retirement. He also supervised SSS's Ph.D. research, which was on a subject, matched asymptotic expansions for boundary layer analysis, unrelated to the magnetosphere. Incidentally, Van Dyke was also born in Minnesota and took shorthand and typing instead of mathematics in his last two years of high school in Portales, New Mexico. He was admitted to Harvard where he enrolled in calculus without the preparation normally expected from high school, or make-up courses provided by the university.)

Van Dyke was determined to solve the "most important problem in aerodynamics", and everyone told him it was the blunt body problem. It was by no means obvious how this was to be done. Such a flow involves not only a bow shock detached from the nose of the body but also a region of subsonic flow embedded in the surrounding supersonic flow that would have to be represented by nonlinear partial differential equations of mixed elliptic-hyperbolic type in three spatial dimensions. No methods existed for solving such equations, and it was clear that solutions would have to be obtained numerically. This was just becoming feasible using a new IBM 650 computer that had been acquired for the Theoretical Aerodynamics Branch in 1955. *Van Dyke* [1958] solved this problem using novel mathematical principles and an effective iteration procedure to produce solutions of useful accuracy with an acceptable, but big, computational effort. In late 1958, the fledgling NASA was continuing the long-standing NACA practice of preparing a nicely bound book of selected research papers of the year as their first official report to the U. S. Congress, and it was decided that a companion paper [*Van Dyke and Gordon*, 1958] was to be NASA Report number 1. It was almost decided at an earlier stage in the deliberations that this paper would be Report number 2, and my paper [*Spreiter and Alksne*, 1958] on transonic aerodynamics would be number 1. On further discussion in my presence, the order was reversed because Van Dyke's topic was perceived, correctly, to be representative of the NASA future and mine to that of the NACA past. Little did I anticipate that Van Dyke's topic was to play an important role in my future work modeling solar wind flow past the magnetosphere. The next year Van Dyke left Ames to be a Professor at Stanford, and his attention turned to other applications. The work he started was continued at Ames by *Fuller*, [1961] and *Inouye and Lomax*,[*1962*] who developed the first of several generations of improved methods suitable for use with a succession of increasingly powerful computers at Ames. These developments were very familiar to me because we all worked closely together in the same branch and discussed our progress and difficulties almost constantly. I also served on the editorial committees that met with the authors and, following NACA traditions, scrutinized their work to a degree almost inconceivable today.

As documented more fully by Hartman [1970], the transition from NACA to NASA resulted in little immediate change at Ames. I recall asking the Director, Smith J. DeFrance, and other administrators what they thought we might do if we became part of the national space agency. The response was always essentially, "The same as we've always done.". I didn't quite believe that and began reading everything I could find relating to the earth, sun, solar system. I joined the American Geophysical Union in 1958, and read with enthusiasm the many fine articles appearing in the *Journal of Geophysical Research* and elsewhere. The International Geophysical Year (IGY) was in progress and much was being written about it. I was Chairman of the Ames Library Committee, funds to buy books and journals were plentiful, and my office was just outside the entrance to the library, so it was quite easy to acquire the literature to pursue these interests.

For me, a decisive moment came very shortly after the transition to NASA, and only a few days after I returned from a trip to Japan, my first trip abroad, where I had been invited to present a paper on transonic aerodynamics at the Japanese Congress of Applied Mechanics [*Spreiter*, 1959]. Two men from NASA Headquarters, John Clark, later Director of NASA Goddard Space Flight Center, and Gerhard Schilling came to Ames and spoke to a small gathering of perhaps 30 researchers to inform us what the new NASA was all about. Schilling spoke first and began rather dramatically, "The two principal questions confronting this agency today are the origin of the universe and the origin of life!". I didn't think I could add much to those subjects, but asked myself "Why am I worrying so much about the pressures on wings and bodies when I can be thinking about all these other things?". Heaslet, Allen, and others were encouraging, and I began immediately to seek topics to which I might be able to make some contribution.

EARLY YEARS OF THE DISCOVERY PERIOD AT NASA AMES: 1958-63

Aeronautical research continued at Ames after the change to NASA, but attention was turning to higher speeds for application to spacecraft design, particularly for re-entry into the atmosphere. At the same time, some of us began to consider the space environment itself. For example, Al Seiff began studying the properties of the upper atmosphere to support the re-entry aerodynamic studies that he later extended to apply to other planets. Mike Bader developed a probe based on mass spectrometer principles to measure plasma flows in space, and was later placed in charge of a new airborne observatory in a converted Convair 880 airliner. Responsibility for the plasma probes was assigned to John Wolfe, who served as Principal Investigator for the Ames plasma probes on numerous spacecraft including IMP-1, the first to explore the earth's magnetosphere in detail, and several Pioneers that flew to most of the other planets and beyond. Carr Neal designed an instrument using a stack of blackened razor blades as a blackbody sensor to be used in engineering studies of the thermal effects of different colored paint on a spacecraft that he soon developed into a sensitive instrument that could measure the temperature of whatever it looked down upon. In addition to what might be normally expected of such an instrument on a spacecraft, it was used to track the meandering of ocean currents such as the Gulf Stream in the Atlantic Ocean, and even to provide guidance to fisherman regarding where to find specific kinds of fish that tend to inhabit water in a narrow temperature range. I turned my attention to theoretical analysis of conditions anticipated in space.

My first topic was concerned with electrostatic fields in a model ionosphere in which the electrical conductivity becomes increasingly anisotropic with increasing height. Ben Briggs, another Theoretical Aerodynamics branch member interested in expanding his horizons, joined me in this work. Our analysis was a direct outgrowth of work by *Farley* [1959, 1960] that developed quantitatively the suggestion by *Martyn* [1955] that electrostatic coupling between the E and F_2 regions (about 100 and 300 km altitude) could explain the strong tidal flows in the F_2 region. In our study, [*Spreiter and Briggs*, 1961, 1962], the analysis was extended to greater heights, the upper boundary condition was altered, and numerical solutions of improved accuracy were obtained. The results showed that effective coupling extended to even smaller scales than indicated previously. In retrospect, the assumption that the electric current along the field lines vanishes at great heights appears unduly restrictive because it does not allow for possible effects of Birkeland currents. Along with many others at the time, we were influenced unduly by Chapman's extensive analyses of possible, but not necessarily unique, spherical-shell current systems in the ionosphere, see e.g., [*Chapman*, 1936, 1951, Chapman *and Bartels*, 1940, and *Mitra*, 1952].

Perhaps the greatest consequence of this work was that it enabled me to attend my first AGU meeting in Washington and soon thereafter the 1960 General Assembly of the International Union of Geophysics and Geodesy (IUGG) in Helsinki, Finland, and still another space plasma meeting in Copenhagen along the way. I presented our work at the IUGG meeting, and met many people from around the world whose names were becoming familiar through their scientific publications. These included Chapman, Ferraro, and Troitskaya, whose names appear prominently throughout this monograph. All this was in marked contrast to our experiences in the NACA years when the Ames travel budget was very meager, travel to any meeting was rare, and foreign travel was virtually a once in a lifetime event. In his history of Ames, *Hartman*, [1968] notes that only eight Ames employees presented papers at meetings abroad during the 18 years of the NACA era, and none did so twice. Our work also led to my appointment to the NASA Radio and Ionospheric Physics Subcommittee, although my knowledge of the subject was obviously minimal at the start. I learned much during several years of service on this committee, both from the discussions at our meetings and from informal conversations with the other members, particularly John Clark, Bob Helliwell, and Francis Johnson.

The next topic to which I turned was the shape of the boundary, now called the magnetopause, of the earth's magnetic field in the solar wind. Since no spacecraft had yet gone far enough from the earth to observe the magnetopause, normally a minimum of about 10 earth radii, the subject was rather speculative. A leading possibility proposed long before *by Chapman and Ferraro*, [1931, 1932, 1933], (see also, [*Chapman and Bartels*, 1939; and *Chapman*, 1963]) to explain certain features of a geomagnetic storm was that the earth's dipole field is temporarily confined in a limited region of space by a passing cloud or rotating beam of collisionless charged particles, protons and electrons, and electrically neutral in overall charge density except in certain thin regions such as the magnetopause. To describe what might occur, they carried out extensive theoretical studies that produced, among many other things, a plot of the combined field of the geomagnetic dipole and the electric currents induced in the planar surface layer of an advancing neutral ionized cloud. Although one of the most reproduced plots in space plasma physics, it is incorrect! They state the correct value of $\sqrt{2}$ for the ratio of the distance between the two neutral points where the magnetic field vanishes on the planar front to the distance of the front from the dipole, but in their plot that ratio is doubled. The only place I know where this figure is drawn correctly is in *Dungey* [1958], and even he seemed puzzled when I mentioned this difference in our historical session at the 1995 AGU meeting.

A significant question regarding the interaction at that time was the nature of the solar wind itself. Chapman and Ferraro had regarded the solar plasma flow as intermittent, either isolated clouds passing through the otherwise vacuum of interplanetary space or beams rotating with the sun that swept past the earth. That the flow is a continuous, rather than intermittent, high-speed radial outflow of ionized particles from the sun had been proposed by *Biermann* [1951, 1953] to explain the behavior of comet tails. On the other hand, *Chapman* [1957] had just presented an alternative view in which he assumed that interplanetary space is filled with stationary plasma having high thermal conductivity and concluded "The earth is a cool speck in the hot extended atmosphere of the sun.". *Parker* [1958] quickly pointed out that Chapman's static model indicated impossibly high pressures at great distances from the sun. He proceeded to present his own theory based on equations for the steady radial outflow of a continuum gas including the effects of solar gravity and extended heating through the use of polytropic pressure-density relation of the form p/ρ^n, where n has a value less than the adiabatic ratio γ of specific heats at constant pressure and volume, i.e., $\gamma = c_p/c_v$. Of the possible solutions, he chose the one indicating acceleration of the solar plasma from low subsonic speeds near the sun to supersonic speeds of several hundred kilometers per second at greater distances, in keeping with Biermann's proposal. This choice was challenged vigorously by *Chamberlain*, [1961] who argued that other entirely subsonic solutions indicating flows of tens of kilometers per second past the earth were the appropriate ones. He called this slow flow the "solar breeze" in contrast to Parker's "solar wind". To confuse the matter further, the same equations and solutions, but with reversed sign for the velocity, had been used shortly before by *Bondi* [1953] and *McCrea* [1956] to describe inflow to a stationary star of interstellar gas at rest at infinity. To the few of us at Ames who were aware of it, the acrimonious dispute between Parker and Chamberlain seemed amusing because their solutions were similar to those for flow of a compressible gas through a converging-diverging pipe, often called a de Laval nozzle, that had been used extensively in the design of both subsonic and supersonic wind tunnels, as was in fact pointed out by *Clauser* [1960]. If the chronology appears somewhat out of order here, it reflects the usual time delay between when ideas were spoken about at meetings and when they were published in refereed archival journals.

By 1960, Chapman and Ferraro's concept that the pressure of a flux of solar particles impinging on the magnetopause is balanced by the magnetic pressure of the geomagnetic field confined therein was generally accepted. The pressure is given at each point on the magnetopause by $K\rho_\infty v_\infty^2\cos^2\psi$, in which ρ_∞ and v_∞ are the density and velocity of the undisturbed solar wind, ψ is the angle between its direction of flow and the local normal to magnetopause, and K is a constant. The most appropriate value was uncertain, (compare statements in *Spreiter and Briggs*[1961, 1962a, 1962b), but K was usually equated to 1 or 2, although 1/2 was sometimes used. The balancing magnetic pressure is given by $B_s^2/8\pi$, where B_s is the intensity of the confined geomagnetic field at the same point. It came as a surprise, however, to learn that the numerous drawings of the shape of the magnetopause appearing at the time were merely sketches, and not accurate graphs of an actual solution. The first actual solution for a model of this type was that of *Zhigulev and Romishevskii*, [1959, 1960], but it was for a two-dimensional, rather than three-dimensional, dipole.

A decisive step was made by *Beard* [1960] who presented an analysis in which the full three-dimensional character of the magnetopause is retained, and simplification is achieved by approximating one of the boundary conditions. This consisted of supplementing the condition that the normal component of $\boldsymbol{B}$ vanishes at the magnetopause with the condition that B_s is twice the tangential component of the earth's dipole field at that point, as it would be if the curvature of the magnetopause could be disregarded at each point. Chapman and Ferraro had previously used this approximation in a more limited way for the nose and along the equatorial trace of the magnetopause, but Beard applied it to the entire magnetopause. This simplification enabled him to obtain a partial differential equation for the radial coordinate of the boundary in terms of spherical angular coordinates θ and ϕ in which the dipole singularity is at the origin. He restricted attention to the case in which the dipole axis is at right angles to the oncoming flow and presented results for the location of the magnetopause in the equatorial and noon-midnight planes.

Upon study, it became evident that the results for the noon-midnight plane were incorrect because of a sign error in the analysis. I was familiar with the type of simplification Beard used, and knew that it provided a good approximation in other applications. I therefore set out to obtain the correct solution of his equation for the noon-meridian plane, and was again joined by Briggs. Figure 1 reproduces results from *Spreiter and Briggs* [*1961a, 1962*]. The most notable difference with Beard's results concerns the location of the neutral points where the magnetic field vanishes in the Chapman-Ferraro model. Our solution indicated they are sunward of the poles, whereas Beard's results indicated they are anti-sunward of the poles.

Our first submission of these results to the *Journal of Geophysical Research* was rejected because the referee insisted, without citing any proof, that the neutral points could not be sunward of the pole. The accepted version cited above included additional results for the noon-midnight trace of the magnetopause for various angles between the earth's dipole axis and the incident stream. They indicated that the magnetopause would rotate like a windsock to follow the direction of the solar wind when presented in a coordinate system in which the dipole direction is fixed. When the same results are presented in a coordinate system aligned with the solar wind as was done in later reviews [*Spreiter, Alksne and Summers,* 1968; *Spreiter and Summers*, 1969] and illustrated here in Figure 2, they show that the magnetopause location is almost independent of the dipole tilt. Evidence that the results obtained using the Beard approximation would be of useful accuracy was also provided by determining the corresponding approximate results for a two-dimensional dipole and comparing them with the newly available exact solutions of *Zhigulev and Romishevskii* [1959, 1960]. The latter were in turn confirmed by comparison with the results of independent analyses of *Hurley* [1961] and *Dungey* [1961], after correction of an error in the calculation of the x coordinate of the boundary in the latter.

Our next step was to complete the solution of Beard's differential equation to provide the coordinates for the entire three-dimensional surface of the magnetopause for the case in which the dipole axis is perpendicular to the direction of the incident solar wind [*Briggs and* Spreiter, 1963]. These are illustrated in Figure 2. About the same time, *Midgley and Davis*, [1963] and *Mead and Beard*,,[1964] presented their results for the same orientation obtained without using

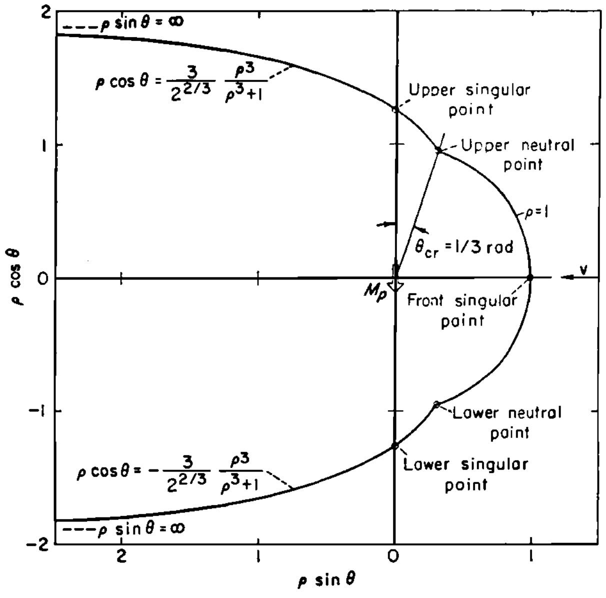


Figure 1. Form of the magnetopause in the noon-midnight plane when the dipole axis is perpendicular to the direction of the solar wind, [*Spreiter and Briggs* [*1961, 1962*].

Beard's approximation. They termed their results "higher order", although convergence of the harmonic expansion in terms of singularities at the dipole center used to represent effects of the electrical currents in the magnetopause is questionable when applied to points at or near the magnetopause. In any case, the differences between the three sets of results are small, (see e.g., *Spreiter, Alksne, and Summer*, [1968] or *Spreiter and Alksne*, [1969]), and probably not very significant considering all the approximations in the underlying C-F model and the uncertainties in the values selected to represent conditions in the solar wind. More accurate solutions of *Olson,* [1969] published near the end of the discovery period and numerous others in the years following have confirmed the same trend. Because of these and other reasons including simplicity, we have continued to approximate the shape of the earth's magnetopause in most of our subsequent modeling with a body of revolution obtained by rotating the equatorial trace of the magnetopause about an axis aligned with the direction of the incident solar wind flow and passing through the dipole. In this, we have used the coordinates for the equatorial trace given by *Spreiter and Briggs*, [1961], which differ only slightly from the earlier but slightly less accurate results of *Beard*, [1960].

By the end of 1961, our modeling efforts were progressing well and interesting others in the Theoretical Aerodynamics Branch at Ames. Most notable among these was Alberta Alksne, with whom I had collaborated in theoretical studies in supersonic and transonic aerodynamics since 1947.

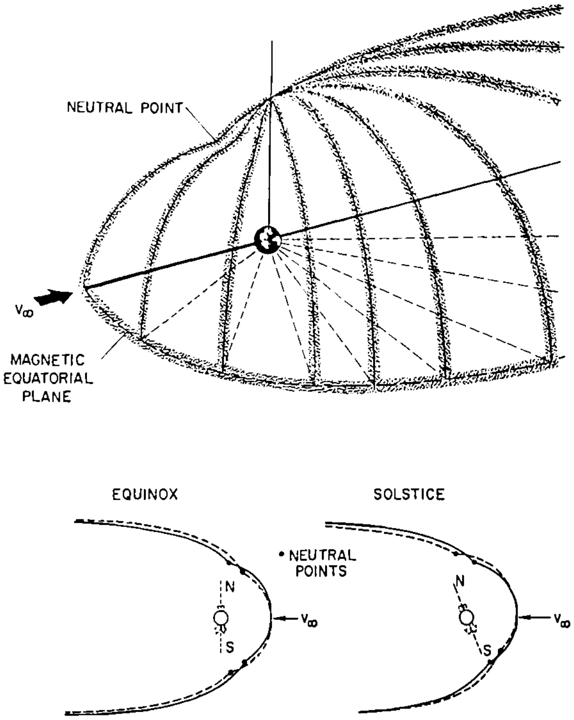


Figure 2. The magnetopause boundary of the geomagnetic field and its diurnal and season variations, as viewed in a coordinate system aligned with the flow direction of the solar wind.

Raised in rural California near the town of Hollister and educated in mathematics and physics at Stanford, she started employment at Ames during the wartime period as a computer, when they were persons, usually young women, and not machines. She was an energetic and adventurous person with widespread interests. Years of active participation in the local astronomy club in Palo Alto provided useful background for our space research. She had spent weekends flying around the deserts of California prospecting for uranium with her aeronautical engineer pilot husband, until he was killed when his airplane crashed into a canyon wall. After retiring from Ames a decade later, she joined the Peace Corps and spent two years in Kenya teaching mathematics at a high school for girls. In 1961, however, we were ready to taper down our remaining work on transonic aerodynamics, and focus attention on the magnetosphere and solar wind. Jeanne Hyett, a younger member of the branch who worked with us on transonic aerodynamics, also elected to make a similar change about the same time, as did Bill Jones.

Our work progressed rapidly to incorporate features not included in the previous analysis, but that were being discussed at meetings and in the rapidly growing literature about conditions in space. Some of these, such as a bow shock discussed at greater length in the next section, have found a permanent place in magnetospheric physics. Others, such as the analysis of the unsteady effects of the passage of large-scale changes in the solar wind [*Spreiter and Summers*, 1965a] and the stability of the magnetopause [*Spreiter and Summers*, 1965b], provided useful insights to us, but did not seem to be noticed much by others. Still others are little remembered or missed, even though the solutions are accurate mathematically, because they were based on assumptions not supported by the new data. Examples include the effects of an equatorial ring current, unduly strong as was later recognized, on the shape of the magnetopause [*Spreiter and Alksne*, 1962] and on VLF whistlers [*Spreiter and Briggs*, 1962c]. We undertook these analyses after hearing an enthusiastic presentation in which a ring current having a dipole moment of magnitude comparable with that of the main dipole field of the earth was proposed to explain Explorer 6 and Pioneer 5 observations on the night side at a geocentric distance of about 6 earth radii, [*Smith, et al.*, 1960]. Unfortunately, their interpretation was incorrect because only two of the three components of the magnetic field were measured and the true character of the missing component was not realized. The low apparent magnitude of the observed field was not the result of the proposed ring current but was due to the spacecraft being near the earthward edge of the tail current sheet, whose existence at that time was, of course, was still to be recognized.

Another extension that experienced a short lifetime, although its main novel feature was used later in other applications, was that of *Spreiter and Hyett*, [1963] and *Spreiter and Summers*, [1963] in which the C-F relation for the pressure of the solar wind on the magnetopause was modified by adding a uniform static pressure p_∞ so that the pressure is given by $p_\infty+K\rho_\infty v_\infty^2\cos^2\psi$. This relation already had a established place in hypersonic aerodynamics and was referred to as the Newtonian approximation, although it did not actually originate with Newton. It is a source of some confusion that this relation is frequently referred to in space physics as the modified Newtonian pressure formula, and the C-F formula without p_∞ is referred to as the Newtonian pressure formula. Our application generalized an earlier analysis by *Slutz*, [1962] in which the shape of a magnetopause under the influence of static pressure of a stationary plasma was shown to resemble an apple. Such an assumption would be appropriate if Chapman's static corona model in which $v_\infty = 0$ were correct, but that model never gained ascendancy as described above. We developed our analysis for two other reasons. One was to provide another check on the usefulness of the Beard approximation for the strength of the magnetic field at the magnetopause, and that was accomplished when our shape for the magnetopause was found to agree satisfactorily with that of Slutz. The other was to show what the shape of the magnetopause would be if the solar plasma flow were much slower than indicated by Parker's solar wind theory, as was being argued by Chamberlain. *Slutz and Winkelman* [1964] arrived independently at their results for the same conditions, and comparisons again confirmed the usefulness of the Beard approximation. In these analyses, it was assumed that the nightside geomagnetic field would be confined within a limited region of space, and that this would be trailed by a long wake filled with quiescent plasma exerting a constant pressure on the boundary with the surrounding flowing solar plasma. We had not anticipated the extended magnetic tail of the magnetosphere soon to be discovered from the data from IMP-1 (Explorer 18) for the first half of 1964 [*Ness*, 1965]. Not all was wasted effort, however. We were later to make use of the augmented pressure formula to show how the indentations in the magnetopause near the neutral points would transform into cusps projecting into the magnetosphere [*Spreiter and Summers*, 1967], and also in the analysis of conditions along the extended geomagnetic tail [*Spreiter and Alksne*, 1969a].

Numerous checks on the accuracy of the calculations and the internal consistency of the theory were made in this way, but the final test of any model of the magnetopause must be based on comparisons with observations actually made in space. As discussed in greater detail by Heppner and Cahill in this monograph, data from Explorers 10 and 12 confirmed that the geomagnetic field terminated abruptly in accordance with the C-F theory, and at distances from the earth compatible with the predictions, considering the uncertainties still remaining because other conditions in the solar wind were still unmeasured. By the end of 1962, data were available from a number of spacecraft, computers and methods for numerical analysis had improved dramatically from only a few years before, and progress in modeling could proceed with increasing confidence that the results would be in useful accordance with observations.

The biggest influence of these comparisons on our modeling came from the observation that the magnetic field fluctuates irregularly in magnitude and direction throughout a region several earth radii thick beyond the magnetopause. Only at even greater distances from the earth did the magnetic field become steadier and smaller, as it was anticipated to be in the solar wind away from the earth's influence. The fluctuating region was called the transition region, but before long it came to be known as the magnetosheath.

These observations prompted a reformulation of the model for solar wind flow past the magnetosphere in terms of supersonic flow of a continuum gas instead of the C-F corpuscular flow model. As developed in increasing detail by *Axford* [1962], *Kellogg*[1962], and *Spreiter and Jones*,[1963], the magnetopause would still exist in essentially the same location as before, but the explanation for the pressure on it would change from the C-F particle impact model to the nearly identical Newtonian approximation of hypersonic flow theory. The distant boundary marking the end of the region influenced by the earth was identified as a bow shock standing upstream of the magnetopause and terminating the magnetosheath. We could now calculate the location of the bow shock and conditions in the magnetosheath using methods developed at Ames for supersonic flow past a blunt body by Van Dyke, Fuller, Inouye, and Lomax. This was a welcome change for me and my colleagues, all of whom had formerly worked in aerodynamics. We thought we were leaving continuum gasdynamics when we began our space studies. Now we were back in more familiar territory with tested computational methods ready to apply.

The change was surprising to many at the time, however, because a fundamental condition for using continuum concepts is that the mean free path of the particles between collisions is small compared with the dimensions of the obstacle, a condition that the solar wind with a density of only a few particles per cubic centimeter at the orbit of earth failed grossly to satisfy. This condition is appropriate for a normal gas composed of neutral particles that interact only through direct collisions. It does not necessarily apply to solar wind plasma, however, because it is completely ionized and particles can interact at much longer range through the action of electric and magnetic forces. The precise conditions required for continuum-like behavior of such a plasma remain uncertain, but we now know that continuum gasdynamic models can be applied successfully to not only solar wind flow past the earth's magnetosphere, but even to the approximately 1600 times less dense solar wind flow around Neptune, [*Spreiter and Stahara*, 1994].

Figure 3 illustrates results from *Spreiter and Jones*, [1963] for the calculated location of the bow shock for an axisymmetric obstacle having the shape calculated for the equatorial trace of the magnetopause as described above. There were two reasons why the magnetopause was approximated in this way. One was that our results [*Briggs and Spreiter*, 1963] indicated that the full three-dimensional shape of the magnetopause is, in fact, nearly axisymmetric. The other was that methods for calculating supersonic flow past a blunt body had been developed only for axisymmetric and two-dimensional flows, and that the computing requirements to determine the solution for a more general three-dimensional obstacles would greatly exceed the capabilities of the most powerful computer at Ames, now an IBM 7094. The location of the bow shock is shown on Figure 3 for two different values for $\gamma = c_p/c_v$, the ratio of specific heats at constant pressure and volume. In the reference cited above, results are presented only for $\gamma = 2$, a value thought by many at the time to be representative of conditions in the collisionless plasma. The corresponding results for another leading possibility, $\gamma = 5/3$, were made available in *IG Bulletin 84*, [1964] and are included in Figure 3 because they were soon found to agree better with extensive plasma and magnetic field data from IMP-1, also known as Explorer 18, launched on November 27, 1963.

The fluctuating character of the flow observed downstream of the bow shock, not accounted for by the steady flow downstream of an aerodynamic bow shock, was explained by presenting results for one-dimensional collisionless shock waves calculated by *Auer, Hurwitz and Kilb*, [1961,

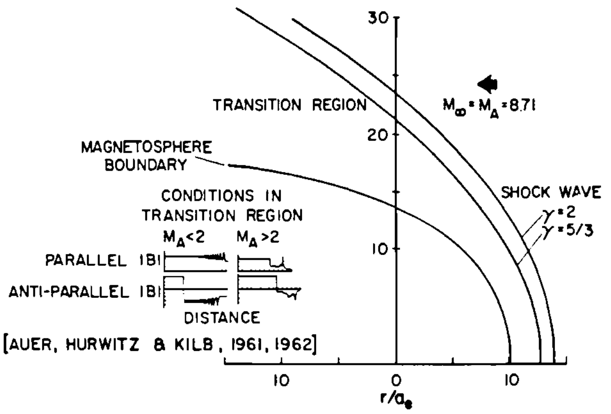


Figure 3. Calculated shape of the bow shock and magnetopause for representative conditions in the solar wind together with small inserts illustrating the nature of the magnetic field in the magnetosheath indicated by collisionless shock solutions of *Auer, Hurwitz, and Kilb*, [1961, 1962]. The results for the ratio of specific heats $\gamma = 2$ are from *Spreiter and Jones*, [1963]. Those for $\gamma = 5/3$ were made available shortly thereafter in *IG Bulletin* ***84***, [1964] because they were found to be in better agreement with the plasma and magnetic field data from IMP-1, or Explorer 18, spacecraft launched on November 27, 1963.

1962]. As illustrated by the small plots on Figure 3, they show that the magnetic field downstream of such a shock is extremely variable. These results indicate the fluctuations are greater when the magnetic field is anti-parallel rather than parallel to that in the obstacle. Applied to the solar wind flow past the magnetosphere, this corresponds to the fluctuations being greater when the interplanetary magnetic field is southward than when it is northward. This result is now so well established by observations that it is virtually a cornerstone in understanding the geomagnetic consequences of the direction of the interplanetary magnetic field.

To increase understanding of collisionless shocks, Jones and Vernon Rossow, another member of the Theoretical Aerodynamics Branch, initiated further calculations. The results [*Jones and Rossow*, 1965] and [*Rossow*, 1965, 1967] confirmed those of *Auer, Hurwitz and Kilb*, [1961, 1962], and demonstrated that the general conclusions regarding fluctuations applied throughout a much wider range of conditions than shown previously. Rossow, from rural Iowa near the town of Correctionville and educated at Iowa State, Stanford, and Zurich, carried out a wide range of research over five decades at Ames. Although his venture into space science was fairly brief, experiments he conducted using transparent plastic channels filled with liquid sodium with imposed magnetic fields provided us with valuable visual insights into the properties of MHD flows and waves.

Two events occurred in 1962 that were to have a major impact on our subsequent work. One was the opportunity for me to attend a course entitled *Geophysique Exterieure* at the Les Houches Summer School of Theoretical Physics. This was a great experience. For two months, we lived, rather uncomfortably, in unheated converted farm buildings high above the town of Les Houches near Chamonix in the French Alps. From our single classroom in a converted granary, we could look down upon glaciers. There were about 20 students from 9 countries, usually two per country including the USSR and Poland of the Eastern bloc. Nearly all were already active in space research. We had many interesting discussions and learned much from each other. Formal instruction at an advanced level was provided by ten of the world's most distinguished experts, including Chapman, Chamberlain, and Dungey whose names appear frequently throughout this monograph. Another was Marcel Nicolet, who had been General Secretary of the recently concluded International Geophysical Year, or IGY, of which Chapman had been President. Written versions of all the lectures were published promptly in *DeWitt, Hieblot, and Lebeau*, [1963]. Chapman invited me to give two of his lectures to describe our work on the magnetosphere, and to write them up for inclusion in the written version of his lectures, [*Chapman*, 1963]. Attendance at this school introduced me not only to leading experts, but also to a number of students who were to gain prominence for their research and leadership in space science. I have enjoyed watching their development, and talking with them about many things when our paths cross occasionally.

The other event was the establishment of a Space Sciences division at Ames under the direction of Charles Sonnett. Raised in California and educated at UCLA, he was a pioneer in measuring magnetic fields in space. He came to Ames from NASA Headquarters in Washington where he oversaw many phases of the early planetary and space program. With no prior knowledge of this development, I was summoned to the Director's office within half an hour of reporting to work on my return from Europe to be informed that my new assignment was to be Chief of the Theoretical Studies Branch in this new division. That was the way Ames worked. People were frequently assigned to a new position without either being consulted or informed of the possibility. Sometimes a person's first knowledge of such a change was from the general announcement that went out to the entire staff. I had mixed feelings about this change. I was very happy with conditions in the Theoretical Aerodynamics Branch directed by Heaslet. Sonnett had not arrived from Washington yet, and I could only guess at what he expected of me and the branch I was to create. On the other hand, it was clear that the new setting would provide a more appropriate place for our magnetospheric modeling, and to expand into other topics of interest in space science.

There was an immediate realignment of the positions of my associates. Alksne was rather obviously assigned to the new branch, but Briggs, Jones, Hyett, and Rossow were to remain in the Theoretical Aerodynamics Branch. Audrey

Summers, a highly regarded computational specialist in one of the wind tunnel groups, requested to join us, and that was approved. She grew up on a nearby farm, studied mathematics and physics at Stanford, and had been employed at Ames for about 15 years at the time. As evidenced by the citations referred to below, she quickly became a effective member of our small team modeling the magnetosphere. Another to be assigned almost immediately to the new branch was Ray Reynolds, who had been hired by Ames a few months before because of his previous theoretical research on the interiors of the giant outer planets and in the management's anticipation of the formation of the new division and branch.

LATER YEARS OF THE DISCOVERY PERIOD AT NASA AMES: 1963-1969

By the beginning of 1963, Ames had changed in many ways from the closely knit and highly focused aeronautical research laboratory that it had been under the NACA. Major emphasis was still on aerodynamics, for which its array of wind tunnels, computers, and other facilities remained uniquely suited. Attention was increasingly turning, however, from airplanes to spacecraft and the problems encountered in their design and operation. Even bigger changes were stemming from the establishment of the Space Sciences Division, and the Life Sciences Directorate devoted to research in fundamental life sciences extending all the way to the origin of life itself. Ames had become a much more diverse and exciting place, and its activities were gaining public attention much more than ever before.

Modeling solar wind flow around the magnetosphere was now in the mainstream of the work of the branch and division in which I worked, but there were many new duties. First of all, there was the matter of establishing the role and recruiting people for the branch, and assisting Sonnett in creating the new division. Much time was spent defining a broad research program in planetary, lunar, and space research, recruiting prospective research leaders, and creating an atmosphere conducive to scientific research. I was assigned the task of organizing a weekly Space Sciences Seminar that convened throughout most of the year, and did that for the remaining six years I was at Ames. The opportunity to invite visitors to speak at these seminars enabled us to expand our knowledge greatly and also to become acquainted with a much wider circle of active contributors to space science than we would have otherwise.

The first addition to the Theoretical Studies Branch who was not previously employed at Ames was Peter Fricker, a Swiss geologist who had held a postdoctoral position at Stanford and was particularly interested in what was still called "continental drift", a highly controversial subject at the time. Before joining Ames, he had climbed without oxygen to over 23,000 feet in the Andes, and survived for a month with a small number of scientific colleagues after their airplane crashed on an island in the far north of Canada. In our branch, he joined with Reynolds and Summers to model the internal composition and thermal history of the moon and planets. After several years, he returned to Switzerland for family reasons, later became the General Secretary of the Swiss National Science Foundation, and now holds a similar position in the European Science Foundation. Joan Hirshberg, who now uses her maiden name, Feynman, joined our branch about the same time to study the connections between observations in space and on the ground, and their implications, see e.g., *Hirshberg*, [1968] for a representative example of her work during this period. We wanted her to accept a regular appointment, but she wished to be committed to somewhat less than the required forty hour week because she had two preschool children. This could only be reconciled by employing her as a consultant. She remained in this capacity until 1969 when arrangements were made for her to carry on her work as a research associate with me at Stanford. She is now at the Jet Propulsion Laboratory at Pasadena, after several years at the NCAR in Boulder and the NSF in Washington. She was raised on Long Island, New York, spent several years living in villages in rural Guatemala assisting her anthropologist husband, and had resumed scientific studies on geomagnetism at Columbia University's Lamont a few years before joining us. David Webster also joined us about this time, officially as a consultant because he was several years beyond the mandatory retirement age for government employees. Like Hirshberg, he entered enthusiastically into all phases of our work, and we all benefited much from his presence. He had no prior experience in space physics, as very few did at that time, but he had a distinguished career as a classical physicist and educator, with particular strength in electricity and magnetism. He came to Stanford in 1921 to be Chairman of the Physics Department where, among many accomplishments, he supervised the work of the Varian brothers in developing the klystron, a high-powered traveling wave tube used in long-range radar and high-energy particle accelerators. Webster also constructed and flew his own biplane glider in 1911, was a test pilot flying French SPADs and other airplanes during World War I, and according to an article in *Physics Today* was remembered by other physicists "for the precision of his test flights and the wild rides he gave his passengers after the tests were completed". I recall that he was very worried when his grandson made a full-scale replica of his early glider and flew it on the hills behind Stanford. Severe restrictions on hiring set in after about a year, but our branch continued to grow slowly with the assistance of a postdoctoral research associate program administered by the National Research Council. Three, in addition to Fricker, who came under this program and later became regular government employees are still with NASA. They are Larry Caroff, Aaron Barnes, and Pat Cassen. Caroff, from Swarthmore and Cornell, was interested in infra-red astronomy and now manages major programs in NASA Headquarters in Washington. Barnes and Cassen are still at Ames and both have served in my former position as branch chief. Barnes, who studied the damping of Alfvén waves in the solar wind in the physics department at Chicago, has become renowned for his continuing studies of the solar wind. For many years, he has been responsible for the Ames plasma probes on Pioneer spacecraft to Venus, the outer planets, and beyond. Cassen studied the geomagnetic tail in the aero department at Michigan. Cassen's advisor was the eminent aerodynamicist, Arnold Kuethe, but Sydney Chapman also took an interest in his research during his annual month-long visits to Ann Arbor. When I inquired about Cassen, Kuethe recommended him highly, of course, and then surprised me by saying that he knew about my home town of Staples because his wife was from there. Cassen's attention has turned to the application of fluid concepts to modeling the formation of stars and galaxies, but a notable achievement along the way done jointly with Reynolds and others was a soundly based prediction of active volcanoes on Jupiter's moon Io. Only a few days after their prediction appeared in *Science*, Voyager spacecraft passed by Io and photographed a huge eruption in progress. Their paper was selected as the outstanding paper of the year in *Science*. Figure 4 is a photograph showing everyone mentioned above at the time I left Ames in late 1968. Two others in the photograph are Darlene Moen, our secretary who was from a farm near Grand Forks, North Dakota. She later became Sonnett's secretary, and then left Ames to be the person in charge of major traveling NASA exhibits. The remaining person, Mae Liu, had just been assigned to the branch as a mathematician employed by an outside contractor. She assisted several of the people for a number of years, but none of us know what happened to her after she left Ames. Alksne, Moen, and Webster are now deceased, and Summers is very ill.

There were also a number of shorter-term visitors. Dungey and his student, Fairfield, spent a couple of summers with us. Several others came for one to three years under the NRC research associate program, as did Peter Fricker originally. Among the first were two I met at the Les Houches Summer School. They were Barbara Abraham-Shrauner from Harvard and Michael Rycroft from Cambridge. Abraham-Shrauner joined in our solar wind magnetosphere interaction modeling, and moved on to become a professor at St. Louis University. More will be said of her work later. Rycroft worked with John Thomas, the new Assistant Division Chief, and Larry Colin of the Electrodynamics branch, who had come from Helliwell's group at Stanford at about the same time, on the analysis and interpretation of extensive data on the upper ionosphere from the very successful Canadian top-side sounder, Allouette. After his stay at Ames, Rycroft returned to England where he has played a leadership role in university and government science, teaching, and administration. He is now a professor at Strasbourg, France. Thomas, who had been one of Rycroft's tutors at Cambridge, returned to Imperial College in England after several years at Ames. Colin later became the project scientist for the long-lived Pioneer Venus Orbiter and remained at Ames until he retired. Others who investigated space plasmas in our branch include Dick Hartle from Penn State who continued his studies of the solar wind at Goddard, Jim McKenzie from Cambridge who carried on his studies of space and astrophysical plasmas as a professor in South Africa and researcher at the Max Planck Institute in Lindau, Karl Westphal from Germany who collaborated with McKenzie, and Clive Marsh from Leeds University where he studied what happens in the surrounding gas when a star "turns on". At Ames, he joined us in the calculation of MHD solutions for solar wind flow past the moon, [*Spreiter, Marsh, and Summers*, 1970]. Still another who participated in the work of the branch as a consultant was Peter Sturrock, a plasma physicist and Stanford professor, who collaborated primarily with Hartle in the development of a two-fluid model of the solar wind, [*Sturrock and Hartle*, 1966] and with me [*Sturrock and Spreiter*, 1965] on forward and reverse shocks observed in the solar wind and their relation to geomagnetic storms. From this beginning, Sturrock developed a distinguished career in solar physics and astrophysics. Two others at Ames who contributed significantly over a long period through analysis of data and occasional modeling are David Colburn and John Mihalov of the Electrodynamics branch. Colburn had a long collaboration with Sonnett studying magnetic fields in space. Mihalov analyzed plasma probe data from a succession of spacecraft in collaboration with John Wolfe at first and Aaron Barnes later. Mihalov still works at Ames, but Colburn has retired.

With respect to our modeling, the extensive magnetometer and plasma data from IMP-1 becoming available, see e.g., [*Ness et al.*, 1964 and *Wolfe et al.*, 1966], were providing conclusive evidence confirming the permanent presence of a magnetopause and bow shock in locations compatible with the calculations. Precise comparisons were not possible for two reasons, however. One was that the observations were by a single spacecraft, so that conditions in the solar wind upstream of the bow shock could not be determined simultaneously with the observations in the magnetosheath or magnetosphere. Reasonable estimates could be made, however, using non-simultaneous IMP-1 data from the part of its orbit beyond the bow shock. The second reason is that solutions had been determined for only a few Mach numbers and values for γ. Each solution required at least a month to determine using a mixture of machine and hand calculations. They had been determined, moreover, to illustrate what the new fluid model would indicate rather than to provide quantitative predictions for all possible conditions in the solar wind. The situation was not too bad, however, because simple scaling laws to account for changes in the density and velocity of the solar wind were known, and the results were not highly dependent on Mach number or γ in the range representative of the solar wind.

about wings and bodies, his was MHD. His lectures and informal discussions with many Japanese, particularly Isao Imai who worked in both compressible flow theory and MHD, as we traveled around Japan for a couple of weeks gave me further insights into MHD. I had no premonition that knowledge of it would ever be useful to me, however.

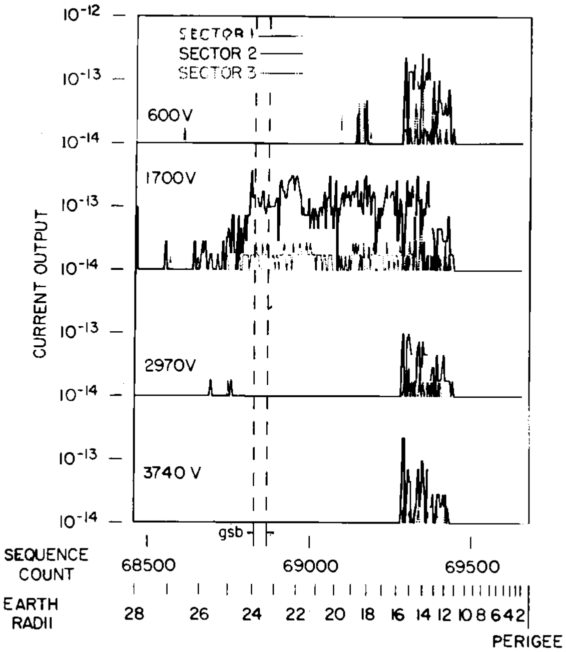
To be both convincing and instructive, we decided to present the case for using MHD in an extended semi-tutorial paper, [*Spreiter, Summers, and Alksne*, 1966a], that (a) reviewed relevant features of ideal steady state MHD theory including shocks and other discontinuities. (b) showed how the former results could be recovered from the MHD equations through introduction of acceptable approximations. (c) introduced simplifications in the MHD model to produce a computationally tractable approximation (d) presented additional results for the flow and new results for the magnetic field in the magnetosheath. In the MHD model, the magnetopause is represented by a MHD tangential discontinuity and the bow shock by a MHD fast shock. Simplification was required because solution of the MHD equations for the full three-dimensional geometry of the magnetosphere flow was not possible with the computers and numerical algorithms then available. In the magnetosphere, the influence of the plasma is disregarded as small, so that the magnetic field is approximated by a vacuum field as in the C-F model. In the solar wind plasma outside the magnetopause, the influence of the magnetic field on the flow is disregarded as small, as evidenced by values typically in excess of 5 for the Alfvén Mach number $M_A = (4\pi\rho_\infty v_\infty^2/B_\infty^2)^{1/2}$. This reduced the exterior flow problem to one of GD rather than MHD. Finally, the magnetic field in the magnetosheath could be determined by solving the remaining Faraday magnetic induction equation of MHD using values for the velocity and density provided by the GD solution. Further simplification by disregarding the small effects of p_∞ and a slight change in the value for K in the Newtonian formula for the pressure of the plasma on the magnetopause returned the specification of the magnetopause shape to the same form as in the C-F model. In this way, previous results for magnetopause location obtained using the C-F model were recovered without change. Only the underlying reasons for using the equations were different. This simplification of the basic MHD model has come to be known as the gasdynamic convected field GDCF model. Figure 5 shows plots of the streamlines, Mach lines, and magnetic field lines obtained in this way. Their counterparts for the density, velocity, and magnetic field strength were also provided in the original presentation.

We were confident that our calculations were accurate and that the iteration solution converged properly, but the observational data from space and the range of conditions for which we had solutions were too limited to establish the validity of the results and the underlying MHD model by extensive comparisons. Elements of the results could be evaluated, however, and this was done in a number of ways to enhance credibility. A notable example was an experiment undertaken to answer the question whether the blunt-body aerodynamic methods we used could provide satisfactory results for a monatomic gas with $\gamma = 5/3$. These methods had been developed and tested extensively for air, a diatomic gas for which $\gamma = 7/5$, but never for a gas with $\gamma = 5/3$. To answer this question, a metal model of the forward part of the magnetosphere was fired at a Mach number of about 4.65 from a 50 caliber light-gas gun through the elongated test section of an otherwise normal supersonic wind tunnel at Ames that was filled with argon, a monatomic gas with $\gamma = 5/3$. Argon was selected because it has the desired value for γ, was available, and had a sufficiently low speed of sound that the desired Mach number could be achieved by firing the projectile into stationary gas. Shadowgraphs of the model in flight show the bow shock to be located exactly where the calculations indicate.

Although the paper [*Spreiter, Summers, and* Alksne, 1966a] in which these results, the MHD model, and its GDCF approximation were first presented has been cited extensively, it did not receive immediate acceptance when submitted for publication. It was rejected quickly by the *Journal of Geophysical Research* with little comment other than that it was not suitable for that journal. We then sent it to an Associate Editor of *Planetary and Space Science*, but received no response for several months. The explanation was that he had been selected to be an astronaut about the time our paper was received, and had been too busy to send it out for review. We then sent it to the Principal Editor, David Bates in Northern Ireland, and he accepted it. We were grateful that he was also willing to accept rather large plots of the results to facilitate their quantitative use.

The new MHD model and the GDCF approximation to it found immediate use in the interpretation of data pouring in from magnetometers and plasma probes in space. *Cahill and Patel*, [1967] analyzed 70 distinct magnetopause crossings by Explorer 12 between August and December 1961 and found that the magnetopause locations were generally consistent with the theoretical predictions. They also concluded that the data were generally consistent with the closed magnetopause envisioned originally by Chapman and Ferraro, although sometimes the boundary was open in the manner suggested by *Dungey*, [1963]. Because apogee of Explorer 12 was only about 13 earth radii, the bow shock was crossed on only three orbits when the magnetosphere was in a highly compressed state. *Kavanaugh*, [1967] examined these and concluded that the location of the bow shock and the change in magnitude of the magnetic field across it were consistent with the fluid model predictions.

While data from Explorer 12 together with more fragmentary data from earlier spacecraft served to suggest the presence of a bow shock upstream of the magnetopause, definitive evidence establishing the permanent presence of these surfaces in about the predicted locations and the validity of the GDCF model generally was provided by more than six



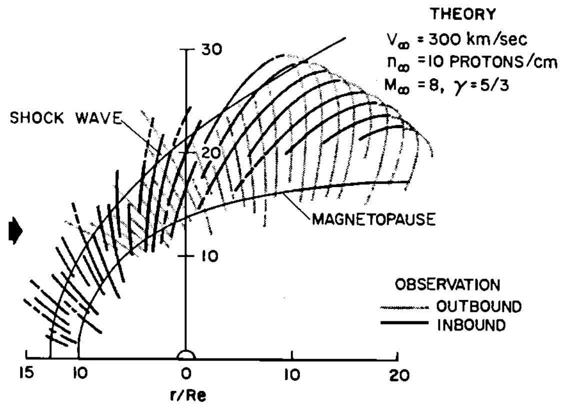


Figure 5. Streamlines, Mach lines, and magnetic field lines for solar wind flow past the magnetosphere for $M_\infty = 8$, $\gamma = 5/3$, and 45° angle between the magnetic field and the flow direction in the solar wind, [*Spreiter, Summers, and Alksne*, 1966a].

months data from the magnetometer [*Ness et al.*, 1964, 1066b] and plasma probe [*Wolfe et al.*, 1966] on IMP-1 spacecraft launched into an orbit with initial apogee of 31 earth radii on November 27, 1963. The left-hand side of Figure 6 from *Spreiter et al.*,[1968] and *Spreiter and Alksne*, [1969a] shows representative data from the plasma probe on a single inbound pass through the magnetosheath. Data from only the four lowest voltage or energy steps of the probe are shown, since current was seldom observed in the energy channels above 3749 volts (846 km/s for protons). The absence of plasma flux at geocentric distances less than about 11.5 earth radii indicates that IMP-1 was within the magnetosphere, and not exposed to the flowing solar plasma. The presence of almost all of the current output in one energy channel at distances greater than about 16 earth radii indicates that IMP-1 was in the solar wind beyond the bow shock, where the random thermal velocities are sufficiently small compared with the directed bulk velocity of the flow that the velocity distribution or energy spectrum of the particles is too narrow to bridge the gap between the energy windows of the plasma probe. The broad energy spectrum observed in the intervening part of the orbit is indicative of the hot plasma in the magnetosheath. The right-hand side of Figure 6 is a summary plot in which are drawn portions of the first 29 orbits in which *Wolfe et al.*,[1966] interpreted their data as indicating IMP-1 is in the magnetosheath. The dashed portions of some of the lines represent uncertainties in the determination of the bow shock or magnetopause. Also included are the locations of these surfaces calculated using the GDCF model and a single representative set of conditions in the solar wind. The value of 5/3 for γ was selected because it provided a better fit to the data than $\gamma = 2$, and also because it was the normal value for a monatomic gas. Although the criterion used to distinguish the various region is different, *Ness, et al.*, [1964, 1966b] presented very similar summary plots based on the magnetometer data from IMP-1. Soon thereafter, *Fairfield*, [1967] examined simultaneous magnetometer data from IMP-1 and IMP-2 when they were on opposite sides of the bow shock. He showed that the magnetic field in the magnetosheath was consistent with the model predictions, and parallel to the magnetopause in accordance with the representation of it by a MHD tangential discontinuity.

These comparisons enhanced confidence in the usefulness of the GDCF model, but they were marginally precise and barely possible to make. Conditions in the solar wind simultaneously with the observations in the magnetosheath needed for model input were unknown, and solutions were available for only one Mach number, 8, and two values for γ, 2 and 5/3. Moreover, the plasma probe resolution was grossly inadequate for determining the density, velocity, and temperature of the plasma in the magnetosheath. None of these quantities were measured directly. They had to be deduced by fitting a velocity distribution function, Maxwel-

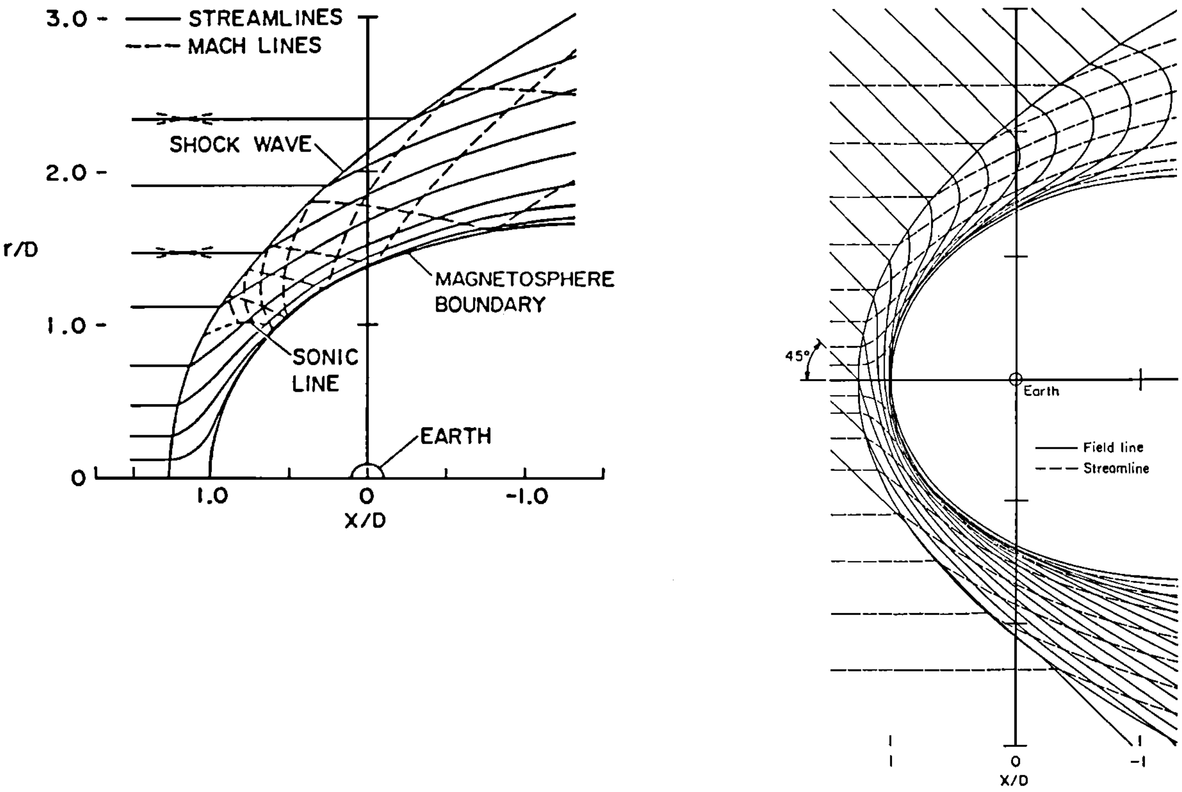


Figure 6. MP-1 plasma probe data for inbound part of orbit 17 on Jan. 31-Feb. 1, 1964, and a summary of IMP-1 magnetosheath crossings for the first 29 orbits, [*Spreiter and Alksne*, 1969a].

lian or otherwise, to the plasma probe data. Instead of tens of measurements in various directions needed for each energy level, the plasma probe measured only the maximum flux of particles in three unequal segments of a complete revolution for each energy level. The plasma probe could not be changed, so an extension of the GDCF model was developed to display the Maxwellian velocity distribution equivalents of the GD values for velocity, density, and temperature [*Spreiter, Alksne, and Abraham-Shrauner*, 1966b]. The results contributed to the interpretation of plasma probe data, but tended to be too complex for detailed comparisons with data and were soon made obsolete to some degree by the introduction of higher resolution plasma probes and improved data reduction algorithms.

The most detailed comparison of calculated and observed conditions in the magnetosheath during the discovery period was that of *Spreiter and Alksne*, [1968a] using data from Pioneer 6. It was launched into an escape trajectory on December 16, 1965 during a period of exceptionally low geomagnetic activity. Data from the magnetometer [*Ness et al., 1966a*] and plasma probe [*Wolfe and McKibbin*, 1968], which had much higher resolution than plasma probes on earlier spacecraft, were shown to be in good agreement with GDCF results calculated specifically for the solar wind conditions measured by Pioneer 6 just after crossing the bow shock. The calculated and observed locations of the bow shock and magnetopause matched perfectly, and the plasma and magnetic field properties matched as well as could be hoped for considering the uncertainties in their determination from the quantities actually measured by the plasma probe. To provide further insight, comparisons of the velocity reported by *Wolfe and McKibbin*, [1968] were made with two different sets of calculated values. One was the obvious set calculated directly using the GDCF model. The other was the maximum particle velocity indicated by the Maxwellian kinetic theory interpretation of the fluid velocity and temperature, [*Spreiter, Alksne, and Abraham-Shrauner*, 1966b]. The was done because it more closely represents what the data analysis procedure actually calculated for the quantity called "velocity". Values reported for the velocity were about midway between the two sets of values, which differed by about 35 to 40 km/s in the magnetosheath and only about 2 km/s in the solar wind.

With the change in paradigm from the particle approach of Chapman and Ferraro to the fluid model came the necessity to reexamine conditions near the neutral points on the magnetopause to which attention is directed in Figure 1. In the C-F model in which the pressure $K\rho_\infty v_\infty^2\cos^2\psi$ on the magnetopause is balanced by the magnetic pressure $B^2/8\pi$, an exact solution would indicate the magnetopause is parallel to the solar wind flow direction at the neutral points because $\boldsymbol{B} = 0$ there. If the static pressure is added to the C-F pressure so that the pressure on the magnetopause is given by $p_\infty + K\rho_\infty v_\infty^2 \cos^2\psi$, as it is in the Newtonian approximation for hypersonic flow, an exact solution would indicate the magnetopause has a sharp cusp directed toward the earth at each neutral point as in the solutions of *Spreiter and Hyett*, [1963], *Spreiter and Summers*, [1963], *Slutz and Winkelman*, [1964], and *Grad and Hu*,[1966]. As pointed out by *Spreiter and Summers*, [1967] and *Spreiter et al.*, [1968], this is not an admissible solution in the fluid models, even though the pressure does not vanish, because the supersonic flow cannot follow the cusped contour without developing a shock which could not be balanced by a jump in magnetic pressure on the magnetosphere side of the boundary. The difficulties were resolved by noting that although the boundary of the geomagnetic field could have cusps pointing toward the earth, the magnetosheath flow would separate from the boundary and reattach downstream of the cusp without forming a shock wave, much like air flowing over the open top of a convertible. The interior of the cusp would be filled with relatively stationary hot plasma that would leak from the point of the cusp and flow down toward the earth. It was noted that such injections of plasma into the magnetosphere had been seen in laboratory experiments designed to simulate the interaction of the solar wind and the geomagnetic field, see e.g., [*Osborne et al.*, 1964; *Kawashima and Fukushima*, 1964; *Cladis et al.*, 1964; and *Waniek and* Kasai, 1966]; and that geomagnetic disturbances in the polar regions observed during the IGY (July 1957 to December 1958) and the Second Polar Year (August 1932 to August 1933) tended to have local maxima where the field lines from the cusps reach the earth, as would be anticipated on the basis of direct invasion of charged particles or propagation of MHD waves into the polar regions along the field lines from the neutral points. These regions are now called the polar cusps, and a great variety of phenomena are associated with them.

The permanent existence of a long geomagnetic tail extending downwind from the magnetosphere in the solar wind like a windsock was recognized by *Ness*, [1965] from IMP-1 magnetometer data extending out to apogee at 31 earth radii. A similar tail-like structure was found to be present and well defined at 80 earth radii in data from the magnetometer, [Ness *et al.*, 1967a and *Behannon*,[1968], and the plasma probe, [*Mihalov et al.*, 1968], on Explorer 33. *Wolfe et al*, [1967] and *Ness et al.*, [1967b] reported indications of the tail on a number of hour-long intervals when Pioneer 7 was in the anticipated vicinity of the tail at 1000 earth radii. That ended a period of uncertainty in which a variety of proposals were made for the shape of the tail and the conditions in it. One of the most popular, called "teardrop", confined the geomagnetic field in a finite tail that tapered inward and closed like the small end of an elongated egg. An alternative, developed into a quantitative model that avoided difficulties in balancing the pressures on the end of the magnetopause tail in the tear-drop configuration, was based on the concept of a long nonmagnetic tail filled with quiescent plasma that exerted a constant pressure on the flanks of the tail and confined the geomagnetic field within a relatively short region, [*Spreiter and Hyett*, 1963; *Spreiter and Summers*, 1963, *Slutz and Winkelman*, 1964]. Since both configurations were inconsistent with the observations, a new first order model was developed within the spirit of the GDCF model that recognized the presence of the extended magnetic tail and the effects of the associated cross-tail current, [*Spreiter and Alksne*, 1969]. It was assumed, as in the treatment of the polar cusps, that the pressure of the magnetosheath plasma on the magnetopause is given by $p_\infty + K\rho_\infty v_\infty^2\cos^2\psi$, rather than $K\rho_\infty v_\infty^2\cos^2\psi$, so as to represent conditions better as $\cos\psi \to 0$ along the distant tail. The magnetic field in the tail was assumed to be directed essentially parallel to the tail, toward the earth in the north half of the tail and oppositely in the south half, and to have a strength inversely proportional to the cross-section so as to preserve the magnetic flux along the tail. The resulting shape for the tail flared out only slightly more than indicated by the original GDCF model, but the logical foundations for the solution were restored in the tail region. Perhaps the main consequence to us was that it indicated that results obtained using continuum fluid concepts could be expected to be in good agreement with observations for at least many tens of earth radii downstream of the earth.

Although all of the preceding discussion is about steady-state phenomena, Chapman was led into his pioneering studies over seventy years ago by his desire to explain transient geomagnetic variations. He considered effects of two classes of disturbances, (a) large plasma clouds ejected from the sun by an explosive event and (b) long-lived rotating beams of plasma emanating from an active region on the sun. Early discussions regarded the clouds and beams as intrusions into the otherwise effective vacuum of interplanetary space. The qualitative manner in which the geomagnetic field would carve a cavity in a finite cloud of solar plasma advancing through a vacuum toward the earth was one of the first topics discussed by *Chapman and Ferraro*, [1931], but it was many years before actual quantitative solutions for their model were given by *Spreiter and Summers*, [1965a]. Since the permanent existence of the solar wind was thoroughly established by that time, these solutions were of significance primarily for their relative simplicity, conceptual value, and historical interest. The concept of clouds and rotating beams of solar plasma remained, but now they were considered to move through the general background of less dense and slower moving solar wind plasma. To provide insight into a more realistic transient situation, solutions were also presented showing how the magnetopause changes in size from one steady state to another as a discontinuity surface in the solar wind passes by. Results for representative conditions indicate that points on the magnetopause may move from their initial to final positions in less than a minute after passage of a discontinuity.

By this time, considerable information was accumulating regarding the nature of interplanetary shock waves and other abrupt changes in the solar wind. To gain a better understanding of the implications for solar-terrestrial relations, Joan Hirshberg undertook a statistical study at Ames of observations of geomagnetic sudden commencements and solar flares. From her results [*Hirshberg*, 1968] and those of *Taylor*, [1969], who analyzed data from IMP-3 as eight "possible shocks" passed by, it was concluded that the shock shape, speed, and transit time from the sun to the earth depend primarily on the energy of the initial disturbance and the properties of the solar wind at the time, and not upon such details as the initial angular extent or distribution of energy within the initiating event on the sun.

Frequently the clearest sign of passing through the bow shock into the magnetosheath is a big increase in fluctuations in the magnetic field and plasma properties. *Spreiter and Jones*, [1963] associated this with fluctuations downstream of a collisionless shock as illustrated in Figure 3. An alternative based on continuum gasdynamics described by *Spreiter et al*, [1968], and *Spreiter and Alksne*, [1969a] is that small disturbances in the solar wind are strongly amplified in passing through a shock, e.g., the amplitude of small pressure fluctuations in a gas with $\gamma = 5/3$ and Mach number M_1 upstream of a shock increases by a factor of about $0.4\ M_1^2$ when passing through a normal shock. For a representative value of for M_1 of 8 at the nose of the bow shock, this indicates an amplification of about 25. *McKenzie and Westphal*, [1968] extended the analysis to apply to oblique shocks and showed that similarly large amplifications were a general property of high Mach number shocks.

The stability of the magnetopause is another topic that has not been touched upon in the preceding discussion. It seemed to us that it must be stable in at least some gross sense, since otherwise the generally good agreement between observations and results of the steady-state models could hardly be expected to occur. On the other hand, a number of studies published during the early years of the discovery period arrived at the opposite conclusion. *Dungey*, [1958, 1963] considered that the flow of solar plasma along the magnetopause would generate surface waves in the same way that wind generates waves on water through the action of a Kelvin-Helmholtz instability. He and *Parker*, [1958] gave a theory of such waves, and concluded that the magnetopause is unstable. The applicability of the results to the magnetopause was not assured, however, because of the neglect of known features of the phenomena such as the compressibility of the plasma, supersonic speeds of the flow along the magnetopause at points away from the nose, the curvature of the boundary, and the effects of nonlinear terms in the governing equations. A different analysis based on the strict application of the C-F theory was developed by *Spreiter and Summers*, [1965b]. They found, if the wavelength and amplitude are sufficiently small that curvature and higher-order effects can be disregarded, all perturbations, except those having wave fronts aligned with the direction of the local magnetic field within the magnetosphere, damp exponentially with time as they move along the magnetopause with the local velocity of the adjacent flow. Aligned waves, which neither damp nor amplify in this approximation, were examined further by inclusion of curvature and higher-order effects. The analysis showed that curvature introduces a destabilizing effect in small indented regions of the magnetopause, and that isolated columns of solar wind plasma aligned with the magnetospheric magnetic field could penetrate the magnetopause and be injected into the outer magnetosphere. It is well within the range of possibility that these columns could be identified with the flux ropes containing magnetosheath plasma often observed just inside the magnetopause.

By 1969, our modeling of solar wind flow past the earth's magnetosphere had reached a definite plateau. The MHD model and the GDCF approximation to it were generally accepted and providing a firm foundation for the interpretation of observations. The MHD equations were still too difficult to solve, but solutions of the GDCF approximation achieved using state-of-the-art methods of computational fluid dynamics had been shown to be in general agreement with numerous observations. They were, in fact, virtually the only quantitative theoretical results for the location of the magnetopause, bow shock, and conditions in the magnetosheath. They also provided a rationale for the continued use of vacuum magnetic field models for the geomagnetic field in the magnetosphere and in the magnetotail. A significant improvement in the calculation of the magnetic field in the magnetosheath was made at this time by the development of a new decomposition procedure by *Alksne and Webster*, [1970]. Without further approximation, it simplified a complicated three-dimensional calculation in the GDCF model to a single two-dimensional calculation, plus two additional components that could be calculated by applying very simple formulas to the gasdynamic results.

As confidence grew in the MHD and GDCF models, it was natural to inquire how they should be modified to apply to other planets and the moon. If the basic model is assumed to remain valid, the differences must be the result of changes in the boundary conditions representing the solar wind and the planetary obstacle. For Venus and Mars, the solar wind speed is about the same as at earth and the density differs by a factor of about two up or down in accordance with its approximate proportionality to the inverse square of the distance from the sun. Such a change in density is well within the range experienced by the earth, and nothing significant would be expected in response except for a slight shrinking of the size of the obstacle as the density increases, and vice versa. The nature of the planetary obstacle may be quite different from that for earth, however, as illustrated in Figure 7 from *Spreiter and Alksne*, [1970] based on modeling results of *Spreiter et* al., [1970a]. For the earth, the strong geomagnetic field forms a large magnetospheric obstacle in the solar wind, and a bow wave forms upstream of it because the flow is supersonic. For average solar wind conditions, the nose of the magnetopause is at about 10 earth radii, and the nose of the bow shock is at about 13 earth radii. Neither Venus nor Mars has a significant magnetic field, but the density and electrical conductivity of the upper ionosphere of these planets are sufficiently great to deflect the solar wind around the ionosphere and to form a bow shock upstream of it. Because the ionospheric pressure falls off much more rapidly with height than the magnetic pressure of the geomagnetic field, the ionopause boundary between the solar wind and ionospheric plasmas is wrapped closely around the planet. For both Venus and Mars its nose is only a few hundred kilometers above the planetary surface and the bow shock is about a third of a planetary radius upstream of it as illustrated. Although there has been a continuing debate whether or not Mars has a weak magnetic field, Venus and Mars remain the only planets known to have an ionosphere type interaction with the solar wind. With the possible exception of Pluto, yet to visited by spacecraft, all of the other planets are now known to act as magnetic obstacles in the solar wind.

The moon, having neither a magnetic field nor an or ionosphere, presents yet another type of obstacle in the solar wind, as illustrated in Figure 6 based on MHD modeling [*Spreiter et al.*, 1970b]. It was assumed as an idealization of observations, that the solar wind plasma flows directly onto the lunar surface, that the magnetic field is continuous from the solar wind into the moon, that no significant electrical currents flow in the moon in the steady state, and that there exists a void in the solar wind downstream from the moon in which neither particles nor electrical currents are to be found. Two alternatives exist for describing conditions at the surface of this void, depending on whether or not an electric current sheet forms. If one does form, the boundary of the void must be represented by an MHD tangential discontinuity. If no current sheet forms, the magnetic field in the moon and the trailing void must join continuously with that in the surrounding flow. Both theoretical considerations and data from Explorer 35 were cited to indicate that either possibility may occur, depending particularly on the orientation of the interplanetary magnetic field. The upper sketch illustrates results for the special case in which the solar wind velocity and magnetic field vectors are parallel. This was our first solution of the full MHD equations. It made use of an important simplification of *Imai*, [1960] and *Grad*, [1960] that $\boldsymbol{B} = \lambda \rho \boldsymbol{v}$ and $(\mathbf{v}.\nabla)\lambda = 0$ in a nondissipative MHD aligned flow. This enables the MHD equations to be reduced to those of GD of a pseudo gas having a nonphysical equation of state by introducing a set of new variables for the pseudo velocity, density, and pressure. The point of this was that a class of MHD solutions could be obtained by using established and simpler methods of GD. This possibility and the required transformations had been given in detail in *Spreiter et al*, [1966a], but this was the first time they were used to obtain actual MHD solutions for a space application. The lower sketch indicates results for other cases in which the velocity and magnetic field vectors are not aligned. They are based on a combination of MHD concepts and GDCF modeling approximations.

To supplement the modeling described above, studies of more complex plasma models and special features of MHD were also investigated within our branch. *Sturrock and Hartle*, [1966] developed a two-fluid model of the solar wind in which the protons and electrons were treated separately. Barbara Abraham-Shrauner commenced her studies of wave propagation in the *Chew, Goldberger, Low*, [1956] model of an anisotropic plasma at Ames, and continued them at Washington University after leaving Ames. Compared with the comparatively simple results of MHD theory, commonly displayed using Friedrich's diagrams for the phase and group velocities, her results displayed a great richness in possibilities when the plasma is anisotropic. Other properties of plasmas, MHD and otherwise, were studied by *Morioka and Spreiter*, [1968, 1969, 1970a,b].

It was recognized during this period that the MHD concepts we had been developing had an even wider field of application. Although no quantitative results were given, applications to comets in the solar wind and the interaction of the solar wind and the interstellar medium were discussed in *Spreiter et al.*, [1968a] and *Spreiter and Alksne*, [1970].

As a final note on this period, I would like to cite a virtually unknown example of a consequence of our work that illustrates how knowledge developed in one field can have a significant impact on a seemingly unrelated field. One day when Alberta Alksne's neural surgeon son, John Alksne, was visiting, he told her of his poor success rate in brain surgery on severely injured patients. The problem stemmed not so much from the injury itself or the primary action of the surgery, he said, but from the size of the incision required to enable all the small blood vessels to be tied off. Almost without thinking, she suggested using a focused magnetic field to stop the flow of blood long enough for it to clot and seal itself. He developed the idea and introduced the techniques in his operations with such great success that it was featured in an article with many pages of photographs in *Life* magazine. It is likely that the economic and humanitarian benefits of this advance greatly exceeds the total cost of all of our magetosphere modeling.

THE POST-DISCOVERY PERIOD: 1969-1997

Many changes occurred at the beginning of this period. I was appointed professor of applied mechanics and aeronautics and astronautics at Stanford in late 1968, succeeding my former advisor, Irmgard Flügge-Lotz. It is only 10

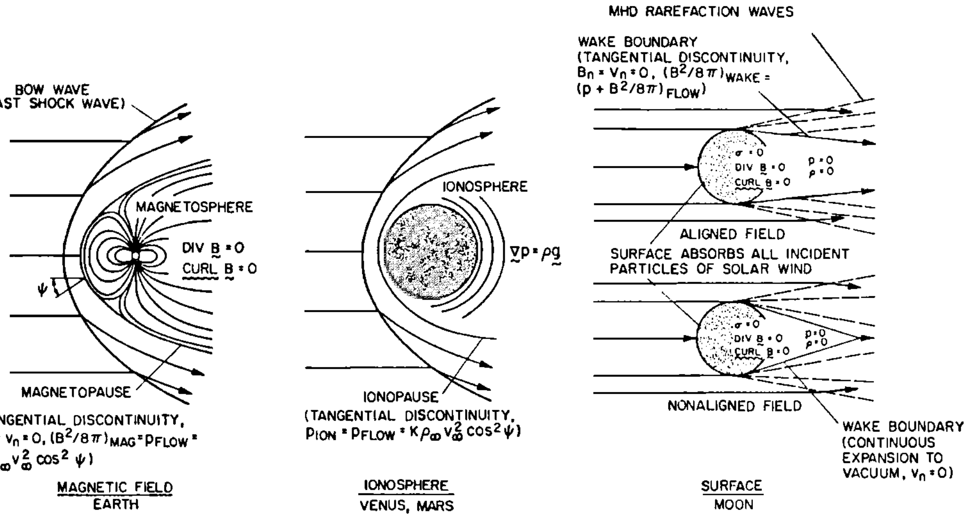


Figure 7. Types of interaction of the solar wind with the earth, moon, and planets, [*Spreiter and Alksne*, 1970].

miles from Stanford to Ames, so it was easy to maintain close relations with Ames for several years, and for my research students to use the facilities of Ames with the support of NASA grants. However, Alksne retired in a couple of years, joined the Peace Corps, and went to Kenya to teach mathematics in a girls high school. Webster stopped when Alksne left. Several who had come as postdoctoral research associates finished their one to three year appointments. These included Abraham-Shrauner, Rycroft, Marsh, Morioka, McKenzie, Westphal, and Hartle. Sturrock's services had already been discontinued when Ames management needed to reduce the use of consultants. For the same reason, arrangements were made for Hirshberg to join me at Stanford as a senior research associate under a NASA grant. Ray Reynolds was appointed to my former position as chief of the Theoretical Studies branch and work turned toward planetary interiors and atmospheres. Audrey Summers soon joined in those studies with the same effectiveness she displayed several years before in our plasma studies. She was appointed assistant chief of the branch a few years later. It was not long before Barnes was about the only permanent member of the branch carrying on with the solar wind studies. All such work had not ceased at Ames, however, because Sonnett, Colburn, Mihalov, Wolfe, and their associates continued extensive studies of magnetometer and plasma probe data for many years.

At Stanford, my first Ph.D. advisee was Art Rizzi, who has just been awarded his M.S. degree. Flügge-Lotz had been his advisor and, with his approval, passed him on to me. We were both anxious to get started promptly, so arrangements were made that he could use the computers at Ames and join in our analysis of solar wind interaction with nonmagnetic planets, with application to Venus and Mars, [*Spreiter et al.*, 1970a; *Spreiter and Rizzi*, 1972, 1974a]. His main project, however, was to develop the first MHD solution for flow past the earth's magnetosphere. It was for the special case of aligned flow in which the velocity and magnetic field vectors are parallel everywhere. As described above, the MHD equations can be reduced to those of GD for this case through the introduction of new variables, except that the equation of state for the gas is different and more complicated. This application was more difficult than the aligned MHD solution for the moon, however, because of the presence of the bow shock and a region of subsonic flow embedded in an otherwise supersonic flow. The results for a representative range of Mach numbers and Alfvén Mach numbers [*Spreiter and Rizzi*, 1974b] are similar to those of the simpler GD model for Alfvén Mach numbers greater than about 5, but differ significantly for lower values. Since the Alfvén Mach number is usually greater than 5, these results help explain the generally good agreement between observations and the GDCF results. Since values lower than 5 are observed at times, these results also demonstrated the need for MHD solutions if the interesting extreme events are to be modeled. When Rizzi finished his Ph.D. studies, he joined the Computational Fluid Dynamics Branch at Ames where he applied his knowledge to develop perhaps the biggest computational fluid dynamics code at the time for application to the aerothermodynamics of the shuttle reentry into the atmosphere. After many years at Ames, he moved to Sweden where he is now Professor of Aeronautics at the Royal Institute of Technology.

After Rizzi, only a few of my advisee worked on space plasmas for their Ph.D. research, and none did magnetosphere modeling in the sense discussed above. The closest were Richard Desautel, Charles Lombard, Joe Reagan, and Carlton Ruthlin. Desautel analyzed the dynamical motion of the distant geotail, Lombard modeled the effects of ion mass loading in the distant solar wind. Reagan analyzed energetic particle data in the magnetosphere, and Ruthlin studied MHD wave propagation in nonuniform magnetic field. Other advisees studied a variety of topics including transonic aerodynamics, helicopter aerodynamics and noise, natural and rocket induced motions in the atmosphere, climate dynamics and chaos theory, and even a viscous fluid problem arising in the manufacture of scotch tape. They all continued their careers doing research and/or teaching computational fluid dynamics, aeronautics, or mathematics.

Not counting the course in space physics that he took in 1964, my long professional association with Steve Stahara began in early 1969, and has continued unbroken now for 28 years. He was born and raised in Youngstown, Ohio where he developed a lifelong interest in model airplanes. He enrolled in Case Institute of Technology where, after a year of high grades, he was one of a select few admitted to their Engineering Science program. He graduated Summa Cum Laude in 1964, and accepted a Hughes fellowship for graduate studies in the Aeronautics and Astronautics department at Stanford. He obtained M.S. and Ph.D. degrees in that department under the direction of Milton Van Dyke, and graduated in January, 1969. During those years, he spent two summers at Hughes Aircraft in El Segundo, California on the pre-flight testing of the Surveyor spacecraft, the first soft-lander on the Moon. Upon graduation, he accepted a staff scientist position at Nielsen Engineering & Research, Inc., a small research and development firm in nearby Mountain View, California. It was headed by Jack Nielsen, a close colleague of both Van Dyke and me from our early days at Ames. The position was created to work with me on transonic aerodynamics. That came about as the result of a chance meeting with Nielsen the first week of my tenure at Stanford. He invited me to be a consultant with his firm, and to seek support from Ames to incorporate the new and rapidly-evolving computational advances into the transonic aerodynamic research I had done in earlier years at Ames. I was receptive to the idea, and my department chairman and dean were encouraging because it was compatible with the one condition they had imposed on me that my Stanford research should not be in aerodynamics. This may seem strange in view of my dual appointment in the Aeronautics and Astronautics department, but my primary base was in a new Applied Mechanics department and they did not want it to look like a second aero department. I asked Van Dyke if he knew of any good candidates, preferably a recent graduate with a Ph.D. He recommended his student, Stahara. Nielsen and I liked him, and Stahara began working at Nielsen in February 1969. We set to work on transonic aerodynamics with applications to wings, bodies, and turbomachinery. Our last publications in that field were in 1982. Steve was always fascinated by the solar wind and magnetosphere studies and their connection with aerodynamics, however, so it was quite natural that we talked about such matters from the beginning of our association. NASA support for theoretical studies of space plasmas was on the decline at that time, however, so we saw little opportunity to do sponsored research on the magnetosphere or solar wind at Nielsen. In any case, I was carrying on my space plasma modeling efforts at Stanford in the face of diminishing support for student research assistants.

An unexpected event occurred in 1975 that enabled me to refocus on space research and be joined by Stahara. Norman Ness called and informed me that I had been invited to join him and four others, Herb Bridge of MIT, Sig Bauer of NASA Goddard, and Alex Dessler and Paul Cloutier or Rice, to participate in a Bilateral US/USSR Seminar on Mercury, Mars, and Venus at the Space Research Institute, IKI, in Moscow. It was a very interesting experience. About thirty prominent Soviet space scientists joined the six of us in very intense and questioning discussions. The meetings were extremely contentious, not so much between us and the Soviets, but among the Soviets themselves. The big question was whether Dolginov's magnetometer had detected a small Martian magnetic field as it orbited that planet, or not. I was in an interesting position because my models were being used to support both sides of the argument. Only now have I learned from *Breus*, [1997] that they "expected the Americans . . . to behave as a kind of Supreme Court, judging our discussions and internal disagreements on the problem of the intrinsic magnetic field of Mars". At the end of the day and on the weekend, our hosts, particularly Tamara Breus, Oleg Vaisberg, Igor Podgorny, his wife and son, Konstantin Gringauz, Alex Galeev, Valerie Troitskaya, Roald Sagdeev, and Igor Belotserkovskii in various combinations, graciously and expertly guided us around the sights, institutions, and even some of the subtleties of Moscow life. Altogether, it was an unforgettable experience.

On returning home, I inquired about the possibility of obtaining a NASA grant to do further modeling of solar wind interactions with the earth and other planets. The results we possessed were accurate solutions, but they were limited in number and covered only a fraction of the parameter range of interest. More could not be generated readily because our old programs could not be used on the new computers, and the old computers no longer existed. Moreover, the algorithms we used for the gasdynamic flow had been superseded by new and vastly improved ones developed in the Computational Fluid Dynamics Branch at Ames for application to the space shuttle. NASA was quite receptive to supporting this modernization, but insisted that there be "no new physics" and that the project be completed promptly. That seemed more like a task to be undertaken with Stahara at Nielsen Engineering than with a graduate student at Stanford, so it was arranged to be done that way. The task was fairly straightforward. It was a matter of obtaining the programs, adapting them to the planetary applications, adding a component to calculate the magnetic field, evaluating them to determine which of the several alternatives would be most effective for our applications, and then, of course, to compile a set of solutions for general use. In addition, considerable attention was paid to simplifying the processes so that solutions could be obtained using readily available work stations instead of a supercomputer. The results, including a complete listing of the program, were reported in *Stahara et al.*,[1978, 1980] and *Spreiter and Stahara*, [1980a,b]. The new results coincided with those calculated before, but additional results could now be obtained easily for other conditions, both by us and by others with modest computing facilities. This was the first time others could easily use our models to calculate their own solutions.

This development brought us into direct collaboration with a larger circle of space scientists. Noteworthy amongst those in the eighties are Chris Russell, Janet Luhmann, Nancy Crooker, George Siscoe, Bob Holzer, Ray Walker, Jim Slavin, and Kurt Moore of the Institute of Geophysics and Planetary Physics at UCLA; Ed Smith of JPL; and John Mihalov of Ames. Model results were being used increasingly in the interpretation of data, and for comparative studies of the various planets. Meanwhile, we continued our model development to include features revealed in the observations but not included in previous models. These include effects of centrifugal flattening of the magnetosphere with application to Jupiter and Saturn, [*Stahara et al.*, 1989]; effects of mass loading by interaction with a comet or planetary atmosphere as believed important at Venus and Mars, and the transient effects of a passing interplanetary shock wave, [*Spreiter and Stahara*, 1992]; and full treatment of MHD effects, [*Spreiter and Stahara*, 1994]. Results of the GDCF model have been calculated for the solar wind conditions observed near each of the planets except Pluto which has yet to be visited by a spacecraft. The good accordance with the observed locations of the bow shock, magnetopause, and conditions in the magnetosheath have established the utility of the GDCF model over a wide range of conditions. Of particular note are the comparisons with Voyager 2 observations for Neptune, [*Spreiter and Stahara*, 1994]. The success of the model is remarkable in view of the extremely low density, only 0.005 protons/cm^{-3}, of the solar wind approaching Neptune. We believe that is the most rarefied gas flow that has ever been modeled successfully using continuum fluid concepts. It will be a long time, if ever, before *insitu* measurements are made in a lower density flow past an obstacle. That includes the flow of the local interstellar medium past the heliosphere, which current estimates indicate to be considerably denser than the solar wind at Neptune. The scale of that application is far larger than for planetary magnetospheres, of course, but the same methods with appropriate changes in the boundary conditions can be used to model the interaction, as we noted long ago, [*Spreiter et al.*, 1968; *Spreiter and Alksne*, 1970]. In 1993, Romana Ratkiewicz of the Space Research Center in Poland came to Ames for two and a quarter years as a National Research Council Senior Research Associate to work with Aaron Barnes and us to develop quantitative models for the conditions, both steady and unsteady, in the solar wind in the outer heliosphere and its interaction with the local interstellar medium flow by. Many interesting results have been been calculated, and several papers describing them are now in press.

The last remaining topic to mention is the space weather forecasting program sponsored by NASA, NSF, and the Air Force that Stahara and I have participated in since the early nineties. Its goal is to forecast, based on observations of the sun and solar wind, the total global geospace environment of the earth. Its purpose is to provide at least a few tens of minutes warning that a geomagnetic storm is about to occur with all of its attendant consequences on earth and in the surrounding space. These include temporary to permanent disablement of spacecraft, particularly those in geosynchronous orbit used for communications and weather forecasting, disruptions of telecommunications, and major electric power outages on earth caused by geomagnetically induced currents. A dramatic example of the latter occurred on March 13, 1989 when the entire Hydro Quebec system serving more than 6 million customers was plunged into blackout by such currents. *Kappenman et al.*, [1997] reports that "its impact was felt over the entire North American continent. Most of the neighboring systems in the United States came uncomfortably close to experiencing the same sort of cascading outage scenario.". Our role was to provide a robust computer model to forecast the location of the magnetopause, bow shock, and conditions in the magnetosheath from a knowledge of solar wind conditions measured by a spacecraft monitor stationed near the Lagrangian L_1 point at a distance of about 230 earth radii on a line to the sun. Others were to develop forecasting models for the magnetosphere, thermosphere, and ionosphere. The growing number of problems brought a sense of urgency to the project, and a validated model was wanted in a relatively short time. This meant that tested "off the shelf" models should be used as much as possible, so we chose the most recent version of our GDCF model The biggest challenge was to accelerate the process dramatically so that the forecast could be made in much less time than the approximately half hour it takes for the disturbances observed in space to reach the magnetosphere. We achieved forecast times of a few second by precalculating a large family of solutions and using a fast interpolating scheme to calculate the forecast. We made more than 4,000 simulated forecasts of conditions observed by ISEE 2 spacecraft based on solar wind conditions observed by ISEE 3 to demonstrate that the model could indeed provide useful forecasts. It is now in use as part of the first generation of space weather forecasting programs. What started out as speculative calculations in 1960 have become a useful tool in our technological society.

CONCLUDING REMARKS

We have told the story of our efforts in modeling solar wind flow past the magnetosphere, and of the preparation and events that enabled us to quickly move from aerodynamics to space physics. For both of us the preparation was a solid education in fluid mechanics and mathematics, many years of experience developing aerodynamic theory and testing it by comparisons with data from wind tunnels and actual flight. In addition, there is the matter of being in the right place at the right time, and being able to take advantage of the opportunity when it comes. Modeling space plasmas has come a very long way during the period discussed here. It has a long way to go before it can describe all matters of interest. Much of what is done on the research frontier is barely possible, but capabilities of computers and computational algorithms are still increasing at a rapid pace and great advances may be anticipated if the effort is sustained. Applications such as space weather forecasting hold the potential for vast economic benefits, but the problems to be faced are daunting and considerable effort will have to be applied to achieve an effective forecasting capability.

Finally, we would like to pass on a few thoughts relating to modeling fluid motions that are worth pondering. First, is the optimistic, perhaps arrogant, remark of *Lagrange*, [1788] following Euler's development of the equations of fluid dynamics of an inviscid fluid. He wrote "if the (Euler) equations involved were integrable, one could determine completely, in all cases, the motion of a fluid moved by any force". Although solutions of these equations had found many important uses including the design of jet airplanes by his time, the Harvard mathematician, *Birkhoff*, [1960] expressed a strong contrary view, "Euler's equations have been integrated in many cases, and the results found to disagree grossly with observation, flagrantly contradicting the opinion of Lagrange." He also wrote in the same book "The possibility of paradoxes cannot be admitted as a universal principle in fluid mechanics because the results of many experiments can be predicted with practical certainty.". With the aid of computers, many more examples of the latter exist today, but there are still many situations where prediction and reality differ significantly. Turbulence in fluid motions is ubiquitous. Sometimes its effects are benign, other times they are dominant and unpredictable. Our ability to deal with it in a reliable way is so limited that turbulence is sometimes said to be one of the great unsolved problems of physics. For these reasons, fluid dynamics remains both a science and an art. Theoretical models must be validated by comparison with observations. Anyone doubting this should visit Ames and see the gigantic wind tunnels and know, from the deafening noise of the air driven around inside them by motors of several hundred thousand horsepower, that they are being used. They stand as monuments to our continuing inability to reliably model the forces on a body in an airstream under all conditions.

Acknowledgments. Much of the funding for the work described here was provided by the National Aeronautics and Space Administration, the National Science Foundation, and the Air Force Office of Scientific Research. We both owe thanks to our colleagues, teachers, friends, and families in addition to those mentioned in this paper who have contributed to the success of our work.

REFERENCES

Alksne, A. Y., The steady-state magnetic field in the transition region between the magnetosphere and the bow shock, *Planetary Space Sci.*, *15*, 239-245, 1967.

Auer, P. L., H. Hurwitz, Jr., and R. W. Kilb, Low Mach number magnetic compression waves in a collision-free plasma, *Phys. Fluids*, *4*, 1105-1121, 1961.

Auer, P. L., H. Hurwitz, Jr., and R. W. Kilb, Large-amplitude magnetic compression of a collision-free plasma, II. Development of a thermalized plasma, , *Phys. Fluids*, *5*, 298-316, 1962.,

Axford, W. I., The interaction between the solar wind and the earth's magnetosphere, *J. Geophys. Res.*, 67, 3791-3796, 19602

Beard, D. B., The interaction of the terrestrial magnetic field with the solar corpuscular radiation, *J. Geophys. Res.*, 65, 3559-3568, 1960.

Behannon, K. W., Mapping of the earth's bow shock and magnetic tail by Explorer 33, *J. Geophys. Res.*, *75*, 907-930, 1968.

Biermann, L., Kometenschweife und solare korpuskular strahlung, *Z. Astrophys.*, *29*, 274-286, 1951.

Biermann, L., Physical processes in comet tails and their relation to solar activity, *Extrait des Mem. Soc. Roy. Sci. Liege* Quatr. Ser. *13*, 291-302, 1953.

Birkhoff, G., Hydrodynamics A study in logic, fact and similitude, Princeton University Press, 184 pp., 1960.

Bondi, H. On spherically symmetrical accretion, *Monthly Notices Roy. Astron. Soc. 112*, 195-204, 1952.

Breus, Tamara K., An unforgettable personality, *J. Geophys. Res.*, *102*, 2027-2034, 1977.

Briggs, B. R., and J. R. Spreiter, Theoretical determination of the boundary and distortion of the geomagnetic field in a steady solar wind, NASA TR-R-178, 1963.

Cahill, L. J., Jr., and V. L. Patel, The boundary of the geomagnetic field, August to December, 1961, *Planetary Space Sci.*, *15*, 997-1033, 1967.

Chew, G., M. L. Goldberger, and F. E. Low, The Boltzman equation and the one-fluid hydromagnetic equations in the absence of particle collisions, *Proc. Roy. Soc.*, *A236*, 112-, 1956.

Chamberlain, J. W., Interplanetary gas. III: A hydrodynamic model of the corona, *Astrophys. J.*, *133*, 675-687, 1961.

Chapman, S., *The Earth's Magnetism*, 127 pp., Methuen, London, 1st ed., 1936, 2nd ed., 1951.

Chapman, S., Solar Plasma, Geomagnetism and Aurora, in *Geophysique Exterieure*, pp. 373-502, Gordon and Breach, New York, 1963.

Chapman, S., and J. Bartels, *Geomagnetism, vol. 1 and 2*, 1049 pp., Oxford University Press, Oxford, 1940.

Chapman, S., and V. C. A. Ferraro, A new theory of magnetic storms, , *Terrestrial Magnetism & Atmospheric Electricity*, *36*, 77-97, 171-186, 1931; *37*, 147-156; 421-429, 1932; *38*, 79-96, 1933.

Cladis, J. B., T. D. Miller, and J. R. Baskett, Interaction of a supersonic plasma stream with a dipole magnetic field, *J. Geophys. Res.*, *69*, 2257-2272, 1964.

Clauser, F. H., Varenna Conference on Cosmical Aerodynamics, *Nuova Cimento*, Suppl. *22*, No. 1, 1960.

DeWitt, C., J. Hieblot, and A. Lebeau, *Geophysique Exterieure*, 624 pp., Gordon and Breach, New York, 1963.

Dungey, J. W., The steady state of the Chapman-Ferraro problem in two-dimensions, *J. Geophys.* Res., *66*, 1043-1047, 1961.

Dungey, J. W., *Cosmic Electrodynamics*, 183 pp., Cambridge University Press, 1958.

Dungey, J. W., Structure of the exosphere or adventures in velocity space, in *Geophysique Exterieure*, pp. 503-550, Gordon and Breach, New York, 1963.

Fukshima, N., Gross character of geomagnetic disturbance during the International Geophysical Year and the Second Polar Year, *Reports of Ionosphere and Space Research in Japan*, *16-1*,37-56, 1962.

Fuller, F. B., Numerical solutions for supersonic flow of an ideal gas around blunt two-dimensional bodies, NASA TN D-791, 1961.

Grad, H., Reducible problems in magneto-fluid dynamic steady flows, *Rev. Mod. Phys.*, *32*, 830-847, 1960.

Grad, H., and P. N. Hu, Neutral point in the geomagnetic field, *Phys. Fluids*, *9*, 1610-1611, 1966.

Hartman, E. P., Adventures in Research - A History of Ames Research Center 1940-1965, 555 pp., NASA SP-4303, 1970.

Hirshberg, J., The transport of flare plasma from the sun to the earth, *Planetary Space Sci.*, *16*, 309-319, 1968.

Hurley, J., Interaction between the solar wind and the geomagnetic field, New York University College of Engineering Report, 1961.

Imai, I., On flows of conducting fluids past bodies, *Rev. Mod. Phys.*, *32*, 992-999, 1960.

Inouye, M., and H. Lomax, Comparison of experimental and numerical results for the flow of a perfect gas about blunt-nosed bodies, NASA TN D-1426, 1962.

Jones, W. P., and V. J. Rossow, Graphical results for large-amplitude unsteady one-dimensional waves in magnetized collision-free plasmas with discrete structure, NASA TND-2356, 1965.

Kappenman, J. G., L. J. Zanetti, and W. A. Radasky, Space weather from a user's perspective: Geomagnetic storm forecasts and the power industry, *IOS*, *78*, 37-45, 1997.

Kawashima, N. and N. Fukushima, Model experiment for the interaction of solar plasma stream and geomagnetic field, *Planetary Space Sci.*, *12*, 1187-1201, 1964.

Kellogg, P. J., Flow of plasma around the earth, *J. Geophys. Res.*, 67, 3805-3811, 1962.

Lagrange, J. L., *Mechanique analytique*, X, p. 271, Paris, 1788.

Lebeau, A. Sur l'activité magnétique diurne dans les calottees polaires, *Ann. Geophys.*, *21*, 167-218, 1965.

McCrea, W. H., Shock waves in steady radial motion under gravity, *Astrophys. J.*, *124*, 1956.

McKenzie, J. F., and K. O. Westphal, The interaction of linear waves with oblique shock waves, *Phys. Fluids*, *11*, 2350-2362, 1968.

Mead, G. D., and D. B. Beard, Shape of the geomagnetic field solar wind boundary, *J. Geophys.* Res., *69*, 1169-1179, 1964.

Midgley, J. E., and L. Davis, Jr., Calculation by a moment technique of the perturbation of the geomagnetic field by the solar wind, *J. Geophys. Res.*, *68*, 5111-5123, 1963.

Mihalov, J. D., D. S. Colburn, R. G. Currie, and C. P. Sonnett, Configuration and reconnection of the geomagnetic tail, *J. Geophys. Res.*, *73*, 943-959, 1968.

Mitra, *The Upper Atmosphere*, 2 nd ed., 713 pp., The Asiatic Society, Calcutta, 1951.

Morioka, S., and J. R. Spreiter, Evolutionary conditions for shock waves in a collisionless plasma and stability of the associated flow, *J. Plasma Phys.*, *2*, 449-463, 1968.

Morioka, S., and J. R. Spreiter, Magnetohydrodynamic shocks in non-aligned flow, *J. Plasma Phys.*, *3*, 81-96, 1969.

Morioka, S., and J. R. Spreiter, The effect of finite Larmor radius on the perturbation flow mixing of a collisionless plasma, *J. Plasma Phys.*, *4*, 403-424, 1970.

Morioka, S., and J. R. Spreiter, Effect of dissipation due to firehose instability on perturbation half-jet flow of a collisionless plasma, *J. Plasma Phys.*, *4*, 629-641, 1970.,

Ness, N. F., K. W. Behannon, S. C. Cantarano, and C. S. Scearce, Observations of the earth's magnetic tail and neutral sheet at 510,000 kilometers by Explorer 33, *J. Geophys. Res.*, *72*, 927-933, 1967a.

Ness, N. F., C. S. Scearce, and S. C. Cantarano, Probable observations of the geomagnetic tail at 103 earth radii by Pioneer 7, *J. Geophys. Res.*, *72*, 3769-3776, 1967b.

Ness, N. F., C. S. Scearce, and J. B. Seek, Initial results of IMP-1 magnetic field experiment, *J. Geophys. Res.*, *69*, 3531-3569, 1964.

Ness, N. F., C. S. Scearce, and S. C. Cantarano, Preliminary results from Pioneer 6 magnetic field experiment, *J. Geophys. Res.*, *71*, 3305-3313, 1966a.

Ness, N. F., C. S. Scearce, J. B. Seek, and J. M. Wilcox, Summary of results from IMP-1 magnetic field experiment, *Space Res.*, *6*, 581-628, 1966b

Olson, W. P., The shape of the tilted magnetopause, *J. Geophys. Res.*, *74*, 5642-5651, 1969.

Osborne, F., M. P. Bachynski, and J. V. Gore, Laboratory studies of the variation of the magnetosphere with solar wind properties, *J. Geophys. Res.*, *69*, 4441-4449, 1964.

Parker, E. N., Dynamics of the interplanetary gas and magnetic fields, *Astrophys. J. 128*, 664-676, 1958.

Rossow, V. J., Magnetic compression of collision-free plasmas with charge separation, *Phys., Fluids*, *8*, 358-366, 1965.

Rossow, V. J., Magnetic compression waves in collisionless plasma - Oblique ambient magnetic field, *Phys. Fluids*, *10*, 1056-1062, 1967.

Slutz, R. J., The shape of the geomagnetic field boundary under uniform external pressure, *J. Geophys.Res.*, 67, 505-513, 1962.

Slutz, R. J., and Winkelman, J. R., Shape of the magnetospheric boundary under solar wind pressure, *J. Geophys. Res.*, *69*, 4933-4948, 1964.

Smith E. J. , P. J. Coleman, D. L. Judge, and C. P. Sonnett, Characteristics of the extraterrestrial current system: Explorer VI and Pioneer V, *J. Geophys. Res.*, 65, 1858,-1861, 1960.

Spreiter, J. R., Aerodynamics of wings and bodies at transonic speeds, *J. Aero-space Sci.*, *26*, 465-487, 1959.

Spreiter, J. R., and Alksne, A. Y., Slender body theory based on

approximate solution of the transonic flow equation, NASA TR-R-2, 1958.
Spreiter, J. R., and A. Y. Alksne, On the effect of a ring current on the terminal shape of the geomagnetic field, *J. Geophys. Res.*, *76*, 2193-2205, 1962.
Spreiter, J. R., and A. Y. Alksne, Plasma flow around the magnetosphere, *Revs. Geophys.*, 7, 11-50, 1969a.
Spreiter, J. R., and A. Y. Alksne, Effects of neutral sheet currents on the shape and magnetic field in the magnetosphere tail, *Planetary Space Sci.*, *17*, 233-246, 1969b.
Spreiter, J. R., and A. Y. Alksne, Solar-wind flow past objects in the solar system, *Annual Review of Fluid Mechanics*, *2*, 313-354, 1970.
Spreiter, J. R., A. Y. Alksne, and B. Abraham-Shrauner, Theoretical proton velocity distributions in the flow around the magnetosphere, *Planetary Space Sci.*, *14*, 1207-1220, 1966b.
Spreiter, J. R., A. Y. Alksne, and A. L. Summers, External aerodynamics of the magnetosphere, in *Physics of the Magnetosphere*, edited by R. L. Carovillano, J. F. McClay, and H. R. Radoski, pp. 301-375, Reidel, Dordrecht, Holland, 1968.
Spreiter, J. R., and B. R. Briggs, Analysis of the effect of a ring current on whistlers, *J. Geophys. Res.*, *67*, 3779-3790, 1962c.
Spreiter , J. R., and B. R. Briggs, Theory of electrostatic fields in the ionosphere at equatorial latitudes, *J. Geophys. Res.*, *66*, 2345-2354, 1961a.
Spreiter , J. R., and B. R. Briggs, Theory of electrostatic fields in the ionosphere at polar and middle geomagnetic latitudes, *J. Geophys. Res.*, *66*, 1731-1744, 1961b
Spreiter, J. R., and B. R. Briggs, Theoretical determination of the form of the hollow produced in the solar corpuscular stream by interaction with the magnetic dipole field of the earth, NASA TR-R-120, 1961.
Spreiter, J. R., and B. R. Briggs, Theoretical determination of the form of the boundary of the solar corpuscular stream produced by interaction with the magnetic dipole field of the earth, *J. Geophys. Res.*, *67*, 37-51, 1962a
Spreiter, J. R., and B. R. Briggs, On the choice of conditions to apply at the boundary of the geomagnetic field in the steady-state Chapman Ferraro problem, *J. Geophys. Res.*, *67*, 2983-2985, 1962b.
Spreiter, J. R., and W. Flügge, Irmgard Flügge-Lotz (1903-1974), in *Women in Mathematics*, edited by L. S. Grinstein and P. J. Campbell, pp. 33-40, Greenwood Press, N. Y., 1987.
Spreiter, J. R., and B. J. Hyett, The effect of uniform external pressure on the boundary of the geomagnetic field in a steady solar wind, *J. Geophys. Res.*, *68*, 1631-1642, 1963.
Spreiter, J. R., and W. P. Jones, On the effect of a weak interplanetary magnetic field on the interaction between the solar wind and the geomagnetic field, *J. Geophys. Res.*, 68, 3555-3564, 1963.
Spreiter, J. R., M. C. Marsh, and A. L. Summers, Hydromagnetic aspects of solar wind flow past the moon, *Cosmic Electrodynamics*, *1*, 5-50, 1970.
Spreiter, J. R., and A. W. Rizzi, The Martian bow wave — Theory and observation, *Planet. Space Sci.*, *20*, 205-208, 1972.
Spreiter, J. R., and A. W. Rizzi, Interplanetary space — A new laboratory for rarefied gas dynamics, in *Rarefied Gas Dynamics*, edited by K. Karamcheti, pp. 497-521, Academic Press, 1974a.
Spreiter, J. R., and A. W. Rizzi, Aligned magnetohydrodynamic solutions for solar wind flow past the earth's magnetosphere, *Acta Astronautics*, *1*, 15-35, 1974b.
Spreiter, J. R., and S. S. Stahara, Gasdynamic and magnetohydrodynamic modeling of the magnetosheath: A tutorial, *Adv. Space Res.*, *14*, (7)5 - (7)19, 1994.
Spreiter, J. R., and A. L. Summers, Effect of uniform external pressure and oblique incidence of the solar wind on the terminal shape of the geomagnetic field, NASA TR-R-181, 1963.
Spreiter, J. R., and A. L. Summers, Dynamical behavior of the magnetosphere boundary following impact by a discontinuity in the solar wind, *J., Atmospher.Terrest. Phys.*, 27, 357-365, 1965a.
Spreiter, J. R., and A. L. Summers, On the stability of the boundary of the geomagnetic field, NASA TR-R-206, 1965b.
Spreiter, J. R., and A. L. Summers, On conditions near the neutral points on the magnetosphere boundary, *Planetary Space Sci.*, *14*, 787-798, 1967.
Spreiter, J. R., A. L. Summers, and A. Y. Alksne, Hydromagnetic flow around the magnetosphere, *Planetary Space Sci.*, *14*, 223-253, 1966a
Spreiter, J. R., A. L. Summers, and A. W. Rizzi, Solar wind flow past nonmagnetic planets—Venus and Mars, *Planetary Space Sci.*, *18*, 1281-1299, 1970a.
Spreiter, J. R., and S. S. Stahara, A new predictive model for determining solar wind-terrestrial planet interactions, *J. Geophys. Res.*, *85*, 6769-6777, 1980a.
Spreiter, J. R., and S. S. Stahara, Solar wind flow past Venus - Theory and observations, J. *Geophys. Res.*, *85*, 7715-7739, 1980b.
Spreiter, J. R., and S. S. Stahara, Computer modeling of solar wind interaction with Venus and Mars, in *Venus and Mars: Atmospheres, Ionospheres, and Solar Wind Iinteractions*, edited by J. G. Luhmann, M. Tatrallyay, and R. O. Pepin, AGU Monograph *66*, pp. 345-383, 1993..
Spreiter, J. R., and S. S. Stahara, Gasdynamic and magnetohydrodynamic modeling of the magnetosheath: A tutorial, *Adv. Space Res. 14*, (7)5-(7)19, 1994.
Stahara, S. S., D. S. Chaussee, B. C. Trudinger, and J. R. Spreiter, Computational techniques for solar wind flows past terrestrial planets - Theory and computer programs, NASA CR 2941, 1978.
Stahara, S. S., D. Klenke, B. C. Trudinger, and J. R. Spreiter, Application of advanced computer procedures for modeling solar-wind interactions with Venus, NASA CR 3267, 1980.
Stahara, S. S., R. R. Rachiele, J. R. Spreiter, and J. A. Slavin, A three dimensional gasdynamic model for solar wind flow past nonaxisymmetric magnetospheres: application to Jupiter and Saturn, *J. Geophys. Res.*, *94*, 13,353-13,365, 1989.
Sturrock, P. A., and R. E. Hartle, Two-fluid model of the solar wind, *Phys. Rev. Letters*, *16*, 628-631, 1966.
Sturrock, P. A., and J. R. Spreiter, Shock waves in the solar wind and geomagnetic storms, *Geophys. Res.*, *70*, 5345-5349, 1965.
Van Dyke, M. D.., The supersonic blunt-body problem – Review and extensions, *J. Aero-space Sci.*, *25*, 485-496, 1958.
Van Dyke. M. D., and H. D. Gordon, Supersonic flow past a family of blunt axisymmetric bodies, NASA TR-R-1, 1958.
Waniek, R. W., and G. H. Kasai, Interaction of a plasma flow with a three-dimensional magnetic dipole, in *Proceedings of the 7th International Conference on Phenomena in ionized Gases, Belgrad*, *2*, 209-214, 1966.
Wolfe, J. H., R. W. Silva, and M. A. Myers, Observations of the solar wind during the flight of IMP-1, *J. Geophys. Res.*, *71*, 1319-1340, 1966.

Wolfe, J. H., R. W. Silva, D. D. McKibbin, and R. H. Mason, Preliminary observation of a geomagnetospheric wake at 1000 earth radii, *J. Geophys. Res.*, *72*, 4577-4581, 1967.

Zhigulev, V. N., and E. A. Romishevskii, Concerning the interaction of currents flowing in a conducting medium with the earth's magnetic field, *Doklady Akad. Nauk SSSR*, *127*, 1001-1004, 1959 [English translation: *Soviet Physics, Doklady 4*, 859-862, 1960].

J. R. Spreiter, 1250 Sandalwood Lane, Los Altos, CA. 94024.

S. S. Stahara, RMA Aerospace, Inc. 10432 Noel Ave., Cupertino, CA.

Early Ground Based Approach to Hydromagnetic Diagnostics of Outer Space

"Facts which at first seem improbable will, even on scant explanation, drop the cloak which has hidden them and stand forth in naked and simple beauty."

Galileo Galilei
Dialogues Concerning Two New Sciences {1638}.

V.A.Troitskaya

NASA, Goddard Space Flight Center, Greenbelt. Maryland, 20771,
(On leave from La Trobe University, Bundoora, Victoria, Australia, 3083.)

INTRODUCTION.

Looking back 40 years, I still wonder at how fortunate it was, that the birth of the internationally organized and coordinated studies of electromagnetic pulsations coincided with the International Geophysical Year (IGY). These studies revealed a wealth of information about the state of outer space, which could be checked on a global network of ground-based stations, most of which were organized by the beginning of the IGY. Moreover with the advent of the sputnik-satellite era, these results could be verified by measurements "in situ." Of course pulsations of the magnetic field were recorded and analyzed well before the IGY, but the comparison of results, and the communications among the scientists involved, were rare, partly because the dates of these excellent publications differed by tens of years [*Stewart*, 1861; *Angenheister*, 1912; *Terada*, 1917; *de Moidrey*, 1917; *Sucksdorf*, 1932; *Harang*, 1936] and partly because of the absence of a generally accepted classification of pulsations.

From the beginning of the fifties, interest in pulsations of the natural electromagnetic field grew significantly in the former USSR, mainly due to a number of relevant problems, that required information about their frequency range, their distribution in time, and the regularities of their occurrence in different regions of the Earth (E.g., magneto-telluric sounding of the Earth's crust, the search for precursors of earthquakes, the distinguishing in the context of classified problems, of some small and short-lasting signals on a background, which consisted of the geomagnetic and Earth's current pulsations). In the first half of the fifties also, appeared pioneer papers on the existence of plasma surrounding the Earth [*Storey*, 1953], and soon after that, theory of oscillations in plasma magnetized in a dipole field like the Earth's [*Dungey*, 1954]. In the latter paper, for the first time, magnetic pulsations observed at the Earth's surface, were interpreted as oscillations of entire flux tubes. As a result of this, a significantly broader scientific community became interested in these tiny oscillations of the magnetic field, hardly seen on standard records of observatories. Beginning from the second half of the fifties, the number of publications on magnetic pulsations rose dramatically, and at the end of the fifties, this direction of research was established in the ambit of the International Association of Geomagnetism and Aeronomy (IAGA).

Being asked by the editors, to give some information about my own schooling before exposing my personal view of my work and the work of my colleagues, "roughly in the period 1957-1967", I would like to say the following: I was a pupil in one of the best schools of Leningrad - "Peterschule" which required being fluent in the German language before admission. All subjects in this school were given in German and it

Discovery of the Magnetosphere
History of Geophysics Volume 7

turned out that knowledge of this language happened to be extremely useful during the Second World War. After graduating I passed the entrance examinations of the Physical Faculty of one of the oldest Universities of Russia -Leningrad (Petersburg) University. Many prominent scientists lectured us in Mathematics (e.g., Prof. V. Smirnov.), Physics (Prof. Fock, Prof. L. Artzimovitch, Prof. I. Tamm) and I was carried away by physics and decided to continue my education after graduating with honors from University. I got an offer to do a PhD in Geophysics and started my studies, which however, soon were interrupted for a long time, first by the World War-2, blockade of Leningrad, and then by family problems. During this time I worked as a teacher of German language in the highest military academy and then participated in the work of a naval institute. The Laboratory in which I worked dealt with problems of clearing the Finnish Strait and Baltic Sea of German mines. This required working on a submarine during last year of the war and was far from safe. Only in 1950, could I return to science, and passed once more, the examination for a position of PhD at the Geophysical Institute in Moscow with Academician A. N. Tichonov as my mentor. There I presented my PhD thesis and then the state DSc degree - both of them dedicated to geomagnetic pulsations.

I have to admit that since the beginning of my studies, which were concerned mainly with relevant problems, I was constantly amazed by the great variety and real beauty of oscillations which appeared on our highly sensitive records. Quite naturally the curiosity of myself and of my colleagues led us from relevant problems to the problems of the origin of these oscillations and consequently to the physics of the magnetosphere. Organizationally, this transition of the direction of studies was not easy, because the interest and usefulness of this research during the first half of the fifties, was still questionable, especially within the charter of the Institute of the Physics of the Earth (Moscow), where we were working. Moreover these studies required significant financial support. The Institute, however, was interested mainly in solid Earth geophysics, and I doubt whether we would have been able to develop so widely the investigations of pulsations of the magnetic field of the Earth, if not for the unprecedented development and funding of geophysics, which took place during preparations for, and carrying out of, the program of IGY. It became clear that the development of studies of magnetic pulsations, in which my colleagues and I were soon engaged, required the establishment of a wide network of stations. In order to register all period ranges occupied by pulsations (from fractions to thousand of seconds) the records at these stations had to be approximately by two- three orders of magnitude more sensitive, than the standard records of magnetic field. As a result of great efforts, at the beginning of IGY, registration of pulsations was established at 17 stations. Five of these stations were located in the Arctic and the other 10 distributed in different longitudes and latitudes over the vast territory of the USSR. In addition, observations of pulsations were implemented at two stations in Antarctica which later were complemented by the organization of pulsations observations at the geomagnetic pole (Vostok). For development of this network of stations it was necessary to obtain huge financial support, to build equipment and stations, to get substantial amount of staff, and train it. Till now I am wondering how we succeeded to fulfill it in such short time (around one and a half years). It required the selfless, dedicated work of many of my colleagues. The central figure in this activity was L. N. Baranskyi. Drs. K. Zybin*, N. Maltzeva* and R. Schepetnov were responsible for the work of newly built observatories Borok, Lovozero (Kola region) and Petropavlovsk Kamchatsky. The records of vertical component of magnetic field by means of big loops on the Earth surface were supervised by Dr. G. N. Petrova and Prof. A. G. Kalashnikov*. Low latitude pulsations studies were established in cooperation with India and Cuba. In the first decade of IGY, with the cooperation of French scientists there were organized observations in the conjugate points Sogra (Archangelsk region)-and Kerguelen (French island in the Indian ocean). These stations were located in subauroral regions of the Northern and Southern hemispheres, respectively. In cooperation with American geophysicists, joint studies of pulsations were conducted in the Arctic and the Antarctic, including studies at the geomagnetic poles Thule and Vostok. Temporary simultaneous observations were carried out between antipodal points Dallas (USA) and Garm (Tadjikistan). The main central observatory for pulsation studies, Borok, was built in the Jaroslavskaya region, some 300 km North-East from Moscow. To Borok, records from many stations were sent, and there an archive of quick run records of magnetic and Earth currents was established. At the beginning of IGY the pulsations were recorded mainly in Earth currents, which later were replaced by measurements of the magnetic field using induction fluxmeters of great sensitivity (of order $\sim 10^3$nT). The processing of data was really back-breaking work. At that time we had no magnetophones, no special analyzers, in order to extract information about pulsations in different frequency and amplitude ranges. Therefore we introduced at each station registrations on photopaper (with appropriate sensitivity) on three time scales: (1) usual -20 mm/hour, (2) quick run -90 mm/hour, for studying mainly Pc3-4 - and Pi-2, and (3) ultra quick -30 mm/min, for

*deceased

studying the shortest part of the pulsation spectrum -less then 15 sec period. The amount of information we obtained by comparison of these records, not only with each other, but also between different stations was in significant part new and immense. And so was the amount of time and efforts spent on these studies.

The first results of these investigations, especially for different types of pulsations with periods in the range from fractions of a second to approximately 15 sec., were met with great skepticism. We were even accused of measuring artifacts and not natural phenomena and that we could not distinguish one from the other. When for instance I showed the beautiful examples of series of pulsations with periods around 2 sec. and amplitudes of several tens of milligammas (called "pearls", because their series reminded of a necklace of pearls) in central institute for geomagnetism (IZMIRAN, Moscow) - the very existence of them, even there, was met with mistrust. In some way, detection of the effects of USA high altitude nuclear explosions (1958, code-named "ARGUS"), which produced oscillations in the same frequency range as short period pulsations, helped in the acknowledgment of the scientific and practical value of their studies. Unfortunately these results could be published only in 1960 [*Troitskaya*, 1960] after they were declassified, which also required significant efforts. The analysis of morphology of magnetic pulsations, their correlations with other geophysical phenomena, first obtained by comparison with results of other ground based measurements, and then with data obtained by satellites determined the fascination of my colleagues and myself with magnetospheric physics. In this paper, after a short description of pulsations and their classification, I shall present some of the discoveries, concerning the magnetosphere and the solar wind, which were obtained using ground based observations of pulsations during approximately the first decade after the beginning of IGY. These results and their successful development led to the formulation of a new direction of research - "Hydromagnetic diagnostics of the magnetosphere and solar wind."

CLASSIFICATION OF MAGNETIC PULSATIONS.

It may seem strange, that in this paper I mention such a specialized topic, but historically, the adoption of a classification scheme played a crucial role, being a specific new "language", which facilitated communication between scientists engaged in different aspects of pulsation studies. Moreover, with the development of solar terrestrial physics, this classification began to be used for oscillations occurring in the ionosphere, the magnetosphere and the solar wind. The concept of two main types of oscillatory regimes (Pc and Pi) was introduced by *Troitskaya* [1953 a,b] This concept was used as a base for classification accepted 10 years later by IAGA. In this classification pulsations were divided into regular continuous, usually stable, oscillations (Pc), and short lasting, usually unstable, irregular oscillations (Pi). Table 1 gives the subdivision of these two types of pulsations by periods.

The most difficult task in compiling this classification was to establish the limits for each of the subdivisions. However, it was solved, drawing upon the intuition acquired by specialists, who in the course of their studies looked through huge amounts of records, and the possibility of exchanging opinions at special meetings, organized before the IGY by Committee 10 of IAGA, and chaired by Father A. Romana. This led to the establishment of the period ranges for each type of pulsations which, in essence, survived even into the nineties. This classification was officially adopted by IAGA in 1963 on the recommendations of a small group of people, and published in 1964 [*Jacobs et al.*, 1964]. Of course this classification should be reconsidered in the light of achievements acquired since its establishment, especially in the range of longer periods, up to 1000 sec. Also different principles could be used, as a base for classification, generic or correlative, taking into account the information which has been obtained on the origin of different pulsations, or their correlations with other geophysical phenomena. Another variant of classification, based on frequency-time dependence of pulsations (hydromagnetic emissions) was suggested by Japanese scientists, but it is applicable only to the short-period part of the pulsation spectrum, and is not widely used [*Nagata 1980 et al.*]

Finally I have to mention, that in the course of the time since IGY, the terminology, indicating these short period variations of electromagnetic field, changed from micropulsat ions to pulsations of the magnetic field, and relatively recently to ultra low frequency waves (ULF-waves).

TABLE 1.

Type	Range of periods (seconds)
Continuous pulsations	
Pc 1	0,2 - 5
Pc 2	5 - 10
Pc 3	10 - 45
Pc 4	45 - 150
Pc 5	150 - 500
Irregular pulsations	
Pi 2	40 - 50
Pi-1	1 - 40

CONNECTION OF CONTINUOUS PULSATIONS PROPERTIES WITH THE PARAMETERS OF THE MAGNETOSPHERE AND THE SOLAR WIND.

1. *Pc2-4 and the direction of the interplanetary magnetic field (IMF)*

Observations on a wide network of stations revealed that in addition to a distinct diurnal variation in local time, the simultaneous global modulation of continuous pulsations undoubtedly takes place. It can be expressed in the "sudden ending" of pulsations on a global scale, or in simultaneous change of their periods at stations separated by large distances or in simultaneous rises and diminishing of their amplitudes. In Fig.(1) are shown the records of Earth currents, from stations located in middle latitudes of the northern hemisphere encompassing 142° in longitude, one station in the Arctic and one in the Antarctic. On all of them pulsations disappeared practically simultaneously. Fig.(2) shows the simultaneous change of periods of pulsations at three middle latitude stations distributed around the globe. Fig(3) shows similar simultaneous change of regime of pulsations on the morning and evening side of the Earth at Dallas (USA) and Borok (Russia). Such cases were enigmatic, and attempts to interpret them using correlations with traditional indices characterizing the state of magnetic field, or other available geophysical data, did not produce any satisfactory result. Besides the puzzle of global modulation, there were "mystical days", when continuous pulsations were absent all around the world. Such days occurred relatively seldom (1-2% of all observed cases), but this fact did not fit the prevailing theories of that time. These theories presumed that Pc pulsation excitation is primarily produced by the continuous flow of solar wind at the magnetospheric boundary, with subsequent generation of resonance oscillations. Therefore one would expect, that they should occur every day, because it was hard to imagine the sudden stopping of solar wind flow. That led to the thought that some changes in the parameters of a distant source of continuous pulsations take place on such "empty" days, which should be crucial for their excitation. Discovery of this parameter was made in the course of comparison of one of the first available detailed

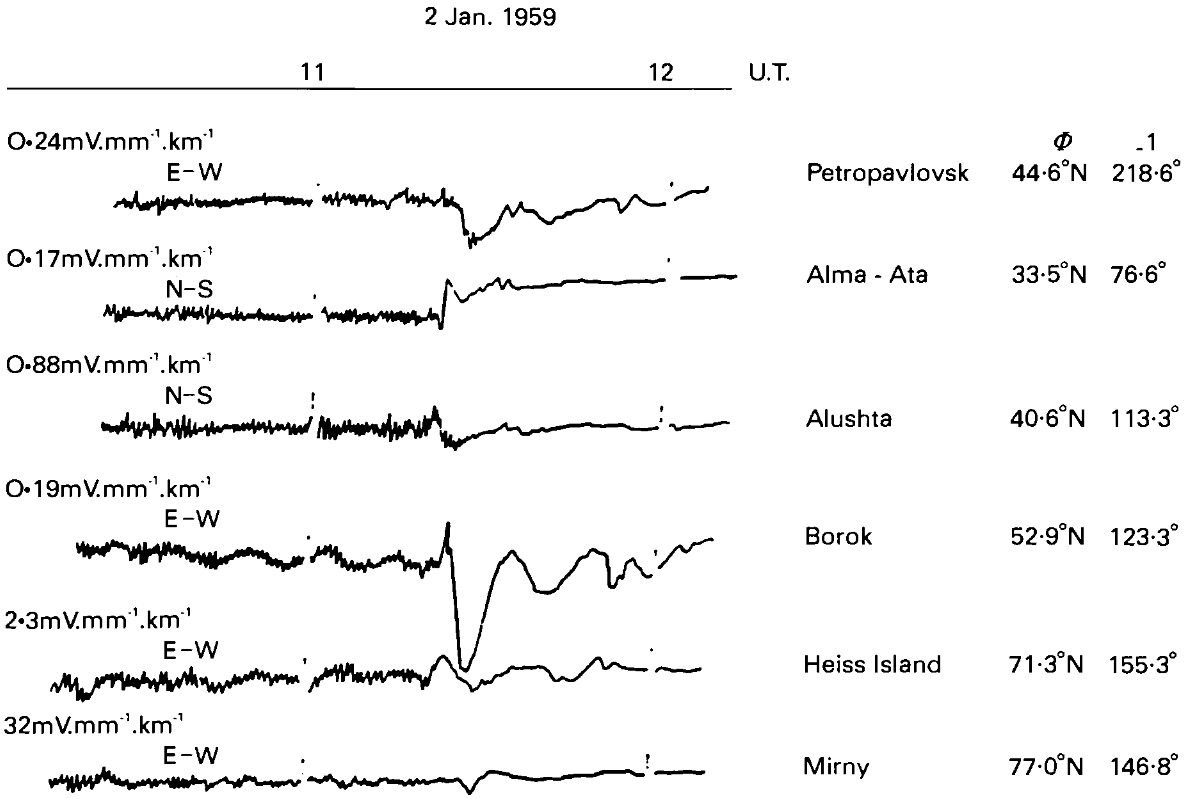


Figure 1. Example of global modulations of continuous pulsations at stations with longitude difference ~105 °and at latitudes extending from 71.3 °N to 77°S. Shown are records of East-West components of Earth currents at four middle latitude and two polar stations. On the left side of the figure is given the sensitivity of records of Earth currents at each station, on the right side the names of the stations and their coordinates, f and l.

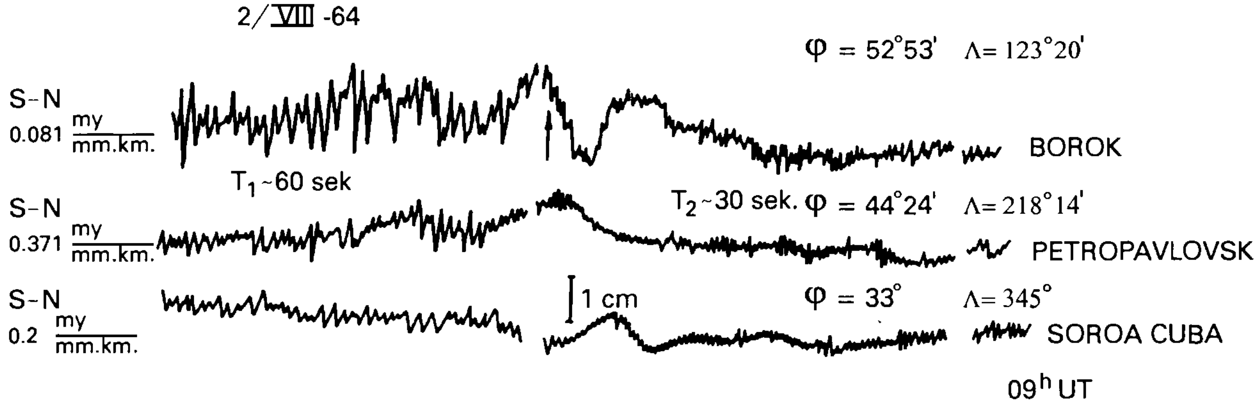


Figure 2. Simultaneous change of periods of continuous pulsations around 8 hours UT from 60 sec. to 30 sec. at three-middle latitude stations, encompassing in longitude the whole globe-Borok (300 km. North- East from Moscow), Petropavlovsk on Kamchatka, and Soroa (Cuba).

data sets of the IMF (for the fall of 1963, kindly provided by N.Ness) with pulsation records. In this set of data, averaged for every 5,46 min, there were days when the predominant direction of the IMF was perpendicular to the Sun-Earth line. To our amazement and delight, the ground based records of continuous pulsations obediently followed the changes in the directions of the IMF. As soon as direction of the IMF turned close to 90 degrees to the Sun-Earth line, the pulsations disappeared. In Fig.(4 a) are shown copies of records of Earth currents for October 12, 1963 and for 14 December 1963 at the station Petropavlovsk Kamchatsky, together with the orientation of the IMF, shown by arrows. The local hours for these cases were around noon, that is, they corresponded to hours of maximum of Pc occurrence. Nevertheless, suddenly the pulsations disappeared, following the change in the direction of IMF. Disappearance of pulsations occurred also on other stations of the network. Such, and many other examples, were the first striking evidence of connections of Pc3-4 occurrence with orientation of the IMF. Fig.(4 b) shows the distribution of all directions of IMF for the available set of data (around 1500 cases) averaged for each consecutive 5.46 min separately for simultaneously observed (1) Pc3, (2) Pc3-4, and (3) absence of Pc3. This picture, in addition to direct comparisons, gave overwhelming statistical evidence of the dependence of continuous pulsation excitation on the direction of the IMF first described in the paper of *Bolshakova and Troitskaya* (1968). The maximal growth rate of Pc3-4 waves upstream of the bow shock which were assumed to penetrate to the ground was estimated by *Kovner et al.* (1976) and was found to occur for the angle which was close to the value obtained from comparison of ground based data to the orientation of the IMF. This result was met with

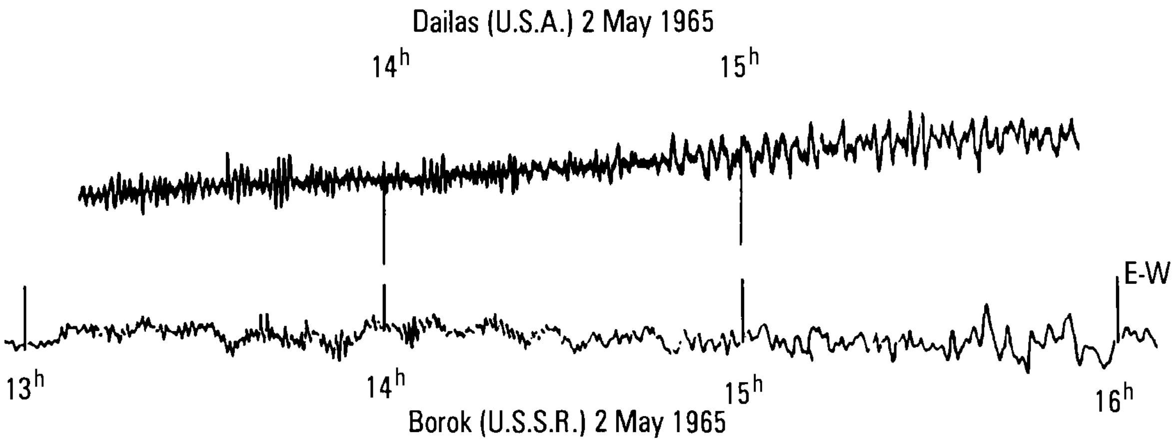


Figure 3. Simultaneous records of continuous pulsations in Dallas (USA, local morning) and in Borok (Russia, local evening) on 2 May 1965. The change of regime of continuous pulsations at both stations is clearly seen.

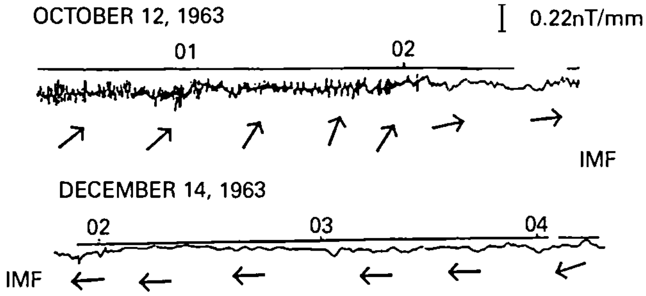


Figure 4 (a). Records of pulsations at the stations Petropavlovsk-Kamchatsky (upper part of the figure), and orientation of IMF in plane of ecliptic shown by arrows. Direction to the sun is upward.

great skepticism in the West and in the East (Japan). When I remember these times, I think that the polite mistrust and doubts with which it was accepted, and which lasted many years, was probably a part of the general situation, when the results of ground based space research were in general considered as second class. However when the waves upstream of the bow shock were measured it was soon discovered, that they follow the same law of modulation of their amplitudes as continuous pulsations on the ground (mainly Pc3) *Greenstadt* 1973; *Greenstadt and Olson*, 1977; *Greenstadt et al.*, 1979. In Fig(5) are presented results which were obtained significantly later but which further illustrate the validity of originally discovered dependence. They show 1) data obtained in 1979 on the spacecraft ISEE-1 of the amplitude of oscillations of upstream waves and the direction of the IMF at the distance 20 R_e; 2) the corresponding behavior of Pc3 amplitudes at stations, widely separated in latitude and longitude [*Troitskaya and Bolshakova*, 1988]. The data from ISEE-1 were kindly provided by C. T. Russell. The velocity of solar wind was constant in this example. This fact is important, because, as it was shown by *Saito* [1964]

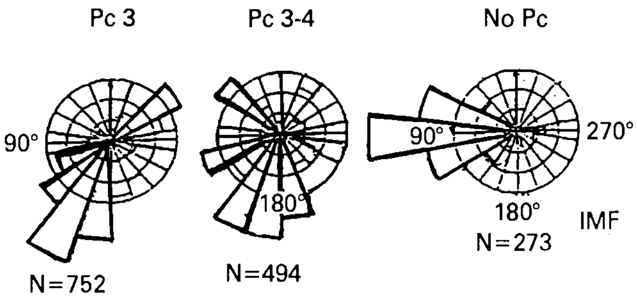


Figure 4 (b). Distribution of directions of IMF for analyzed period organized by three plots: The first (752 cases) corresponds to occurrence of Pc3, the second (494 cases) corresponds mainly to Pc 3-4 regimes, and the third (273 cases) shows the situation when Pc3 are absent. Each case corresponds to the average value of IMF over consecutive 5.46 minutes.

and *Vinogradov and Parkhomov*, [1970] the amplitude of continuous pulsations depends also on solar wind velocity.

Fig.(5) clearly shows that the modulation of amplitudes of waves upstream of the bowshock, and of continuous pulsations at the three stations on the ground follow the same pattern.

The external source for dayside Pc 2-4 [mainly upper part of Pc2, range of periods, Pc3, and lower part of Pc4 period range] is now generally accepted, but the concepts about the linkage between waves internal and external to the magnetosphere, that is, the mechanism of coupling of waves generated in the interplanetary space to those observed on the ground, are still under discussion.

Pc3-4 periods, the density and the value of IMF

Interesting results, stressing the difference of the dependence of Pc2-3 and Pc4 on parameters of the solar wind were shown in the paper of *Gringauz et al.*,[1971]. In this paper

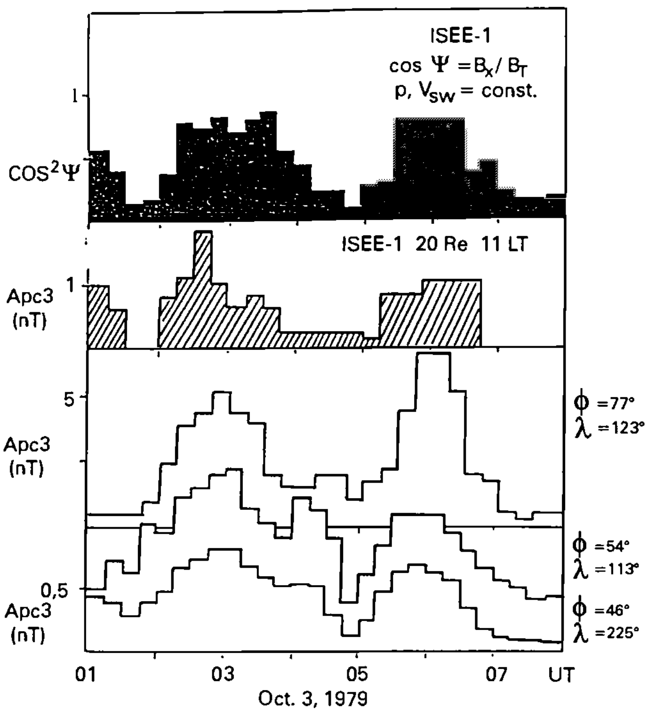


Figure 5. Example of global modulation of amplitudes of Pc3 regime at widely separated ground based stations (lower part of the picture) corresponding (with appropriate time delay) to the amplitude modulation of upstream waves and changes of directions of IF measured at distance ~ 20 R_e and at local time ~ 11 LT. by SEE-1 (upper part of the picture). B_x is the component of B towards the Sun, BT is the amplitude of B. On the vertical axis (below) is given the scale for amplitudes of continuous pulsation's on the ground. On the horizontal axis is time in UT.

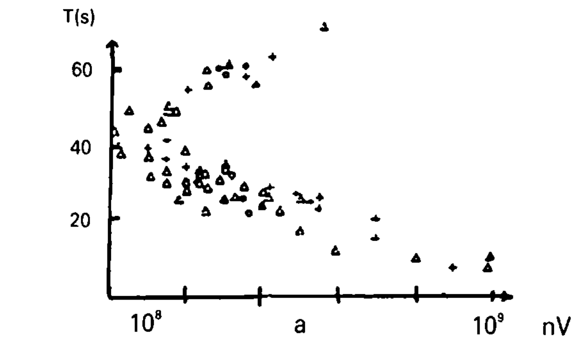


Figure 6 (a). Dependence of Pc 2-4 periods observed on 3 ground based stations (Borok, Petropavlovsk on Kamchatka and Cuba) encompassing the Earth on solar wind flux -nV (part cm-^{2} s^{-1}) obtained using data from Veneras 2-4, 5-6.

the periods of Pc2-3-4 were compared with the density of solar wind fluxes, measured on Venus-2 (1965), Venus-4 (1967), Venus 5 and 6 and on IMP-1 (December 1963 to March 1964). Pulsation periods were determined for three observatories: Petropavlovsk-Kamchatsky, Soroa (Cuba), and Borok, (i.e., practically around the world in longitude, for middle and low latitudes). The results of comparison are shown in Fig.(6 a), where pulsations periods are plotted as a function of ion flux (nV cm^{-2} sec^{-1}) of the solar wind. The dependence of Pc-2-3 on nV is opposite to that observed for Pc-4. This result was unexpected and presented another hint of the different nature of Pc3 and Pc4. At the same time it gave support to the empirically postulated boundary dividing the classes of Pc-3 and Pc4 at ~40-45sec. In order to check the results shown in Fig.(6 a), the same comparison was done using data of IMP-1. The results are given in Fig.(6 b). Two different dependencies of Pc periods on nV were confirmed,

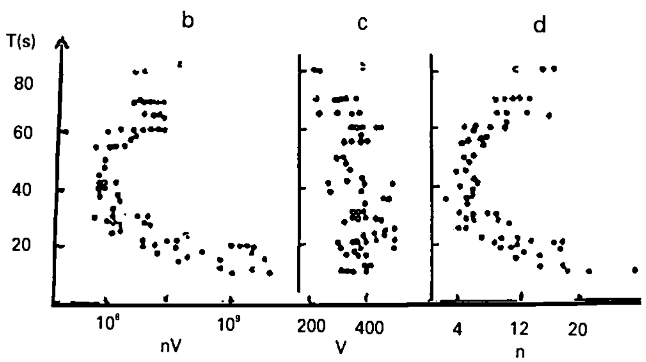


Figure 6 (b). The same dependence of periods of Pc2-4 using data of IMP-1 from December 1963 to March 1964.

Figure 6 (c). Dependence of Pc2-4 periods on solar wind velocity for fixed value of density n.

Figure 6 (d). The same dependence on n for fixed value of V. The symbol (^) in the figures corresponds to the station Borok, (o) to Soroa (Cuba) and (+) to Petropavlovsk.

one for T < 40-45 sec (T diminishing with the growth of nV) and the other for T > 40-45 sec (T rising with the growth of nV). Fig.(6 c) shows the dependence of Pc period on changing solar wind speed (V), for fixed value of n, and Fig.(6 d) gives the dependence of Pc periods on changing value of density of the solar wind n, for fixed value of the solar wind speed. It is seen that the change in Pc periods is influenced mainly by changes of density n. The dependencies obtained indicated the possibility of simultaneous generation of two superposed Pc regimes, which inversely depended on the density of the solar wind. Such cases of superposed regimes are often observed on ground based records, and present a serious difficulty in selecting the periods of oscillations, for instance, in the studies of connection of Pc periods with the parameters of IMF.

The dependence of Pc3-4 periods on the magnitude of the IMF

This dependence was discovered in 1971 [*Troitskaya et al.*, 1971]. It was experimentally and theoretically developed in papers [*Gul'elmi. et al.* 1973, *Gul'elmi A.V. and V.A. Troitskaya*, 1973, *Gul'elmi*, 1974; *Plyasova-Bakunina*, 1972. In the paper of *Gul'elmi et al.*, [1973] is presented one of the experimentally obtained dependencies of Pc 2-4 periods on the value of the IMF, see Fig.(7). The scatter of points is given by the solid lines. The investigations of this dependence and the first results, were based on data of IMF obtained by IMP-3 for September-November 1966, and by IMP-4 for August-November 1966. For ground based data, observations from Borok observatory were used . The results of these investigations showed an inverse dependence of Pc2-4 periods on the magnitude of the IMF. First an empirical relation between these quantities was established in the form [*Troitskaya et al*, 1971]

$$T = \frac{160}{B_{IMF}(nT)} \tag{1}$$

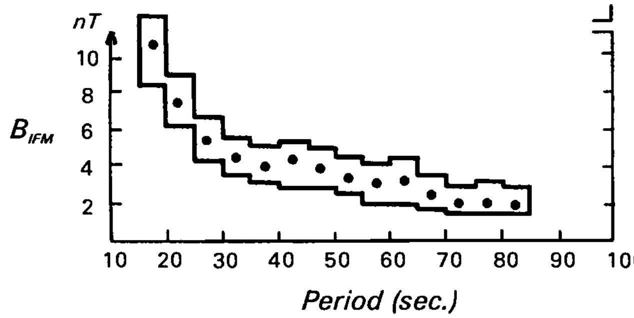


Figure 7. Dependence of periods of Pc 2-4 on the value of IMF. On vertical axis is given the magnitude of IMF in nT, on the horizontal, the periods of continuous pulsations.

The first conclusion from this relation was the confirmation of previously suspected difference between Pc 3 and Pc 4 pulsations. It became clear that not all range of periods of Pc 4 can obey this empirical formula. Indeed oscillations with periods greater than 60-80 seconds should correspond due to (1) to the value of IMF less than ~ 3 -2 nT, which occur rather seldom. The second obvious conclusion from this relation was the supposition that one should find some process in front of the bow shock, which could be responsible for generation of waves with periods falling in the range of continuous pulsations. In several papers [*Troitskaya et al.*, 1971; *Gul'elmi and Troitskaya*, 1973; *Gul'elmi* 1974] the idea of the possibility of excitation of continuous pulsations by the cyclotron instability of protons reflected from the bow shock was developed. Such particles were already observed on satellite Vela [*Asbridge et al.*, 1967] and on Explorer-34 there were discovered hydromagnetic waves generated by these particles [*Fairfield*, 1969].

In the course of time the relation (1) was usually presented as

$$f(MHz) = cB_{IMF}(nT) \qquad (2)$$

Here "c" is the constant which, for the first time, was determined using data of the station -Borok and spacecrafts - IMP 3 and 4. The value of "c" obtained as the result of this investigation was ~ 6.15 mHz/nT. Determinations of the value of "c" were later carried out in many studies and the results obtained in 15 different investigations using ground based data are summarized in table 1 of the paper [*Troitskaya and Bolshakova*, 1988]. The mean value of "c" deduced from these 15 values was 5.9 ± 0.4 mHz/nT. This was in remarkable agreement with the first determination of "c." This relation was used in studies conducted on satellites in more conventional units of Hz and nT, in which it can be written, for ground based determined value of "c", as ~

$$f(Hz) = 0.006(nT) \qquad (3)$$

Direct measurements of wave frequencies and values of IMF in the upstream region of the interplanetary medium around the Earth [*Russell and Hoppe*, 1981; *Russell and Hoppe* 1983, *Russell*, 1994] gave the following relation between these quantities

$$f(Hz) = 0.0058B_{IMF}(nT) \qquad (4)$$

It was really far beyond our expectations, working with the data of Borok station, that these two expressions-one from the ground and the other, from interplanetary space upstream of the bow shock, would practically coincide! But even more remarkable, was the discovery that for upstream waves in front of the bow shocks of all planets visited by spacecraft up to date, the same relation holds between B and f as that obtained from ground based data in 1971 [*Troitskaya V. A. et al.*] is valid (Fig.8) [*Russell*, 1994].

DISCOVERY OF INTERVALS OF PULSATIONS WITH DIMINISHING PERIODS (IPDP).

In the beginning of the IGY a morphologically and physically interesting phenomenon (IPDP) was discovered in the range of pulsation periods from approximately 10 seconds down to fractions of a second. All great storms, observed during the IGY (1957-1958) contained one or several (up to four) IPDP. They occur around the main phase of the storm, with greatest intensity in local evening or night hours, and can be traced practically simultaneously and distinctly at stations located as far away as 70°-80° of longitude. Their beginning corresponds to intense Pi-1 pulsations which turn over to the sequence of Pc-1 ("Pearl pulsations") with diminishing periods. This important phenomenon plays a crucial role in the decay of the ring current and consequently influences the duration of the magnetic storm (and the corresponding Dst variation) The IPDP intervals could not be revealed before systematic studies of pulsations were started, because the average amplitudes of oscillations in IPDP especially for its shortest periods part, is of the order ~ 10^{-2}-10^{-3} nT, that is, they are

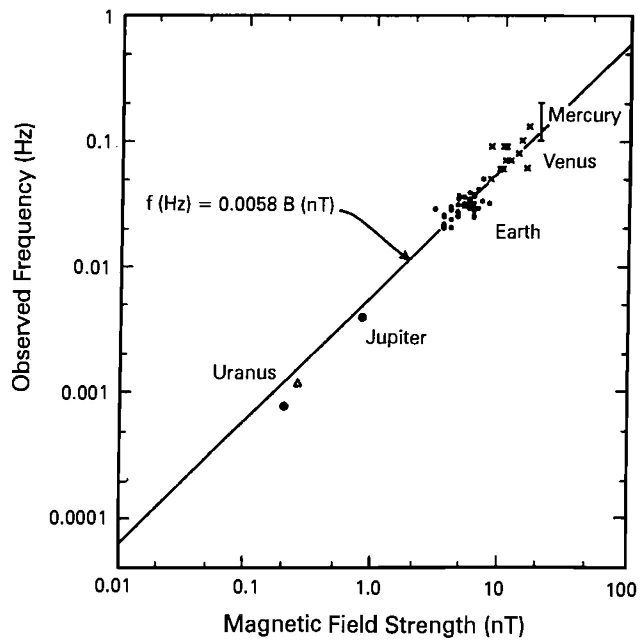


Figure 8. Dependence of the frequency of upstream waves (in Hz) on the value of IMF (in nT) for the planets-Uranus, Jupiter, Earth, Venus and Mercury.

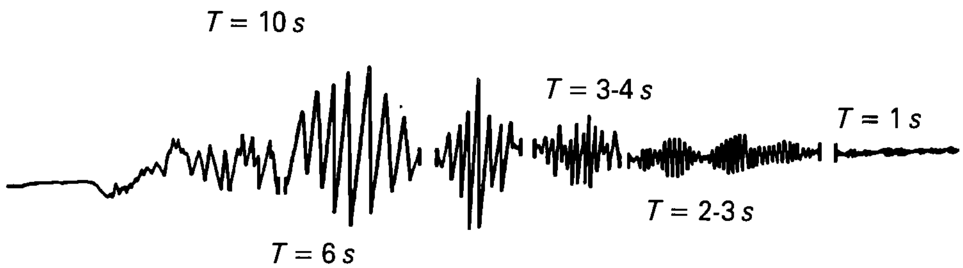


Figure 9. Stylized scheme of complete development of IPDP. Such full development encompassing the change of frequencies from ~0,1 to several Hertz is usually observed in the greatest magnetic storms.

much smaller than the sensitivity of standard magnetometers (around 1-several nT). For pulsations in the range of periods of IPDP a 24 hours registration with appropriate sensitivity and time scale (30 mm/min) was established. The occurrence of IPDP, coincides with periods of maximal intensity of disturbances in the magnetosphere and ionosphere during the storm, and can be considered as a culmination of the magnetic storm development. Fig(9) shows a stylized scheme of IPDP development in the form amplitude versus time on which wave packages of different periods follow each other. However the duration of each group of pulsations having different periods as well as the beginning and last periods of IPDP can change from case to case. The original IPDP record, an "antique" picture (in amplitude-time presentation on which it was discovered) could hardly be reproduced here due to photographic difficulties. The first publication on IPDP was presented in the paper of *Troitskaya and Melnikova* [1959], in which it was noted that the occurrence of IPDP coincided in time with the appearance of low latitude red aurora and the beginning of big disturbances in the state of the ionosphere. In Fig.10 is shown the behavior of the critical frequencies f_oF_2 (vertical axis-frequencies in MHz) for two magnetic storms which occurred in the first months of IGY The abscissa represent the local time at longitude 30° E (Moscow). Both storms had two intervals of IPDP-each one indicated by an arrow. It is seen that IPDP began to develop simultaneously with a sharp fall of critical frequencies of the F-2 layer, followed by diffuse ionospheric reflections and blackout, However both magnetic storms began several hours before severe disturbances in ionosphere and appearence of red aurora at low latitude station Alma-Ata [*Gul'elmi, A.V. and V.A.Troitskaya.*, 1973.] Investigations which were conducted later, confirmed these first results. A variety of correlations of IPDP with phenomena in the upper, and even

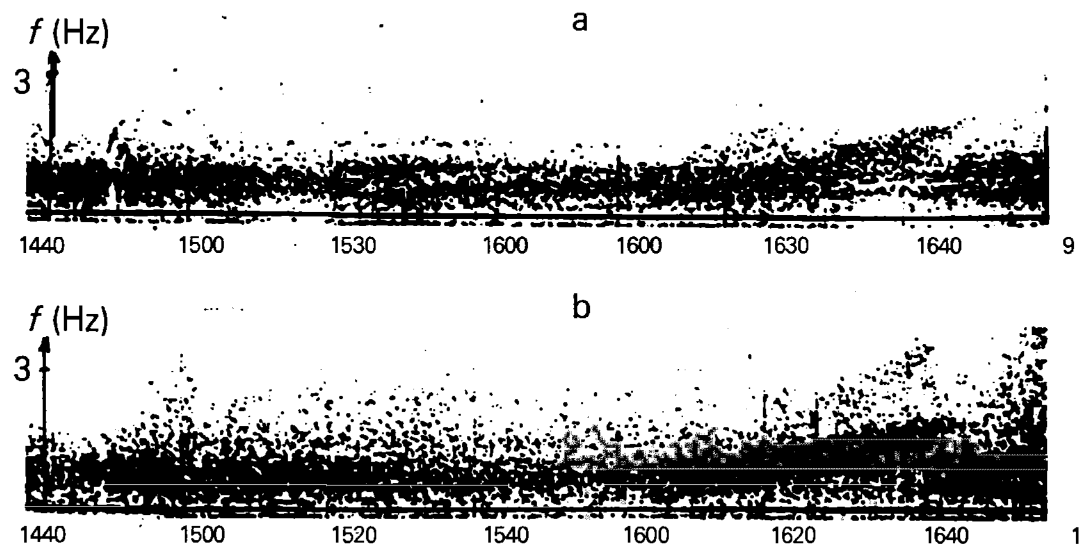


Figure 10. Presentation of IPDP in the form frequency-time at two conjugate subauroral stations Sogra (a) and Kerguelen (b). On horizontal axis is given time in UT, on vertical one - the frequency of observed pulsations. The beginning of the rise of the frequency of pulsations (corresponding to definition of IPDP) is simultaneous at both stations.

the lower atmosphere was established, such as riometric absorption, X-ray bursts in the stratosphere, and even pressure increases on microbarographs [*Troitskaya*, 1961, 1964, 1967 and references therein; *Chrzanowski et al.*, 1961]. IPDP were intensively investigated in subauroral conjugate stations, Sogra-Kerguelen, in the framework of French-Soviet cooperation. [*Gendrin & Troitskaya*, 1965; 1967, *Gendrin et al.*, 1966;]. The appearance of IPDP at these stations, shown in frequency-time presentation, is illustrated in Fig.11 It is seen that contrary to Pc-1 ("pearls") pulsations which appear intermittently in conjugate points, IPDP are observed simultaneously at both of them. Investigations of the connection of IPDP with processes in the magnetosphere showed their coincidence with large intensity changes occurring simultaneously in the radiation belts. During the flight of the second Sputnik, it was already noticed that on November 7, 1957, sharp fluctuations, (more then 50%,) in the intensity of the outer radiation belt electron fluxes occurred [*Vernov et al.*, 1958] during the development of IPDP at a number of stations. This result was confirmed by comparing IPDP appearance with the occurrence of large intensity changes of electron fluxes in the outer radiation belt during the flight of satellite Electron 1 in 1964 (Fig.12). In this figure are shown the changes of position of the boundary of the outer radiation belt and changes in the intensity of electron fluxes in it during occurrence of IPDP shown by dark stripes. The number at the side of the stripes indicates the shortest period of oscillations observed in IPDP. The conclusion from these and many other cases of IPDP, was the establishment of a close connection between the position of the boundary of the outer radiation belt and the intensity of the ring current based on the degree of development of IPDP. The smaller was the last period observed in IPDP, the greater was the shifting of the boundary of the outer radiation belt, and the loss of particles from the ring current. [*Troitskaya et al.*, 1966., *Gendrin et al.*,1967, *Troitskaya and Gul'elmi*, 1967]. At the same time, it was discovered that intense (several tens of nT) long period fluctuations of magnetic field, at the distance of about 5 Earth radii measured on Explorer 26 [*Cahill*, 1966], coincided with IPDP. In figure 13, the occurrence of IPDP is indicated by a horizontal bar under the record of the magnetic field on Explorer 26. The second bar shows the sequence of Pc-1 with changing periods resembling the final stage of IPDP [*Cahill*,1966, *Troitskaya*, 1967]. Comparison of records of magnetic field obtained on satellite GEOS-2 and of many cases of IPDP observed on the ground confirmed the coincidence of occurrence of IPDP with the large scale magnetic fluctuations in the ring current [*Troitskaya*, 1979]. The theory of

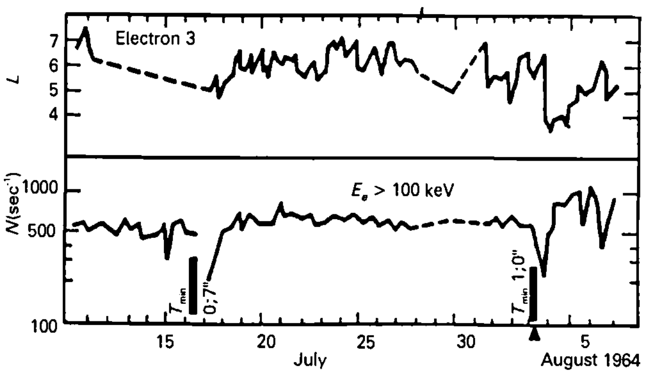


Figure 12. Intensity changes of electron fluxes (E_e > 100keV) (lower part of the figure) and changes of the position of the boundary of the outer radiation belt in L values (upper part of the figure) measured on satellite Elektron-1 in 1964. Black stripes indicate the occurrence of IPDP and T_{min} gives the value of shortest period of Pc1 observed in its development.

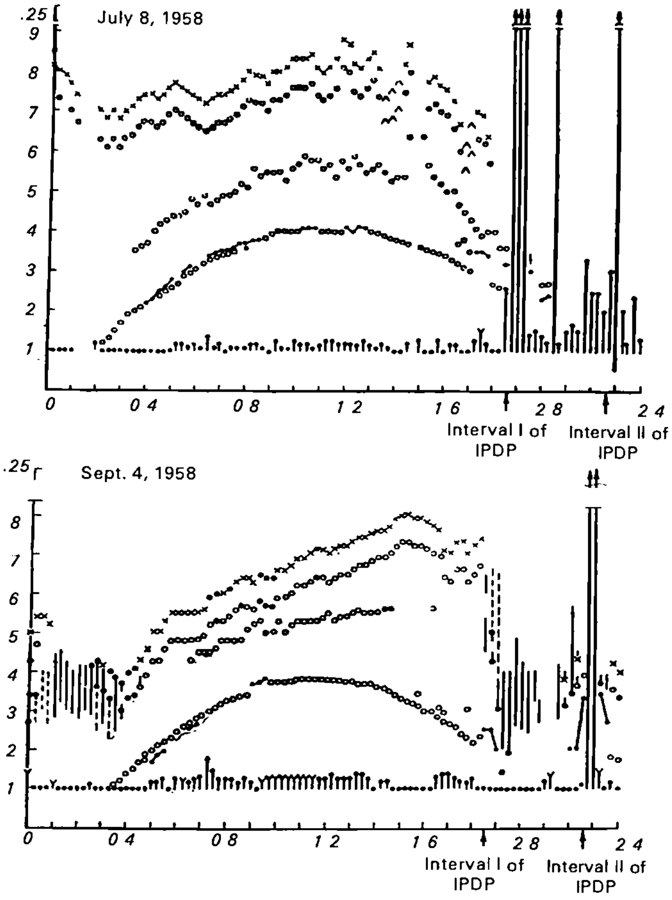


Figure 11. Processes in the ionosphere during occurrence of IPDP. In the figure are presented the plots of critical frequency, fo F-2 (in MHz), at local time of meridian 30° (Moscow) for magnetic storms which occurred at the beginning of IGY in August and September of 1957. By arrows are indicated the starts of IPDP.

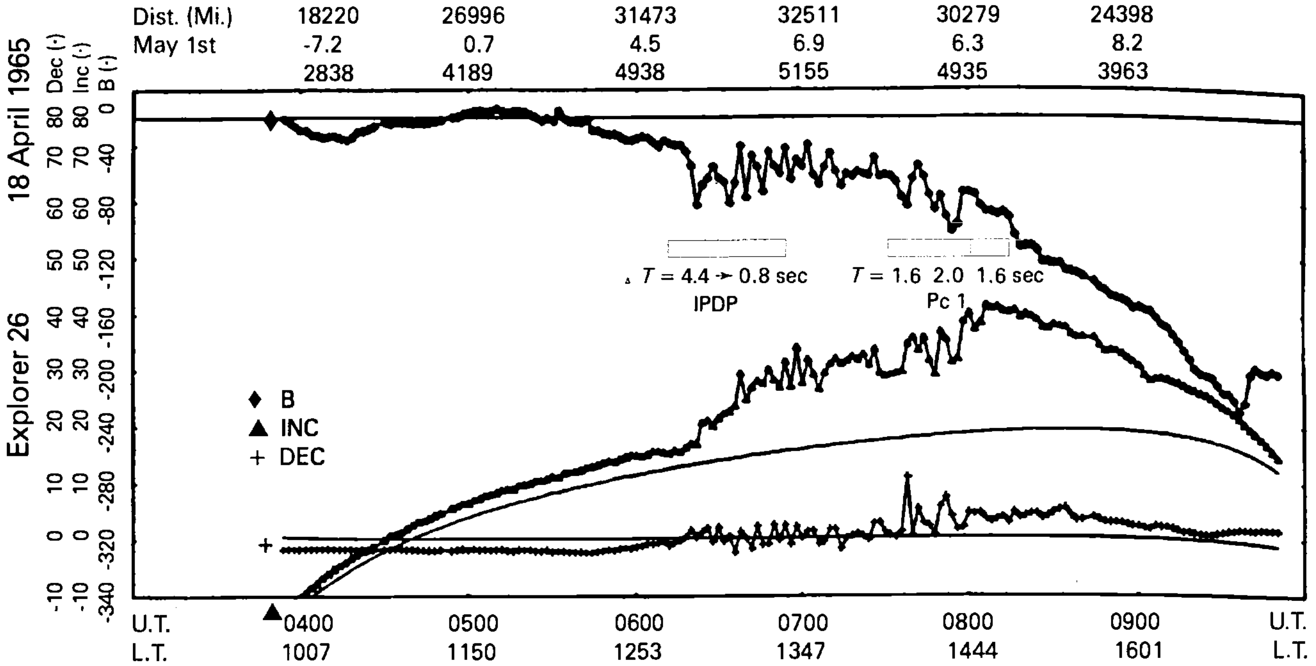


Figure 13. Intense long period variation of the magnetic field measured in the magnetosphere (L~5) by Explorer 26. Horizontal bars under the record of the magnetic field on Explorer 26, correspond: first - to the fully developed IPDP and second to the sequence of Pc-1 usually observed at its end.

IPDP was developed in various papers [*Knafflich and Kennel*, 1967; *Gendrin et al.*, 1967; *Troitskaya and Gul'elmi*, 1967; *Gul'elmi and Troitskaya*, 1973; and references therein]. IPDP were explained as a result of the ion cyclotron instability of energetic protons (10-100 kev) which develops when in the course of their westward drift they encounter the dusk side of the plasmapause. During the following years this phenomenon attracted the attention of many investigators and its morphology and theory were further developed. The results of investigations of pulsations on satellites with periods smaller than ten seconds during occurrence of IPDP were not considered here, because the main aim was to show that IPDP is a unique ground based indicator of dramatic happenings encompassing the magnetosphere, ionosphere and upper atmosphere.

The results of the pulsations studies which were obtained mainly in the first decade after the beginning of IGY, and which I chose to present, are far from a complete picture of what was achieved in this period. The subject of resonance oscillations in magnetosphere for instance was not covered here, in spite of the fact that their periods partly fall in the range of Pc 3-4. But their morphological properties, distributions in space, connections with parameters of the solar wind differ significantly from those of Pc 3-4 described in this paper. I did not mention also the results of investigations of numerous and most informative morphological properties of pulsations in middle and low latitudes, polar regions, in conjugate points, and their meridional and latitudinal profiles. All of them revealed important correlative connections of pulsations with geophysical phenomena in the upper atmosphere, ionosphere and magnetosphere. These results formed the basis for determination not only of some parameters of the magnetosphere and the solar wind, but also gave the possibility to recognize and follow the development of different processes in these unstable media. They were summed up in the monograph "Geomagnetic pulsations and diagnostic of magnetosphere" [*Troitskaya and Gul'elmi*, 1973], and in *Space Science Reviews* in the paper of *Gul'elmi* "Diagnostic of magnetosphere and interplanetary medium by means of pulsations." [1974].

CONCLUSION.

1. Two results of investigations concerned with the parameters of solar wind and its coupling with the magnetosphere were presented in this paper. Both were produced using ground based records of continuous pulsations.

The first result stated the dependence of global modulation of continuous pulsations amplitudes on the direction of IMF, upstream of the bow shock. Correspondingly there could be explained their global disappearance, even for very short intervals, by the situation when the IMF direction turned

Bolshakova, O. V., and V. A. Troitskaya, Relationship between direction of the interplanetary magnetic field and behaviour of continuous pulsations, *Dokl. Acad. Nauk. USSR.*, *180*, 343-346, 1968.

Cahill, L. J., Jr. *J.Geophys. Res.*, *71*, (19) , 4505; 1966.

Chrzanowski, P., G. Greene, K. T. Lemmon. and J. M. Young, Travelling pressure waves associated with geomagnetic activity, *J.Geophys. Res.*, *68*, (11) , 3727-3734, 1961.

De Moidrey, S. J., Pulsations magnetique a Zi-Ka- Wei et Lu-Kiapang, T*err. Mag. Atmos. Elec.*, *22*, 113-120, 1917.

Dungey., J. W., Electrodynamics of the outer atmospheres, Rep. 69, Ions.' Res. Lab. Pa. State Univ. University, Park 1954

Fairfield, D. H., Bow associated waves observed in the far upstream interplanetary medium, *J.Geophys. Res.*, *74*, 3451-3553. 1969.

Gendrin, R. E. and V. A. Troitskaya, Preliminary results of micropulsations experiment at conjugate points. *Radio Science, 69D*, 1107- 1116. 1965.

Gendrin. R. et. Troitskaya. V. A. Experience franko-sovietique effectuees en deux stations geomagnetiquement conjuguees., Revue CNFRA. 21. 1967.

Gendrin, R. E. M. Gokhberg, S. Lacourly, V. A. Troitskaya, and R. V. Schepetnov, Caracterestiques des pulsations irregulieres de periode decroissante (IPDP) et leur relations avec les variations du flux des particles piegees dans la magnetosphere. *Planetary Space Sci.15*, 1239-1259,.1967,

Greenstadt, E. W., I. M. Greene, G. T. Inoue, A. J. Hundhausen, S. J. Bame, and I. B. Strong, Correlated magnetic field and plasma observations of the Earth's bow shock, *J.Geophys. Res.*, *73*, 51-60, 1968.

Greenstadt E. W. and J. V. Olson, Pc3-4 activity and interplanetary field orientation *J. Geophys. Res.81*, 5911-5920, 1976.

Greenstadt E. W., H. J. Singer, C.T. Russell and J. V. Olson, IMF orientation, solar wind velocity, and Pc 3-4 signals. J.*Geophys. Res.*, *84*, 527- 532, 1979.

Gringauz, K. I., V. A. Troitskaya, E. K. Solamatina, and R. V. Shchepetnov, The relationship of the solar wind variables to periods of continuous micropulsations of the electromagnetic field of the earth, *Dokl. Acad. Nauk. USSR.*, *5*, 1061-1069, 1970..

Gul'elmi, A. V., T. A. Plyasova-Bakunina, and R. V. Schepetnov, Relationship of the period of gemagnetic pulsations Pc3,4 and the parameters of the interplanetary medium at the Earth's orbit, *Geomag. Aeron. 13*, 382-384, 1973.

Gul'elmi, A. V., and O. V. Bolshakova, Diagnostics of the IMF from ground-based data on Pc 2-4 micropulsations, *Geomag. Aeron. 18*, 535-539, 1973.

Gul'elmi, A. V., and V. A. Troitskaya, *Geomagnetic pulsations and the diagnostics of the magnetosphere*, *Monograph*, Nauka, Moscow, 288pp., 1973.

Gul'elmi, A. V., Diagnostics of the magnetosphere and interplanetary medium by means of pulsations, *Space Sci . Rev.*, 16, 331-345, 1974.

Gul'elmi, A. V. , T. A. Plyasova-Bakunina, and R. V. Schepetnov, On the relationship between period of Pc 3-4 geomagnetic pulsations and the interplanetary medium at Earth's orbit, *Geomagn. Aeron. 13*, 382-384, 1973.

Kovner, M. S.,V. V. Lebedev, T. A. Plyasova-Bakunina, and V. A. Troitskaya, On the generation of low frequency waves in the solar wind in the front of the bow-shock, *Planet. Space Sci.*, *24*, 261-267 1976.

Harang, L., Oscillations and vibrations in magnetic records at high latitude stations, *Terr. Mag. Atmos. Elect. 41*, 329-326, 1936.

Jacobs, J. A., T. Kato, S. Matsushita and V. A. Troitskaya, Classification of geomagnetic pulsations, *J. Geophys. Res.*, *69*, 180-181, 1964.

Knafflich, B. and J. F. Kenney, IPDP events and their generation in the magnetosphere, *Earth Planetary Sci. Letters*, *2*, 453-459, 1967.

Nagata T.,T. Hirasava, H. Fukunishi, M. Ayukawa,.N. Sato, R. Fujii. and M. K. Kawamura, Classification of Pc1 and Pi1 waves observed in high latitudes, *Mem. Natl. Inst. Polar. Res.,Special issue 16*, 56-711980.

Plyasova-Bakunina, T. A. , Effect of IMF on the characteristics of Pc 2-4 pulsations, *Geomag. Aeron.*, *12*, 675-676, 1972.

Russell, C. T., and M. M. Hoppe, The dependence of upstream wave periods on the interplanetary magnetic field strength, *Geophys. Res. Lett. 8*, 615-617, 1981

Russell, C. T., and M. M. Hoppe, Upstream waves and particles, *Space Sci. Rev.*, *34*, 155-172, 1983.

Russel, C. T. Planetary Upstream waves, Solar Wind Sources of Magnetospheric Ultra-Low-Frequency Waves. Geophysical Monograph 81, 67-74,.1994.

Saito, T., Geomagnetic pulsations, *Space Sci. Rev.*, *10*, 322-401, 1969.

Stewart, B., On the great magnetic disturbance which extended from August 2 to September 7, 1859 as recorded by photograph at Kew Observatory, *Phil. Trans. Roy. Soc. Lond.,11*, 407, 1861.

Storey, L. R. O., An investigation of whistling atmospherics, *Phil. Trans. Roy. Soc.*, *A-246*, 113, 1953.

Sucksdorff, E., Observations of rapid micropulsations at Sodankyla during 1932-1935, *Terr. Mag. Atmos. Elect.*, *41*, 337-334, 1936.

Terada, T., On rapid periodic variations of terrestrial magnetism, *J. Univ. Tokyo Fac. Sci.*, *Sect II*, *27*, 1917.

Troitskaya, V.A., Two oscillatory regimes of electromagnetic field of the Earth and their diurnal variation in Universal time, *Dokl. Acad. Nauk. USSR*, *91*, *N2*, 181-183, 1953. (a)

Troitskaya, V. A., Short period disturbances of the electromagnetic field of the Earth. *Dokl. Acad. Nauk.*, *USSR. XVI*, *2*, 211-213,1953. (b)

Troitskaya, V. A. and M. V. Melnikova, On the characteristic intervals of pulsations diminishing by periods (10-1 sec) in the electromagnetic field of the earth and their relation to phenomena in the upper atmosphere, *Dokl. Acad. Nauk. USSR.*, 228, 455-458, 1959.

Troitskaya, V. A., Effects in Earth currents aroused by high altitude nuclear explosions, *Izvest.Acad. Nauk. USSR*, *Ser. Geofiz.*, *9*, 65-70, 1960.

Troitskaya, V. A., Pulsations of earth's electromagnetic field (T < 15 sec) and their connection with phenomena in the high atmosphere, *J.Geophys. Res.*, *66(1)*, 5-18, 1961.

Troitskaya V. A. Rapid variations of the Earth electromagnetic field, in Research in Geophysics (edited by H.Odishow) I.M.I.T. Press Cambridge. Mass., pp.485-530, 1964.

Troitskaya, V. A., O. V. Bolshakova, and E. T. Matveeva, Rapid variations of the electromagnetic field as an indication of the state of the radiation belts and geomagnetosphere, *Geomag. Aeron., 6,* 292-298, 1966.

Troitskaya, V. A. Micropulsations and the state of magnetosphere in *Solar Terrestrial Physics*, eds. J. W. King and W. S.Newman, Academic Press, pp 214-274, New York, 1967.

Troitskaya, V. A.and A. V. Gul'elmi, Geomagnetic pulsations and diagnostic of the magnetosphere, *Space Sci.Rev.* 7, 689-768, 1967.

Troitskaya, V. A., R. V. Schepetnov, and A. V. Gul'elmi, Effect of sudden disappearance of geomagnetic Pc 2-4 pulsations, *Geomag. Aeron., 9,* 363-366, 1969

Troitskaya, V. A., T. A. Plyasova-Bakunina, and A. V. Gul'elmi, Relationship between Pc2-4 pulsations and the interplanetary magnetic field, *Dokl. Acad. Nauk. USSR, 197,* 13 12-1314, 1971.

Troitskaya V. A. Unpublished report, Moscow, 1979.

Troitskaya, V. A. and O. V. Bolshakova, Diagnostics of the magnetosphere using multipoint measurements of ULF waves, *Adv. Space Res., 8,* 413-429, 1988

Vernov, S. N., G. L. Grigorov, U. J. Logachev, and A. E. Chudakov, Cosmic ray measurements on Sputniks, Dok.Akad. Nauk. USSR 120 (6), (1958)

Vinogradov, P. A. and V. A. Parkhomov, Dependence of amplitudes of continuous pulsations from Solar Wind speed *Issledovaniya po Geomagnetism Aeronomii i Fisike Solntsa,* (Irkutsk) 6, 71-79. 1970

V.A. Troitskaya, NASA, Goddard Space Flight Center, Greenbelt, Maryland 20771, on leave from La Trobe University, Bundoora, Victoria, Australia, 3083.

Energetic Particles in the Earth's External Magnetic Field

James A. Van Allen

Department of Physics and Astronomy, University of Iowa, Iowa City

In the early sections of this chapter the author recalls his introduction to geophysics as an undergraduate student, his graduate work in nuclear physics, his World War II work in helping develop radio-proximity fuzes for naval projectiles and in overseeing their use in the Pacific fleet (1942–45), and his post-War scientific investigations with high-altitude rockets (1946–57). In the context of this background he then describes the development of radiation detectors by his students and himself at the University of Iowa and the successful flights of these instruments in early U.S. satellites Explorers I, III and IV (1958) and lunar probes Pioneers I, II, III and IV (1958–59). The combination of the observational data from these flights yielded the discovery and preliminary survey of the two major natural radiation belts of the Earth and of a succession of artificial radiation belts produced by high-altitude nuclear bomb bursts. In the concluding section the subsequent Iowa program in space physics research by the author's colleagues and himself is suggested by a listing of later satellite and planetary missions on which they have served as principal investigators. A brief bibliography is appended.

PREFACE

During 1981, I spent eight months at the National Air and Space Museum in Washington D.C. combing through a voluminous collection of personal notebooks and journals, and other material (unpublished and published); and writing a book-length manuscript on the early history of the magnetosphere. This opportunity was arranged by Noel Hinners, founding editor of the American Geophysical Union's *Geophysical Research Letters*, and then director of NASM. After some further work back at the University of Iowa, I completed the manuscript and it was published in 1983 by the Smithsonian Institution Press under the title *Origins of Magnetospheric Physics*. The book has been out-of-print for several years and all rights to its content have been returned to me. Therefore, in preparing the present article, I have been free to paraphrase, abridge, and/or reproduce substantial portions of that previous, well-researched account.

1. A PERSONAL INTRODUCTION

The editors of this volume invited me to describe my role in the discovery of the magnetosphere. In the interest of brevity, I might leap into this subject by starting with the launch of the first successful American satellite, Explorer I, in early 1958. But without the prior professional context that enabled me to have a central role in Explorer I's scientific achievements, I consider such a starting point to be woefully inadequate in the expository sense. The purpose of this section is to supply some background on how I came to be at the right place at the right time. Readers who are in a hurry may skip immediately to section 4.

My introduction to geophysics occurred in the summer of 1932 following my freshman year as a physics major at Iowa Wesleyan College in Mt. Pleasant, Iowa.

Discovery of the Magnetosphere
History of Geophysics Volume 7

Thomas C. Poulter, *the* professor of physics there, had been chosen by Admiral Richard E. Byrd to plan and conduct geophysical investigations on the Antarctic continent during the prospective Second Byrd Antarctic Expedition (1933–35), an extension of the Second International Polar Year (1932–33). Poulter, one of the most inspiring and creative experimentalists that I have ever known before or since that time, employed me as a summer assistant—at 35 cents an hour, payable occasionally. He taught me machine shop practice including use of a metal turning lathe and a milling machine, glass blowing, vacuum techniques, and, more importantly, the elements of original experimental research. I helped build and test a tilt-meter for recording the shifting of glacial surfaces. This was a simple and amazingly sensitive horizontal pendulum suspended so that it was in nearly neutral equilibrium. Our tests consisted of walking around the aged laboratory and recording how the pendulum responded to the shifting weight distribution on the floor.

During August 1932, I served as an observer of the Perseid meteor shower in order to determine the heights of appearance and disappearance of individual meteor trails in the upper atmosphere. The basic instrument was a circular reticle of Poulter's design. This device comprised several concentric circles, about a meter in overall diameter, and radial rods corresponding to hour markers on a clock. The reticle was connected to a small circular eyepiece by three rods to form a conical viewing device, thus providing a coordinate system of 50° angular diameter on the sky. Stereoscopic viewing was provided by one instrument in Mt. Pleasant and another in Iowa City, with a north-south baseline of 48 miles. The axes of the two conical fields of view were oriented so as to intersect at a height of about 120 km. I manned one in Mt. Pleasant and Raymond E. Crilley manned the other in Iowa City. During the night of 9 August we succeeded in observing the beginning and ending points of five time-coincident meteor trails within our overlapping fields-of-view. The observational data were analyzed and published by Charles C. Wylie, professor of astronomy at the University of Iowa [*Wylie*, 1932]. The calculated beginning points ranged from 114 to 148 km height and ending points from 84 to 105 km. The observing system was used extensively by Poulter and his colleagues during the subsequent Antarctic expedition.

Among other equipment, Poulter obtained the loan of a field magnetometer of the Department of Terrestrial Magnetism (DTM) of the Carnegie Institution of Washington. He entrusted this beautiful instrument to me to test by measuring the Earth's magnetic field at sites on the college campus and elsewhere within Henry County, and, using the theodolite on the instrument, true north and the latitude and longitude at each site. All of this I did with great care after detailed study of "Directions for Magnetic Measurements" [*Hazard*, 1930]. My results were transmitted to DTM as part of an ongoing magnetic survey of the United States.

Poulter invited me to be a member of the expedition but my parents vetoed this idea. After Poulter's departure, I continued my academic work at Wesleyan. But I followed the progress of the expedition vicariously and was an avid listener to the short-wave radio reports from Little America, conducted by means of the then remarkable equipment developed by Arthur Collins, founder of the Collins Radio Company in Cedar Rapids, Iowa (now Rockwell-Collins) and later a valued friend. One of the highlights of the expedition was Poulter's heroic rescue of Admiral Byrd from his lonely vigil at the South Pole Station [*Byrd*, 1935].

In June 1935, I graduated summa cum laude from Wesleyan, in a class of 38. At Poulter's invitation Byrd was the commencement speaker. The ceremonies were preceded by a gala parade in Mt. Pleasant, honoring these two recently returned explorers. [It was not until twenty-two years later that I myself got to the Antarctic in an expedition on the U.S.S. *Glacier*, firing balloon-launched rockets for high-altitude measurements of auroral particles and cosmic rays; and with Laurence Cahill the Earth's magnetic field in the lower ionosphere.]

I was admitted to graduate study in physics at the University of Iowa (our "family" university) in September 1935 and immersed myself in the world of electrodynamics, analytical mechanics, optics, quantum mechanics, differential equations, statistical analysis and other fundamental subjects, seemingly remote from the more tangible phenomena of geophysics. My 1936 M.S. thesis was in solid state physics with E.P.T. Tyndall. Later that year I switched to the then new field of nuclear physics and helped build a Cockroft-Walton 200 kilovolt generator, an accelerator tube, and various particle detectors. My 1939 Ph.D. dissertation was "Absolute Cross-Section for the Nuclear Disintegration $H^2 + H^2 \rightarrow H^1 + H^3$ and its Dependence on Bombarding Energy" under the guidance of Alexander Ellett, a resourceful and skilled experimentalist [*Van Allen et al.*, 1939].

On Ellett's recommendation, I received a postdoctoral appointment as a Carnegie Research Fellow in Merle A. Tuve's nuclear physics laboratory at the Department of Terrestrial Magnetism in Washington D.C. For about a year and a half, 1939–1941, I continued research in nuclear physics using the DTM's pioneering 1 MeV van de Graff accelerator [*Van Allen and Ramsey*, 1940; *Van Allen and Smith*, 1941a,b]. Also my interests in cosmic rays, atmospheric physics, and geomagnetism were revived by association with the more traditional members of the DTM staff, especially Scott E. Forbush, E. H. Vestine, and John A. Fleming, then director of DTM, and by the extended visits of Sydney Chapman and Julius Bartels who were then completing their great two volume work *Geomagnetism* [*Chapman and Bartels*, 1940].

But all such matters were placed on the back burner by DTM's conversion to "war-work," specifically on the development of radio-proximity fuzes for naval antiaircraft projectiles. I joined full heartedly in this work and continued it when the project was transferred in March 1942 to the newly created Applied Physics Laboratory (APL) of Johns Hopkins University, located in nearby Silver Spring, Maryland. My principal contribution was to the development of vacuum tubes capable of surviving the some 20,000 g acceleration which they experience during firing from a naval gun [*Porter et al.*, 1963]. My consequent experience in building rugged electronics and in internal and external ballistics gave me the confidence to undertake scientific work using high-performance rockets and spacecraft in later years.

On 6 November 1942, two of my colleagues at APL and I were given spot commissions as lieutenants (junior grade) in the U.S. Naval Reserve to take the first load of proximity fuzed projectiles to the Pacific Fleet, issue them to combatant ships and instruct gunnery officers in their virtues (and shortcomings). Our military training consisted of reading a pamphlet on the duties and responsibilities of officers of the U.S. Navy. Thirteen days later we sailed on a transport ship from San Francisco nonstop to Nouméa, New Caledonia, headquarters of the South Pacific Fleet. During the subsequent three and a half years, I served as a gunnery and ordnance specialist on destroyers, at advanced bases, and on the battleship U.S.S. *Washington* for two eight-month tours of duty as APL's man in the South Pacific.

In March 1946, now a lieutenant commander, I was released from active military duty and returned as a civilian to the Applied Physics Laboratory to develop a program of using high-altitude rockets for scientific work. Our group conducted such work with captured and refurbished German V-2 rockets and oversaw the development and corresponding use of the American Aerobee rocket. Our principal achievements were observation of cosmic rays above the appreciable atmosphere and the inference of the primary cosmic ray spectrum using the Earth's magnetic field as a huge magnetic spectrometer in accordance with the theory of Carl Störmer [*Störmer*, 1955]; determination of the interactions of primary cosmic rays with the atmosphere and with localized material; obtaining high-resolution spectograms of the Sun's ultraviolet down to 2285 Å; measurement of the distribution of ozone in the upper atmosphere; and near infrared photography of the Earth's surface and cloud cover from altitudes up to 160 km. Most of these investigations in geophysics had their roots in my early work with Poulter and my incidental associations at DTM. On the technical side, I felt qualified to oversee the development of equipment for rocket flights based on my wartime experience.

In 1950, I was offered a position as professor of physics and head of the department of physics at my Ph.D. alma mater, the University of Iowa, and I was delighted to return there in January 1951.

2. WORK WITH BALLOON-LAUNCHED ROCKETS

At Iowa, I was joined by Melvin Gottlieb, whose Ph.D. was from the University of Chicago, in the conduct of cosmic ray investigations by balloon techniques and soon thereafter by Frank B. McDonald and Kinsey A. Anderson, both University of Minnesota Ph.D.'s; and by a succession of very able graduate students.

My own principal aspiration was to extend our earlier (APL) observations of primary cosmic rays above the atmosphere to polar regions. For this purpose, we adopted the inexpensive technique of launching small, military-surplus rockets from balloons at altitudes of the order of 20 km in order to carry small scientific instruments to altitudes of about 100 km. The U.S. Office of Naval Research approved my proposal for such a program and supplied the modest but essential financial and operational support for such an undertaking. My colleagues and I conducted a large number of successful (and unsuccessful) flights from shipboard in the Arctic and sub-Arctic in the summers of 1952, 1953, 1954, and 1955. We then obtained additional support from the National Science Foundation as part of the International Geophysical Year (IGY) (1957–58) for Arctic, equatorial and Antarctic expeditions in 1957.

In its relevance to magnetospheric physics, the most significant result of our rockoon (balloon-launched rocket) work was our observations of the "auroral soft radiation" as we called it. During our 1953 Arctic expedition, the intensity of radiation above about 50 km (as measured by single, lightly shielded Geiger tubes) was found to be much greater at geomagnetic latitudes $\lambda = 64°$ N and 74° N than at either lower (56° N) or higher (89° N) latitudes. It appeared that the effect was confined to the auroral zone. Our initial working hypothesis was that we were detecting electrons in the high-energy tail of the primary auroral spectrum [*Meredith et al.*, 1955]. If correct, the absolute intensity of electrons $E_e \geq 2$ MeV was of the order of tens $(\mathrm{cm}^2\ \mathrm{s})^{-1}$. Follow-on investigations in 1954 and 1955 used combinations of Geiger tubes with and without additional shields of lead and aluminum and very thinly shielded NaI (Tl) scintillation crystals, cemented on photomultplier tubes and surrounded by current-carrying coils for deflecting electrons. The auroral zone identification of the effect was confirmed but our initial hypothesis as to the nature of the causative radiation was found to be false. It was conclusively supplanted by the finding that the originally observed effect was caused by bremsstrahlung (X rays) of the order of 10 kev energy generated in the atmosphere above about 90 km and in the nose cone of the rockets by the absorption of primary auroral electrons. In a summary paper *Van Allen* [1957] estimated that the intensities of electrons in the primary auroral beam were in the range 10^6 to 10^8 $(\mathrm{cm}^2\ \mathrm{s})^{-1}$ for electrons in the energy range 10–100 kev and that the corresponding energy flux was 0.01 to 1.0 erg $(\mathrm{cm}^2\ \mathrm{s})^{-1}$.

Further confirmatory observations were made during our 1957 Arctic, equatorial, and Antarctic expeditions [*Van Allen,* 1995].

3. PLANNING FOR OBSERVATIONS WITH SATELLITE-BORNE INSTRUMENTS

In early 1946, a group of individuals was assembled to plan and coordinate the use of V-2 rockets for scientific purposes. This group, initially called the V-2 Upper Atmosphere Panel, had no official governmental status and, in principle, was only advisory to the U.S. Army Ordnance Department. But in practice, it came to be the central entity for assigning payload space, exchanging plans and experience and timely reporting of results and interpretations of our findings. E. H. Krause was the first chairman of the panel. I was elected to succeed him in 1947 and continued as chairman until the creation of NASA in late 1958. The original name of the panel was changed to the Upper Atmosphere Rocket Research Panel (UARRP) after we began using the Aerobee and other American sounding rockets and, later, to the Rocket and Satellite Research Panel (RSRP). Membership in the Panel varied over its history but typically comprised about 12 representatives of military or quasi-military laboratories engaged in high-altitude research.

During the early 1950's, rapid advances in the development of inter-continental ballistic missiles in both the United States and the USSR gave realism to the long hoped-for possibility of delivering payloads of scientific instruments into durable orbits about the Earth and beyond. Members of the UARRP led the U.S. effort to define prospective uses of artificial satellites as a natural extension of their work with high-altitude rockets. In parallel with these activities, formal plans for including satellites in the program for the International Geophysical Year (1957–58) were being made by both the U.S. and the Soviet Union. Other members of the UARRP and I served on the rocket and satellite panels of the U.S. National Committee for the IGY and testified on behalf of the IGY program before numerous congressional committees.

Anticipating the desirability of specific suggestions, I wrote a detailed "Proposal for Cosmic Ray Observations in Earth Satellites" and submitted it on 25 September 1955 to Joseph Kaplan, chairman of the U.S. National Committee.

The 43rd meeting of the UARRP in Ann Arbor, Michigan was an invitational but unclassified symposium on prospective uses of satellites for scientific purposes and on related technical considerations. I presented two papers and edited the thirty-three papers of the symposium into book form [*Van Allen,* 1956, 1958].

My two papers were the basis for follow-on proposals for the IGY program. The first was entitled "Cosmic-Ray Observations in Earth Satellites, Part A. Geographical Dependence and Temporal Variations of Cosmic-Ray Intensity in the Vicinity of the Earth. Part B. Relative Abundance of Heavy Nuclei in the Primary Cosmic Radiation." The second was entitled "Study of the Arrival of Auroral Radiations."

In the spring of 1956 graduate student George Ludwig and I began work on specific detector systems and supporting circuitry for an instrument for flight on a satellite. By this time we were beginning to understand the probable mass, power, size, and telemetry restraints of early U.S. satellites. Also, it appeared likely that the propulsive capability of any practical combination of U.S. rockets would restrict early launchings to a nearly due east direction from Cape Canaveral so that orbits would be limited to the approximate latitude range of 33° N to 33° S. For this reason we temporarily abandoned plans for auroral studies with satellite equipment. In May 1956, we received formal IGY approval of our proposed instrument and funding from the National Science Foundation (originally \$106,375, later increased to \$109,225).

We were acutely aware of the necessity for building satellite instruments to rigorous standards of reliability, i.e., insensitivity to temperature over a wide range, immunity to corona discharge in partial vacuum, generous tolerances on all operating elements, and mechanical ruggedness to resist the vibrations and linear and angular accelerations of launch. Other specific requirements for satellite-borne instruments were minimizing the electrical power and the mass and volume of the equipment. All of these considerations led us in the direction of simplicity and meticulous attention to each element of the instrumentation. I settled on a single Geiger tube as the basic radiation detector and obtained and tested samples of the halogen-quenched tubes developed by Nicholas Anton in his laboratory in New York. These tubes were mechanically rugged, had "infinite" lifetime, gave large signals, and operated stably over a wide range of temperature—characteristics critically important for reliable operation on a satellite. Ludwig developed all of the circuits, including high-voltage and low-voltage power supplies, pulse amplifiers, scaling circuits, mixers, and modulators, all using transistors. We also realized that real-time telemetry reception would be rather meager early in the program, whereas the purposes of our investigation demanded the fullest possible latitude, longitude, altitude and temporal coverage and hence the addition of some form of data storage and playback, plus a command receiver. For this purpose Ludwig and our lead instrument maker Edmund Freund designed and built a miniature magnetic tape recorder. But we also included real-time data transmission as a back-up and supplemental mode.

Meanwhile there was an intense competition between the Army Ballistic Missile Agency (ABMA) and the Office of Naval Research/Naval Research Laboratory (NRL) for supplying the launching vehicle. The NRL proposed a new three-stage vehicle called the Vanguard. But on 20 September 1956, the ABMA demonstrated its existing capability of placing a small satellite in orbit using a four-stage combination of military rockets, called Jupiter C. However, for political reasons they were forbidden to carry a live fourth stage. This cold-war sensitivity to using military rockets for peaceful purposes in the IGY was also influential in the governmental decision to adopt the not-yet-developed Vanguard. The Iowa instrument was one of several slated for an early flight on a Vanguard but I decided that it would be prudent to design it so that it would fit in the payload of either a Vanguard or a Jupiter C.

4. DISCOVERY OF THE NATURAL RADIATION BELTS OF THE EARTH, EXPLORERS I AND III

By September 1957 Ludwig had completed construction and testing of our full package, nominally for the Vanguard but readily adaptable to the Jupiter C. The only unsolved aspect of the latter adaptation was the question of proper operability of the tape recorder in the rapidly spinning fourth stage of the Jupiter C.

During July–November 1957 I was principally engaged in our Arctic and equatorial-Antarctic rockoon expeditions, the latter aboard the large icebreaker the U.S.S. *Glacier*, as another part of our IGY program [*Van Allen,* 1995], sponsored by the Office of Naval Research and the National Science Foundation. Development of the Vanguard was proceeding but it did not appear likely that we would be able to fly our instrument during 1957.

As a result of the national trauma that followed the Soviet's successful launching of Sputnik I on 4 October 1957, Wernher von Braun, William Pickering and their associates at ABMA and the Jet Propulsion Laboratory (JPL) had been extremely active in developing a U.S. "response," as I learned later. In the course of these discussions Pickering pointed out that the Iowa cosmic-ray instrument was the only prime IGY instrument that had been configured for the Jupiter C payload, as an alternative to Vanguard. Von Braun, who had endorsed and fostered my decision, replied with mock innocence, "Isn't that interesting?" Following arrangements by Eberhardt Rechtin of JPL, Henry Richter and two others from JPL visited the University of Iowa on 23 October and reviewed the details of our instrument with Ludwig. Rechtin met with Secretary of the Army Wilber M. Brucker and with Mr. Holaday (DOD missile coordinator) on 28 October and received their approval of Jupiter C as a back-up satellite launcher. On or about 30 October Richard W. Porter, chairman of the IGY's Technical Panel for the Earth Satellite Program (TPESP) recommended the Iowa cosmic ray detectors as the first priority payload for Jupiter C and Pickering contacted me by radiogram on the *Glacier*. After an exchange of clarifying messages during the following two weeks, I agreed.

Ludwig packed all of his technical equipment and payload components (and his family), canceled his university registration, and moved between 18 and 20 November to Pasadena "for the duration." There he worked with the JPL staff in the detailed adaptation of our instrument to the Jupiter C payload. A wire-grid for the detection of micrometeoroids was supplied by the Air Force Cambridge Research Laboratory; all other elements of the payload were developed and supplied by JPL/ABMA. The entire arrangement was called "Deal I." Because of the high rate of spin of the Jupiter C final stage, as compared with that of Vanguard, Ludwig omitted his magnetic tape data storage unit from Deal I; as a result we were to be exclusively dependent on real-time telemetry on the first flight. But he immediately undertook modifications of the tape recorder to reduce its susceptibility to centrifugal force and arranged for its on-axis positioning in the payload. Our plan for the tape recorder was that we could store the data from our detector over a complete orbit, then play it back on command within six seconds over a chosen receiving station. This configuration was called "Deal II." As of early December 1957 the planned launch readiness dates were 1 February 1958 for Deal I and 1 March 1958 for Deal II. The date for Deal I was compatible with von Braun's 8 November promise to President Dwight Eisenhower that the first Jupiter C with scientific payload would be ready to launch within ninety days. As mentioned above the reasonable nature of this promise had been demonstrated in September 1956 by the flight of Jupiter C (with inert fourth stage) to a range of 5,300 km. The completed payload for the first Jupiter C launch was delivered to Cape Canaveral in mid-January 1958. Ludwig accompanied the JPL group to the Cape for preflight checkout work and for the launch.

The Deal I (*Explorer I*) payload was lifted off the launch pad at Cape Canaveral 10:48 P.M. EST on 31 January 1958 (03:48 GMT of 1 February). The Jupiter C vehicle consisted of four propulsive stages. The first stage was an upgraded ABMA Redstone liquid-fueled rocket. The second stage consisted of a cluster of eleven Sergeant motors of JPL development; the third, a cluster of three such motors; and the fourth and final stage, a single such motor to which the payload (instrumentation 4.82 kg, total orbital body 13.97 kg) was permanently attached. The burning of all four stages was monitored by down-range stations and judged to be nominal. The final burnout velocity of the fourth stage was somewhat higher than intended, and there was a significant uncertainty in the final direction of motion. Hence, the achievement of an orbit could not be established with confidence from the available data. The telemetry transmitter was operating properly, and the counting rate data from our radiation instrument during the first five minutes of flight corresponded to expectations, thus establishing its survival of the launching sequence. It was not until almost two hours later that the reception of telemetry signals by stations in Southern California proved that a durable orbit had been achieved.

Announcement of this fact was made at a large press conference in the early hours of 1 February in the Great Hall of the National Academy of Sciences. The successful launch of *Explorer I* was an event of major national and international interest, coming as it did after a launch failure of Vanguard in December 1957.

It was not until many weeks later that we were able to assemble a sufficiently comprehensive body of data to achieve a clear understanding of the scientific results from *Explorer I*. Meanwhile, we all went back to work preparing for the scheduled launches of two Deal II payloads. Ludwig returned to Pasadena, and I returned to Iowa City, where Ernest Ray and I started organizing the initially rather meager flow of data from our cosmic ray instrument. Also, I resumed my normal university duties. In a brief memorandum to the IGY staff and committees on 28 February we made a preliminary report of generally favorable operations.

Carl McIlwain, an advanced graduate student, was principally engaged at this time in preparing auroral particle detection instruments for rocket flight from Fort Churchill, but he also joined in the examination of *Explorer I* data. During February McIlwain conducted two Nike-Cajun rocket flights into visible aurorae and established that the "major fraction of the auroral light was produced by electrons with energies of less than 10 keV," thus confirming and greatly improving the inferences that we had made from our soft auroral radiation observations during the preceding several years [*McIlwain*, 1960].

Explorer II with our full-up payload (Deal II) including the magnetic tape recorder was launched on 5 March, but an orbit was not achieved because the fourth-stage ignition failed.

On 11 and 12 March Ludwig and I met in Pasadena with Pickering, Jack Froelich, and Henry Richter of JPL and Wolfgang Panofsky of Stanford to review the radiation intensity data from *Explorer I*. We also discussed techniques for building miniature detectors that were suitable for small satellites but capable of particle identification and the measurement of energy spectra and angular distributions—characteristics not possessed by the single Geiger tube that I had adopted for the early measurements. There were vague allusions at this meeting to the possibility of radiation experiments at high altitudes by the Atomic Energy Commission (AEC). It was some four weeks later before I learned that a series of high-altitude bursts of small nuclear bombs was being planned as a test of an idea that Nicholas Christofilos of the Livermore Radiation Laboratory had proposed. He had visualized that energetic, electrically charged particles could be artificially injected into the earth's magnetic field and that they would then be trapped therein as in a laboratory magnetic mirror machine. In fact, he was engaged in developing a dipolar magnetic field machine, called the Astron, for demonstrating confinement of hot plasma in the laboratory. There were also military implications of the possible production of high-intensity radiation regions around the earth and of the transient ionospheric effects of the bursts. Panofsky had been asked by the AEC to help assess the physical effects to be expected. Pickering had suggested that the Iowa group would be a suitable one for observing the effects. The proposed tests were classified as secret at that time. Pickering and Panofsky had developed the plan that an unclassified program of satellite observations with an improved version of our *Explorer I* instrument might be placed within an IGY context as a logical follow-on to *Explorer I* but that such an instrument would also serve the classified purpose.

I was very eager to join in this enterprise purely as an extension of our observations with *Explorer I*, but I was also intrigued by my then vague perception of the other possibilities. Ludwig and I started to plan an improved system of detectors. We were soon joined by McIlwain, who contributed his valuable experience in developing instrumentation for auroral rocket flights.

Explorer III, carrying our full Deal II instrument, was successfully orbited by a Jupiter C on 26 March 1958. The tape recorder, the first such device ever flown in a satellite, functioned beautifully in response to ground command and fulfilled our plan of providing complete orbital coverage of radiation intensity data.

The assembly of data from *Explorer I* was proceeding rather slowly. The long slender body equipped with four whip antennas had passed from its original axial spin mode into the minimum kinetic energy state of a flat spin about a transverse axis as deduced from the modulation of the received signal—an impressive and humiliating lesson in the elementary mechanics of a somewhat nonrigid body. This modulation produced periodic and substantial fade-outs of the signal. In addition, the r.f. signals (at 108 MHz) from the 60 milliwatt and the 10 milliwatt transmitters were quite weak, at the best, and there were many other technical problems in getting reliable, noise-free recordings at the receiving stations. The entire telemetry system was, of course, undergoing its first shakedown under fully realistic conditions. The overall result was that during the first few weeks of *Explorer I*'s orbital flight, we had only a sparse set of data consisting of segments of the order of one minute's duration from many different and somewhat uncertain positions in latitude, longitude, and altitude. The Naval Research Laboratory tracking team under Joseph Siry's supervision was making a heroic effort to calculate a reliable orbit but was also having shakedown problems, partly because of the larger-than-expected eccentricity of the orbit.

During some segments of the data the counting rate of our single Geiger tube was of the order of 12 to 80 counts per second, generally within the range that we had expected for cosmic radiation, and the instrument appeared to be operating reliably. In other segments of data there were no counts observed for as long as two minutes. On one noteworthy pass the apparent rate underwent a transition from zero to a reasonable value within about twenty seconds. There was no conceivable way in which the cosmic-ray intensity could drop to zero at high altitudes. On the other hand we had a high level of confidence in the Geiger tube and the associated electronic circuitry because of its conservative design and the rigorous thermal and mechanical testing to which it had been subjected before flight. The puzzle hung over our heads as we tried to see if there was any systematic dependence of the apparent failure on passing through earth shadow, on payload temperature, on altitude, latitude or longitude. Noise-free data accumulated slowly. Also, the higher power transmitter failed after eleven days of flight. This premature failure spawned the hypothesis that a vital component had been disabled by a micrometeoroid hit, a possibility of neurotic concern at that time. In defiance of the hypothesis the transmitter resumed operation a few days later, but its operation thereafter was desultory. Throughout this early period we were heavily occupied, as indicated above, in preparing for the *Explorer II* and *III* launchings, in working out data reduction and analysis techniques (to which we had given relatively little attention before flight), in formulating plans for subsequent flights, and in coping with a steady flow of urgent phone calls on practical arrangements and on inquiries on our progress. Indeed, our original plan to accumulate a comprehensive body of data on the distribution of cosmic-ray intensity around the earth did not dictate urgency in analysis of data. We had high hopes for eventually getting a worldwide survey of the auroral soft radiation, but the decision to launch *Explorer I* nearly due east from Cape Canaveral in order to obtain the maximum advantage of the rotation of the earth yielded an orbital inclination of only 33.3° to the equator. Hence, the maximum geomagnetic latitude was only about 45°, far below the auroral zone.

Following the 26 March launching of *Explorer III*, I flew to Washington to confer with Siry, John Mengel, and others at the Naval Research Laboratory and to pick up preliminary data from *Explorer III*. Contrary to some popular accounts the Vanguard group fully supported the Explorer program in many vital ways. The first successful launch of Vanguard had occurred on 17 March, and the NRL team was operating on an around-the-clock basis, coping now with tracking and data acquisition for three satellites. From NRL I returned to the Vanguard data reduction center on Pennsylvania Avenue and picked up a complete orbital record of a successful playback from our *Explorer III* tape recorder. The playback had been received at the San Diego minitrack station on 28 March. I put the record in my briefcase and returned to my hotel room, where, with the aid of graph paper, a ruler, and my slide rule, I worked out the counting rate of our Geiger tube as a function of time for a full 102-minute period and plotted the data.

The data provided a beautiful explication of the still fragmentary information from *Explorer I*. The counting rate at low altitudes was in the expected range of 15 to 20 counts per second. There was then a very rapid increase to a rate exceeding 128 counts per second (the maximum recordable rate of our on-board storage system). A few minutes later, the rate decreased rapidly to zero. Then after about fifteen minutes, it rose rapidly from zero to greater than 128 counts per second and remained high for forty-five minutes, then again decreased rapidly to 18 counts per second as the orbit around the earth was nearly completed. At 3:00 A.M. I packed my work sheets and graph and turned in for the night with the conviction that our instruments on both *Explorers I* and *III* were working properly, but that we were encountering a mysterious physical effect of a real nature. Early the following day, I flew back to Iowa City and proudly displayed my graph to Ernest Ray, Carl McIlwain and Joseph E. Kasper. During the previous day McIlwain had made tests with our prototype Geiger tube and circuit using a small x-ray machine and demonstrated that a true rate exceeding about 25,000 counts per second would indeed result in an apparent rate of zero. The conclusion was then immediate—at higher altitudes the intensity was actually at least a thousand times as great as the intensity due to cosmic radiation. Ray's famous (though consciously inaccurate) remark summarized the situation, "My God, space is radioactive!" Our realization that there was actually a very high intensity of radiation at high altitudes rationalized our entire body of data.

George Ludwig returned from JPL to the University of Iowa on 11 April, and the four of us worked feverishly in analyzing the data from *Explorers I* and *III* (by primitive hand reduction of pen-and-ink chart recordings) and organizing them on an altitude, latitude, and longitude basis. A crucial aspect of the data was the repetitive, systematic dependence of the Geiger tube's counting rate on these parameters. I promptly informed Porter, Odishaw, Newell, and Pickering of our results. The latter informed Panofsky, whose reaction was that the Soviets had beaten us to the punch in conducting a Christofilos-type test. Odishaw admonished me to make no public announcement of our findings, pending a formal IGY report, which he would schedule as soon as we had our results in order. We agreed on a 1 May date for the public report.

During mid-April we prepared graphs and a short written statement of our raw findings [*Van Allen et al.*, 1958], and I mulled over the meaning of the results. I entertained two quite different lines of thought: *(a)* that we might be detecting high-energy x rays or γ rays, possibly from the sun, or *(b)* that the high intensity radiation might be akin to the auroral soft radiation that we had studied during the preceding several years with rockoon flights at high latitudes and most recently with McIlwain's rocket flights from Fort Churchill and that we had identified as being electrons having energies of the order of 10s of keV [*McIlwain*, 1960]. I quickly rejected hypothesis *(a)* on the conclusive grounds that "the effect" was present during both daylight and dark conditions, that it exhibited a strong latitude effect, and that the extremely sharp increase in intensity with increasing altitude was impossible for any plausible type of electromagnetic or corpuscular radiation arriving directly from a remote source. Specifically, the rapid increase occurred within an altitude range of less than 100 kilometers at an altitude of the order of 1000 km; the decrease in atmospheric thickness within that increment of altitude was totally negligible compared to the some 1.5 g cm^{-2} of material in the nose cone and wall of the counter. I concluded that the effect had to be attributed to electrically charged particles, mechanically constrained by the earth's external magnetic field from reaching lower altitudes. By virtue of my familiarity with an early paper of *Störmer* [1907] and with magnetic field confinement of charged particles in the laboratory during my 1953–54 work building and operating an early version of a stellarator at Princeton, I further concluded that the causative particles were present in trapped orbits in the geomagnetic field, moving in spiral paths back and forth between the northern and southern hemispheres and drifting slowly around the earth. The intensity of such trapped particles would be diminished at low altitudes by the cumulative effect of atmospheric absorption and scattering.

The foregoing account of observations and interpretation is essentially the one I gave in a joint session of the American Physical Society and the National Academy of Sciences in the latter's auditorium on 1 May 1958. Fortunately, a tape recording was made of my lecture and of the ensuing question-and-answer period, though I did not learn that until a year or more later. A written transcription of this tape provides a documented, published record of this lecture, complete with grammatical errors and colloquial language [*Van Allen*, 1961].

I had adopted at that time the working hypothesis that the trapped radiation consisted of "electrons and likely protons, energies of the order of 100 keV and down, mean energies probably about 30 keV." In this vein of thought the response of our Geiger tube would be attributed to bremsstrahlung produced as the electrons bombarded the nose cone of the instrument. If this bremsstrahlung interpretation were correct, I estimated an omnidirectional intensity of 10^8 to 10^9 (cm^2 sec)$^{-1}$ of 40 keV electrons would be required to account for the observed counting rate at altitudes of $\sim$ 1500 km over the equator. However, in my 1 May lecture as well as in response to a question at the end, I emphasized that we had no definitive identification of particle species and that the particles might be penetrating protons or penetrating electrons. I did, however, regard protons and electrons of energies necessary to penetrate the Geiger tube directly, namely $E_p > 35$ MeV and $E_e > 3$ MeV, as unlikely in view of our auroral zone measurements with rocket-borne equipment. A few months later, we showed that this opinion was mistaken.

Of the some 1,500 real-time recordings [*Ludwig*, 1959] of *Explorer I* telemetry signals in the period 1 February to 9 May 1958, 850 contained readable cosmic-ray data. The tape-recorded data from *Explorer III* had an upper rate limit of 128 counts per second, a value that we had judged to be adequate for the originally planned cosmic-ray investigation. However, the real time transmissions from both *Explorers I* and *III* had a much greater dynamic range.

Data reduction continued during 1958 and 1959. Virtually all of the *Explorer I* data were reduced and analyzed [*Yoshida et al.*, 1960; *Loftus*, 1969].

Following our 1 May report and news conference we received many inquiries for further details. One that lingers in my memory is the following telephone conversation.

"This is John Lear, science editor of the *Saturday Review of Literature*, calling from New York." Heavy emphasis on "calling from New York," then a long pause waiting for me to recover from the thrill of hearing from such an important person, in New York, no less. Actually I did know who he was and had often characterized him as the anti-science editor of the *Saturday Review*. He continued: "I read of your recent report of the discovery of radiation belts of the Earth and thought that I would do a piece on this subject. What I found remarkable was that such important work had been done at a midwestern state university." Heavy emphasis on "midwestern state university." Well, I don't think that I responded with any profanity but I did manage to convey a suggestion as to what he could do with his piece and hung up. The next day, the president of my university, Virgil M. Hancher, called me to report that Mr. Lear had called him to complain about my discourtesy. I then gave a brief explanation of my reaction, at the end of which Hancher replied "I promised Lear that I would call you and you may now consider that I have done so. And, by the way, Van, my congratulations!" I never heard from the matter again. It's great to have a boss like that.

In our early reports, I used the term "geomagnetically trapped corpuscular radiation." At the press conference following the 1 May 1958 lectures at the National Academy of Sciences, I described the distribution of the radiation as encircling the earth. A reporter asked "Do you mean like a belt?" I replied: "Yes, like a belt." This was the origin of the term radiation belt. At a meeting sponsored by the International Atomic Energy Agency in Europe in the summer of 1958, Robert Jastrow first used the term *Van Allen radiation belt*.

The more inclusive term *magnetosphere* was suggested by Thomas Gold [1959]: "It has now become possible to investigate the region above the ionosphere in which the magnetic field of the earth has a dominant control over the motions of gas and fast charged particles. This region is known to extend out to a distance of the order of 10 earth radii; it may appropriately be called the magnetosphere." This term is now used almost universally in referring to a large body of geophysical phenomena as well as to corresponding phenomena at other planets and other celestial objects (e.g., pulsars).

5. ARTIFICIAL RADIATION BELTS, EXPLORER IV

As I have noted in the preceding chapter, another development of major importance to us was moving along in parallel with our work on the data from *Explorers I* and *III*. In December 1957 Nicholas Christofilos, a Greek engineer-scientist at the Livermore Radiation Laboratory, had proposed exploding one or more small nuclear fission bombs at a high altitude ($\sim$ 200 km) to test two effects that he envisioned: *(a)* the prompt enhancement of ionospheric ionization and the consequent disruption of radio communications at VHF frequencies and *(b)* the injection of large numbers of energetic electrons ($E_e \sim 2$ MeV) into durably trapped orbits in the earth's magnetic field. The electrons (and positrons) from the decay of radioactive fission products would be the principal source of both effects; an additional source of ionospheric influence would be the prompt ultraviolet, x- and γ-radiations from the burst itself.

In mid-April 1958 I informed Pickering and Panofsky of my by-then reasonably firm interpretation of the observations by *Explorers I* and *III*—namely that there was a huge population of electrically charged particles already present in trapped, Störmerian orbits in the earth's external magnetic field. In the context of our earlier studies of the primary auroral radiation, I considered it likely that these particles had a natural origin. At this time, I was introduced to the secret plans for the conduct of the high-altitude bomb tests, later called Argus. I also learned that, despite the absence of definitive information, some officials of the Atomic Energy Commission believed that the Soviet Union might have already conducted such tests. Panofsky suggested that, if this were in fact true, the radiation discovered by *Explorers I* and *III* might be of such artificial origin. A complementary line of thought was that, in either case, the Argus tests would provide the United States with the necessary competence to detect high-altitude nuclear bomb tests by the Soviet Union or by other countries.

Using unclassified information on the fission yield of nuclear (atomic) bombs and electron spectra of the radioactive decay products and estimates of injection efficiency and the geometry of geomagnetic trapping, I estimated the resulting intensities and spatial distribution of trapped electrons.

My principal interest was to have an opportunity to follow up on our earlier work by flying detectors of the proper dynamic range and of more discriminatory capability than that of the Geiger tubes on the original *Explorers*. At the same time I was eager to participate in the Argus tests because of their apparent national importance and more particularly because of the possibility of distinguishing between a natural and an artificial population of geomagnetically trapped particles and of making direct observations of the residence times and diffusion rates of a known spectrum of electrons, injected at a known place at a known time. Ludwig, McIlwain, Ray, and I set to work designing, building, and testing a system of detectors and associated electronics for further satellite missions. Our principal working relationship in developing a practical payload for a Jupiter C vehicle was with Ernst Stuhlinger, Joseph Boehn, and Charles Lundquist of the Army Ballistic Missile Agency in Huntsville, Alabama, and with JPL.

Porter and Kaplan were instrumental in arranging for IGY sponsorship and U.S. Department of the Army support for our work as an extended part of the IGY satellite program. Such sponsorship put our work on an unclassified level, as was altogether proper, but it also shrouded the classified aspect. In short, IGY sponsorship was the truth, but not the whole truth. On 1 May 1958, we received an informal go-ahead with assurance of financial support. The proposed tests were being planned for August and September, a considerable challenge to say the least.

JPL had upgraded the performance of the high-speed stages of the Jupiter C so that a significant increase in the mass of the payload was possible, but we continued to labor under very severe constraints of mass, power, and telemetry capacity. Our optimized design provided battery power for about two months of flight. Two satellite flights were planned, as were some nineteen rocket flights by the Kirtland Air Force Base and a variety of ground- and ship-based observing programs. Three rocket flights of nuclear bombs were planned in the Argus program. From a geomagnetic point of view the best site for the injection of electrons into durable orbits was over the South Atlantic (the South Atlantic anomaly). Because of the eccentricity of the earth's magnetic field, a site at that longitude would minimize the altitude of injection in order to produce durably trapped orbits. Launching from a ship in an isolated site was desirable because it allowed the secrecy of the operation to be safeguarded. Two satellite launches and three bomb injections were judged to be the minimum effort to give reasonable assurance of success. The Navy's guided missile ship, the U.S.S. *Norton Sound*, which we had "initiated" with Aerobee rocket launchings in 1949, was selected to launch the rockets. Needless to say, intensive coordination of the efforts of hundreds of people was required to assure operational success. Herbert York, director of the Advanced Research Project Agency of the DOD, was the central person in this coordination. In addition to the three Argus bursts of small ($\sim$ 1 kiloton) bombs, the AEC planned tests of two large ($\sim$ 10 megaton) bombs at altitudes of 40–70 km over Johnston Atoll in the central Pacific.

Visitors to the University of Iowa during the spring and summer of 1958 were astonished to find that a crucial part of this massive undertaking had been entrusted to two graduate students and two part-time professors, working in a small, crowded basement laboratory of the 1909 Physics Building. But we knew our business and were in no way intimidated by representatives of huge federal agencies. We settled on an array of four basic radiation detectors: a thin plastic scintillator on the face of an end window photomultiplier, a thin slice of cesium iodide crystal on the face of another photomultiplier, and two miniature Geiger tubes, one lightly shielded and the other heavily shielded. Every effort was made to provide the necessary dynamic range to cope with the intensity of the natural radiation and with the estimated intensity of the additional radiation that was to be artificially injected. Upgraded high-speed stages of the Jupiter C made it possible to plan an increase in inclination of the satellite orbit from the 33° of the orbits of *Explorers I* and *III* to 51° in order to provide improved coverage in latitude. On 1 July 1958, Ludwig, McIlwain, and our new electronics engineer, Donald Enemark, took our completed No. 1 and No. 2 payloads to Huntsville for final environmental testing there and then to Patrick Air Force Base in Cocoa Beach, Florida, for spin balancing.

Explorer IV was launched successfully up the east coast of the United States on 26 July 1958. It carried Iowa payload No. 2, the best of four units that we had built. In my journal I wrote on 29 July: Apparatus seems to be working quite well! Confirms existence of 'soft radiation'—except it doesn't seem to be so soft (?). And on 30 July:
President Eisenhower yesterday signed the National Aeronautics and Space Act.

McIlwain brought our payload No. 3 (with a bad PM tube) back to Iowa City while Ludwig remained at Cape Canaveral checking out No. 4. The No. 1 payload had been fully checked out and was also available for flight on *Explorer V*. The entire system of orbit determination and data acquisition (all real time, no storage) had been improved markedly during the preceding six months,

and by 31 July we had examined the data from several recorded passes by various stations. My journal summary was as follows:
General situation on data: Explorer IV (satellite 1958ϵ)

- (1) All detectors working splendidly.
- (2) 1.4 g cm^{-2} of Pb around the shielded counter reduces the counting rate only *mildly* at the higher altitudes, i.e., $\sim$ factor of 2!
- (3) Detector A (pulse scintillation channel) giving fluxes $\sim 10^4$ (cm^2 sec sr)$^{-1}$ at $\sim$ 1500 km altitude, latitude $\sim 12^\circ$N (E > 600 keV if electrons).
- (4) Detector B (CsI detector) gives $\sim$ 1 erg (cm^2 sec sr)$^{-1}$.
- (5) Rapid altitude dependence between 600 $\rightarrow$ 1500 km observed as in Explorers I and III
- (6) GM tube rates, shielded 184 sec^{-1}, unshielded 308 sec^{-1} at 1500 km, 12°S.

Based on data so far received and read ($\sim$ 6 to 8 passes here and there):

- I. Previous results [from satellites 1958 α and γ] well confirmed.
- II. Evidence for a high energy tail or perhaps another phenomenon than aurora. However, Kinsey Anderson has previously had evidence for high energy $\sim$ 500 keV electrons. Due for publication in Phys. Rev. on 15 August.

On August 4 my wife and I packed our then four children into our station wagon and drove to Long Island, New York, for a long planned, but brief, vacation. In my journal for 9 August:

I recorded a list of phoned suggestions to McIlwain for changes in the Explorer V detectors in light of Explorer IV data. Among other reasons I cited the following (verbatim transcript) from my journal:
(c) The high intensity radiation may very well be dominantly *protons*. Injection mechanism may be—decay of albedo (upward moving from top of atmosphere) neutrons at high altitudes into *protons* and *electrons*. Since "lifetime" of trapped particles in the earth's magnetic field is probably limited dominantly by *scattering out of* trapped orbits which are mirrored at high altitudes to ones less inclined to the magnetic lines—which are therefore mirrored at lower altitudes and hence result in more rapid loss of energy by ionization—protons probably a *very* much greater lifetime than electrons of comparable energy. The factor may well be of the order of mass ratio—at least 2000 to one.

The modified payload No. 4 was launched as *Explorer V* on 24 August 1958, but the final stage of the Jupiter C failed to ignite and an orbit was not achieved. Our apparatus functioned properly during its brief ten-minute flight before falling into the sea.

On the basis of the first few weeks of data from *Explorer IV*, we had advised ARPA of our discovery of a minimum in the previously present radiation when intensity at constant altitude was plotted against latitude. This finding was utilized in helping select the latitude for the Argus bursts so that the planned artificial radiation belts would enjoy the optimum prospect for detection. This choice of latitude turned out to be the best possible one within the latitude range of *Explorer IV*, i.e., in the "slot" between the previously observed "inner" radiation belt and the newly discovered "outer" radiation belt.

The AEC/DOD task group successfully produced two bursts of 10-megaton yield bombs, called Teak and Orange, on 1 August and 12 August at approximate altitudes of 75 and 45 km, respectively, above Johnston Atoll in the Central Pacific. The three Argus bursts (about 1.4 kiloton yield) were produced successfully on 27 August, 30 August, and 6 September at altitudes of about 200, 250, and over 480 km at locations 38° S, 12° W; 50° S, 8° W; and 50° S, 10° W, respectively.

We observed with *Explorer IV* the effects of all five of the bursts in populating the geomagnetic field with energetic electrons. Despite the large yields of Teak and Orange, the incremental effects on the existing population of trapped particles were small and of only a few days lifetime because of the atmospheric absorption corresponding to the low altitudes of injection.

The three higher-altitude Argus bursts produced clear and well-observed effects and gave a great impetus to understanding geomagnetic trapping. About 3% of the available electrons were injected into durably trapped orbits. The apparent mean lifetime of the first two of these artificial radiation belts was about three weeks and of the third, about a month. In each of the three cases a well-defined Störmerian shell of artificially injected electrons was produced. Worldwide study of these shells provided a result of basic importance—a full geometrical description of the locus of trapping of "labeled" particles. Also, we found that the physical nature of the Argus radiation, as characterized by our four *Explorer IV* detectors, was quite different than that of the pre-Argus radiation, thus dispelling the suspicion that the radiation observed by *Explorers I* and *III* had originated from Soviet nuclear bomb bursts.

A comprehensive ten-day workshop on interpretation of the Argus observations was conducted at Livermore in February 1959. The physical principles of geomagnetic trapping were greatly clarified at this workshop.

To us, one of the principal puzzles had been the durable integrity of a thin radial shell of electrons despite the irregular nature of the real geomagnetic field and the existence of both radial and longitudinal drift forces resulting from gradients in the magnetic field intensity. Previously, we were familiar with Liouville's theorem; we understood the role of the first adiabatic invariant of Alfvén in governing trapping along a given magnetic line of force; and we understood the effects of the radial component of the gradient of the magnetic field intensity B in causing longitudinal drift in an axially symmetric field. But the longitudinal component of the gradient of B in the real geomagnetic field seemed to imply irregular drift in radial distance and hence in radial spreading, contrary to observation. The puzzle was immediately solved by Theodore Northrop in a tutorial lecture at the workshop [*Northrop and Teller,* 1960; *Northrop* , 1963; see also *Kellogg,* 1959a,b]. He invoked the second and third adiabatic invariants of cyclic motion to account for the observations. These invariants had been proven previously by *Rosenbluth and Longmire* [1957] and applied to plasma confined by a laboratory magnetic field. A specific application of these principles was *McIlwain's* [1961] concept of the L-shell parameter for the reduction of three-dimensional particle distributions to two-dimensional ones—a concept that greatly aided interpretation and has permeated the entire subsequent literature of magnetospheric physics. The adiabatic conservation and nonadiabatic violation of the three invariants have proved to be central to understanding trapped particle motion and to play a basic role in all of magnetospheric physics. In effect, they supplant the rigorous integral of motion found by Störmer for an axially symmetric magnetic field and make it possible to understand trapped particle motion and the diffusion of particles when the conditions for conservation of the three invariants are violated by time-varying magnetic and electric fields. The three invariants correspond to the three forms of cyclic motion, with quite different periods, into which the Störmerian motion of a charged particle in an approximately dipolar magnetic field can be analyzed. The first is the gyro motion of the particle around a field line; the second is the latitudinal oscillation of the guiding center (the center of the cylinder on which the helical motion of the particle occurs) of the particle's gyro motion; and the third is the time-averaged cyclic drift of the guiding center through 360° of longitude.

The entire Argus operation was conducted in secret as were the reduction and interpretation of the observations, all under the general supervision of the Advanced Research Projects Agency of the Department of Defense. But in mid-March 1959 after much internal discussion, the Department of Defense and the Atomic Energy Commission decided to declassify the major features of the Argus tests.

The National Academy of Sciences agreed to sponsor an open symposium emphasizing the scientific aspects of the tests and the relevance of the results to understanding the dynamics of geomagnetic trapping and related matters. This "Symposium on Scientific Effects of Artificially Introduced Radiations at High Altitudes" was held on 29 April 1959. Seven papers were presented at that time, and were later published in both the *Proceedings of the National Academy of Sciences* [1959] and the *Journal of Geophysical Research* [1959].

The combination of all this initial work and of many subsequently published papers and still-classified reports clarified many aspects of magnetospheric physics and gave a great impetus to the subject. We also showed conclusively that the trapped radiation observed by *Explorers I, III*, and *IV* before the tests was different in character than was the radiation produced by nuclear bursts and was therefore of natural origin.

6. EARLY CONFIRMATIONS OF THE INNER RADIATION BELT AND DISCOVERY OF THE OUTER RADIATION BELT—EXPLORER IV AND PIONEERS I, II, III, AND IV

As described in the previous sections, *Explorer IV* was the first satellite to carry a set of radiation detectors that had been designed with knowledge of the geomagnetically trapped radiation and of its intensity. For example, the Geiger tubes (Anton Type 302) that we selected for *Explorer IV* had an effective cross-sectional area one hundred times smaller than those on *Explorers I* and *III*. The two scintillation detectors had correspondingly small geometric factors and were directional in order to study the angular distribution of the radiation [*McIlwain,* 1960]. All detectors had the greatest feasible dynamic ranges. Also, the orbit of *Explorer IV* was inclined to the geographic equator by 51°, thus providing latitude coverage that was much broader than that provided by the 33° inclination orbits of *Explorers I* and *III*. A worldwide network of twenty-two regular telemetry stations received data in real time.

Explorer IV yielded a massive body of data on the natural trapped radiation as well as on that injected by the five nuclear bomb bursts—Teak, Orange, and Argus I, II, and III—during the period from launch on 26 July to 19 September 1958. All detectors operated properly and remained within their dynamic ranges. Data were obtained for five days before the Teak test on 1 August and for thirty-one days before the first Argus test on 27 August. The effects of all five of the bursts were clearly recognized in the data and were easily segregated from the effects of the natural radiation.

In early December 1958 the Iowa group completed a paper on *Explorer IV* observations of the natural radiation, omitting those of the artificially introduced radiation. This paper, which was published in the March 1959 *Journal of Geophysical Research* [*Van Allen et al.*, 1959], gave the first comprehensive body of observations on the subject and provided massive confirmation of the results of *Explorers I* and *III*. The initial range of altitude was from 262 to 2,210 km. The diverse detectors on *Explorer IV* yielded both particle intensities and energy fluxes as well as angular distributions and crude intensity-range data for the radiation, all as a function of position along the orbit. Among other findings were the facts that at constant altitude there was a relative minimum value of intensity (the "slot") at about 45° geomagnetic latitude, both north and south, and an apparent high-latitude boundary of the trapped radiation at about 65° latitude. The radiation at low latitudes was markedly more penetrating than that at high latitudes. These two findings presaged our later *Pioneer III* findings of two major radiation belts having particle populations of distinctively different character. The possibility of two distinct radiation belts was considered at the time that we prepared the *Explorer IV* paper but was omitted in the final draft in favor of a single belt with a low-altitude slot at 45°. This mistake in our conjectural extension of isointensity contours was made clear by *Pioneer III* data obtained only a few days after submission of the *Explorer IV* paper. The mistake was not corrected in proof in order to preserve the integrity of the original *Explorer IV* paper.

Despite the variety of information obtained with the four different detectors on *Explorer IV* including temporal variations [*Rothwell and McIlwain*, 1960], it was frustratingly difficult to reach firm conclusions on the identity and energy spectra of the particles responsible for the responses of the detectors. Also, we fully realized that much of the trapped radiation might lie at energies too low to be registered by our detectors. Because of the universal presence of electrons in ionized matter and because of the dominant abundance of hydrogen in astrophysical matter, we made the plausible assumption that the trapped radiation consisted of an admixture of electrons and protons having unknown energy spectra which might very well be different from each other and that were clearly a function of position in space. We made many different trial assumptions but were unable at that time to demonstrate any clear conclusions on these matters. We did state sample absolute intensities based on three alternative interpretations: *(a)* penetrating protons, *(b)* penetrating electrons, and *(c)* nonpenetrating electrons (via bremsstrahlung); and we suggested that the radiation might well be a mixture of all three, with the mix a function of position.

In a memorandum of 23 May 1958, entitled "Radiation Measurements with Lunar Probes," I outlined the desirability of observing the full radial structure and outer boundary of the radiation belt by means of an approximately radial scan that would be provided by a trajectory leading to the moon. This proposal was viewed favorably by the U.S. National Committee for the International Geophysical Year, and we again joined with the Jet Propulsion Laboratory and the Army Ballistic Missile Agency in preparing suitable payloads for the planned flights on Juno II launch vehicles (upgraded versions of Jupiter C).

We were also invited to help prepare radiation detectors for two Air Force moon flights, using Thor-Able launch vehicles. McIlwain equipped a small Anton type 706 ionization chamber with a d.c. logarithmic amplifier for the scientific payload of each of the two spacecraft *Pioneers I* and *II*. In order to provide against any possibility of saturation, he designed the system for a dynamic range of 1 to 10^6 roentgen per hour. *Pioneer I* was launched on 11 October 1958 and reached a maximum geocentric distance of about 19 earth radii before falling back to the earth, thus making the first passage through the radiation belt region to altitudes above 2,200 km. Despite some instrumental difficulties, *Rosen et al.* [1959b] reported results generally confirmatory of those with *Explorer IV*.

The third stage of the vehicle for *Pioneer II* failed to ignite following an otherwise successful launch on 7 November 1958 from Cape Canaveral. Nonetheless, the payload including its radiation detecting ionization chamber was propelled to a maximum altitude of 1,550 km and operated properly. In approximate agreement with *Explorer IV*'s data, the radiation intensity at a nearly constant altitude of 1,525 km increased by a factor of 13 as the spacecraft moved southward from 31° to 24° N. More importantly, the combination of absolute particle intensity data from the Geiger tubes on *Explorer IV* and data from the ionization chamber on *Pioneer II* yielded a rough determination that the average specific ionization was about five times its minimum value for a charged particle. In consideration of this finding and the necessary range of the radiation, it was concluded that the responsible radiation must be dominantly protons of energy $\sim$ 120 MeV [*Rosen et al.*, 1959a]. A similar conclusion on somewhat different grounds was reached by John Simpson, who had a wide-angle triple-coincidence array of semiproportional counters shielded by 5 g cm^{-2} of lead on *Pioneer II*.

Another Iowa student, Louis A. Frank, and I prepared a pair of miniature Geiger tubes and associated power supplies and electronic circuitry with large dynamic range for flight on the first two ABMA moon flights. One of the tubes was the primary radiation detector and the other, smaller one was arranged as an "ambiguity-resolver" in case the radiation was sufficiently intense to drive the first tube over the top of its characteristic curve of apparent counting rate versus true counting rate. The telemetry transmitter (960.05 MHz) and antenna, battery pack, and payload structure and shell were built by the JPL, which also conducted the necessary environmental tests (acceleration, vibration, and thermal vacuum) and established two special telemetry receiving stations—one near Mayaguez, Puerto Rico (10-ft. dish antenna), and the other at Goldstone Lake, California (86-ft. dish). *Pioneer III* was launched from Cape Canaveral on 6 December 1958. Earth escape velocity was not achieved. The payload reached a maximum geocentric distance of 107,400 km (17 earth radii) and fell back to the earth on the following day. Excellent data were obtained on both the outbound and inbound legs of the trajectory. From our point of view the flight was much more valuable than it would have been if it had flown to the moon because of the two quite different paths through the radiation belt that occurred. This flight established the large-scale features of the distribution of geomagnetically trapped particles including its outer boundary, and, in conjunction with the lower altitude data from similar detectors on *Explorer IV*, clearly defined two major, distinct radiation belts, both of toroidal form encircling the earth with their planes of symmetry at the geomagnetic equatorial plane [*Van Allen and Frank,* 1959a; *Van Allen,* 1959]. The inner belt exhibited maximum intensity (as measured with a Geiger tube shielded by 1 g cm^{-2} of material) at a radial distance of 1.4 earth radii, and the outer, more intense and much larger belt exhibited a maximum intensity at about 3.5 earth radii. A "slot" or local minimum intensity was observed between the two major belts. A meridian cross-section of isointensity contours was prepared using *Pioneer III* and *Explorer IV* data. This diagram has continued to have essential validity throughout all subsequent work.

A payload of similar equipment was launched as *Pioneer IV* on 3 March 1959. Earth escape velocity was achieved, but the payload missed the moon by a radial distance of 62,000 km and continued into a heliocentric orbit. Data were received by the Cape Canaveral, Mayaguez, and Goldstone stations. A valuable addition to telemetry reception for this flight was provided by the 250-ft. Jodrell Bank dish, which obtained data of good quality to a geocentric distance of 658,300 km (103 earth radii), far beyond the moon's orbit. The *Pioneer IV* radiation data confirmed the inner zone/outer zone structure of the geomagnetic trapping region and showed a much greater intensity in the outer zone than that in early December 1958 [*Van Allen and Frank,* 1959b; *Snyder,* 1959]. This enhanced intensity was attributed to unusually strong solar activity during the few days preceding the flight of *Pioneer IV*, whereas there was an especially quiet period before and during *Pioneer III*'s flight. Also, the interplanetary value of the cosmic-ray intensity was well determined at points remote from the earth for the first time. No effect of the moon was detected.

7. CONCLUDING REMARKS

During 1958 the Iowa satellite research group consisted principally of three talented and assiduous graduate students (Ludwig, McIlwain, and Kasper), an able, tough-minded young undergraduate student (Frank), a recent Ph.D. assistant professor (Ray), and myself. Both Ray and I were simultaneously teaching courses, while I was also serving as head of the Department of Physics and as a member of various national planning committees and panels. We had the services of three skilled instrument makers, J. George Sentinella, Edmund Freund, and Robert Markee. Within a fourteen-month period we provided the principal scientific instrumentation for *Explorers I, II, III, IV*, and *V* and for *Pioneers III and IV*; and a central portion of the instrumentation for *Pioneers I and II*. Of these nine missions, seven yielded valuable radiation data; only *Explorers II* and *V* failed to do so—both because of failure of launch vehicles. The work of preparing the detector systems for *Explorers I, II*, and *III* began in 1956 and was carried forward principally by Ludwig, who devised many novel circuits using the then new technology of transistor electronics and designed and nurtured the development of the miniature magnetic tape recorder [*Ludwig,* 1959]. He temporarily transferred to the Jet Propulsion Laboratory from November 1957 to April 1958 to adapt the Iowa apparatus to the payloads of *Explorers I, II*, and *III* (Deal I, Deal II A, and Deal II B). Apparatus for the other satellites and moon shots was conceived, designed, built, tested, and calibrated by Ludwig, McIlwain, Frank, and me after late April 1958. Ray and Kasper provided a wealth of theoretical help and assistance with data reduction, orbital calculations, and the like. Our early work was supported in part by the state of Iowa and in part by the Army Ordnance Department, the Office of Naval Research, the Jet Propulsion Laboratory, and the National Science Foundation. But despite the multiplicity of sources of support, we had a minimal burden of paper work and enjoyed extraordinarily free and entrepreneurial working circumstances.

With continuing support by the Office of Naval Research and soon thereafter and on a larger scale by the National Aeronautics and Space Administration, the scope and depth of our terrestrial magnetospheric research has expanded greatly and has been extended to

- 1978: ISEE 3, 12 August;
- 1981: DE 1 (Dynamics Explorer 1), 3 August;
- 1982: PDP (Plasma Diagnostics Package)—flown on Space Shuttle Columbia, 22 March;
- 1984: AMPTE-IRM, 16 August;
- 1985: PDP (Upgraded Version)—flown on Space Shuttle Challenger, 29 July;
- 1989: GALILEO, 18 October;
- 1990: CRRES, 25 July;
- 1992: GEOTAIL, 24 July;
- 1996: POLAR, 17 February.

Hundreds of our published papers have reported the findings of these missions and as of 1996, seventy-five graduate students have earned Ph.D. degrees in space physics at Iowa.

Acknowledgments. I am grateful to Alice Shank for help in preparing this material.

REFERENCES

Byrd, R. E., *Discovery: The Story of the Second Byrd Antarctic Expedition*, G. P. Putnam's Sons, New York, 1935.

Chapman, S., and J. Bartels, *Geomagnetism*, Vols. I and II, Oxford at the Clarendon Press, 1940.

Freden, S. C., and R. S. White, Protons in the Earth's magnetic field, *Phys. Rev. Lett.*, *3*, 9-10 and 145, 1959a.

Freden, S. C., and R. S. White, Particle fluxes in the inner radiation belt, *J. Geophys. Res.*, *65*, 1377-1383, 1959b.

Gold, T., Motions in the magnetosphere of the Earth, *J. Geophys. Res.*, *64*, 219-224, 1959a.

Gold, T., Origin of the radiation near the Earth discovered by means of satellites, *Nature (London)*, *183*, 355-358, 1959b.

Hazard. D. L., *Directions for Magnetic Measurements*, U.S. Government Printing Office, 3rd ed., Washington, D.C., 1930.

Kellogg, P. G., Possible explanation of the radiation observed by Van Allen at high altitudes in satellites, *Il Nuovo Cimento, Ser. X*, *11*, 48-66, 1959a.

Kellogg, P. G., Van Allen radiation of solar origin, *Nature (London)*, *183*, 1295-1297, 1959b.

Loftus, T. A., Disturbance of the inner Van Allen belt as observed by Explorer I, M.S. Thesis, University of Iowa, 1969.

Ludwig, G. H., Cosmic-ray instrumentation in the first U.S. Earth satellite, *Rev. Sci. Instrum.*, *30*, 223-229, 1959.

McIlwain, C. E., Direct measurements of particles producing visible auroras, *J. Geophys. Res.*, *65*, 727-747, 1960.

McIlwain, C. E., Coordinates for mapping the distribution of magnetically trapped particles, *J. Geophys. Res.*, *66*, 381-391, 1961.

Meredith, L. H., M. B. Gottlieb, and J. A. Van Allen, Direct detection of soft radiation above 50 kilometers in the auroral zone, *Phys. Rev.*, *97*, 201-205, 1955.

Northrop, T. G., and E. Teller, Stability of the adiabatic motion of charged particles in the Earth's field, *Phys. Rev.*, *117*, 215-225, 1960.

Northrop, T. G., *The Adiabatic Motion of Charged Particles*, Interscience Publishers, New York, 1963.

Porter, H. H., S. Karrer, R. D. Mindlin, and J. A. Van Allen, Rugged Vacuum Tube, U.S. Patent No. 3,113.235 of 3 December 1963.

Rosen, A., P. J. Coleman, Jr., and C. P. Sonett, Ionizing radiation detected by Pioneer II, *Planet. Space Sci.*, *1*, 343-346, 1959a.

Rosen, A., C. P. Sonett, and P. J. Coleman, Jr., Ionizing radiation at altitudes 3,500 to 3,600 kilometers, Pioneer I, *J. Geophys. Res.*, *64*, 709-712, 1959b.

Rosenbluth, M. N., and C. L. Longmire, Stability of plasmas confined by magnetic fields, *Ann. Phys.*, *1*, 120-140, 1961.

Snyder, C. W., The upper boundary of the Van Allen radiation belts, *Nature (London)*, *184*, 439-440, 1959.

Störmer, C., Sur des trajectories des corpuscles electres dans l'space sons l'action du magnetisme terrestre. Chaptre IV, *Arch. Sci. phys. et naturelles*, *24*, 317-364, 1907.

Störmer, C., *The Polar Aurora*, Oxford at the Clarendon Press, 1955.

Van Allen, J. A., A. Ellett, and D. S. Bayley, Cross-section for the reaction $H^2 + H^2 \rightarrow H^1 + H^3$ with a gas target, *Phys. Rev.*, *56*, 383, 1939.

Van Allen, J. A., and N. F. Ramsey, Jr., A technique for counting high energy protons in the presence of fast neutrons, *Phys. Rev.*, *57*, 1069-1070, 1940.

Van Allen, J. A., and N. M. Smith, Jr., The absolute number of quanta from the bombardment of fluorine with protons, *Phys. Rev.*, *59*, 501-508, 1941a.

Van Allen, J. A., and N. M. Smith, Jr., The absolute cross section for the photodisintegration of deuterium by 6.2 MeV quanta, *Phys. Rev.*, *59*, 618-619, 1941b.

Van Allen, J. A., Direct detection of auroral radiation with rocket equipment, *Proc. Natl. Acad. Sci.*, *43*, 57-62, 1957.

Van Allen, J. A., Editor, *Scientific Uses of Earth Satellites*, University of Michigan Press, 1956 and 2nd ed. 1958.

Van Allen, J. A., G. H. Ludwig, E. C. Ray, and C. E. McIlwain, Observation of high intensity radiation by satellites 1958 alpha and gamma [Explorers I and III], *Jet Propulsion*, September, pp. 588-592, 1958.

Van Allen, J. A., The geomagnetically-trapped corpuscular radiation, *J. Geophys. Res.*, *64*, 1683-1689, 1959.

Van Allen, J. A., and L. A. Frank, Radiation around the earth to a radial distance of 107,400 km, *Nature (London)*, *183*, 430-434, 1959a.

Van Allen, J. A., and L. A. Frank, Radiation measurements to 658,300 km with Pioneer IV, *Nature (London)*, *184*, 219-224, 1959b.

Van Allen, J. A., C. E. McIlwain, and G. H. Ludwig, Radiation observations with satellite 1958 ϵ, *J. Geophys. Res.*, *64*, 271-286, 1959.

Van Allen, J. A., Observations of high intensity radiation by satellites 1958 alpha and 1958 gamma [Explorers I and III], *IGY Satellite Report, 13*, 1-22, National Academy of Sciences, Washington, D.C., 1961. [Reprinted with minor editing in *Space Science Comes of Age*, edited by P. A. Hanle and V. D. Chamberlain, pp. 58-75, Smithsonian Institution Press, Washington, D.C., 1981.]

Van Allen, J. A., Early rocket observations of auroral bremsstrahlung and its absorption in the mesosphere, *J. Geophys. Res., 100*, 14,484-14,497, 1995.

Wiley, C. C., Real paths for five meteors, *Popular Astronomy*, Vol. XL, 500-501, Goodsell Observatory of Carleton College, Northfield, Minnesota, 1932.

Yoshida, S., G. H. Ludwig, and J. A. Van Allen, Distribution of trapped radiation in the geomagnetic field. *J. Geophys. Res., 65*, 807-813, 1960.

Other Writings of an Historical Nature:

Molfetto, W., James A. Van Allen: Space scientist: A career biography through the year 1958, M.S. Thesis, University of Iowa, 1994.

Van Allen, J. A., Dynamics, composition and origin of the geomagnetically-trapped corpuscular radiation, *Transactions of the International Astronomical Union, XIB*, 99-136, 1962.

Van Allen, J. A., *Origins of Magnetospheric Physics*, Smithsonian Institution Press, Washington, D.C., 1983. A major source of the present abridged account.

Van Allen, J. A., What is a space scientist? An autobiographical example, *Ann. Rev. Earth Planet. Sci., 18*, 1-26, 1990.

Van Allen, J. A., Early days of space science, *J. British Interplanetary Society, 41*, 11-15, 1988.

Van Allen, J. A., 1994 Kuiper Prize Lecture: Electrons, protons, and planets, *Icarus, 122*, 209-232, 1996.

Department of Physics and Astronomy, University of Iowa, Iowa City, IA 52242. (e-mail: james-vanallen@uiowa.edu)

From Nuclear Physics to Space Physics by Way of High Altitude Nuclear Tests

Martin Walt

Starlab, Stanford University, Stanford, California

At the time satellites became available I was working in nuclear research at the Lockheed Missiles Systems Division. This work naturally led to participating in the high altitude nuclear weapons effects tests and the creation of artificial radiation belts. Efforts to unravel the phenomena occurring in these experiments have uncovered several features of trapped radiation source and loss mechanisms and have led to a better understanding of the magnetosphere.

1. INTRODUCTION

Like many of my colleagues I was swept into Space Physics by events of the late 1950's. I was trained as an experimental nuclear physicist, receiving a B.S. from Caltech in 1950 and a PhD in 1953 from the University of Wisconsin under the guidance of Professor H. H. Barschall. After graduate school I spent three years (1953-1956) at the Los Alamos Scientific Laboratory (LASL) doing neutron scattering experiments. It was a very productive period for me thanks to the excellent facilities of Los Alamos. However, the living conditions in an isolated community had drawbacks, and in 1956 I left LASL and joined the newly-formed Missiles System Division of Lockheed Aircraft Company. (My closest colleague at Los Alamos, J. R. Beyster, with whom I published a number of papers, also left LASL at that time and later founded Science Applications International Corporation, which is now a $2 Billion/year enterprise.) My initial tasks at Lockheed were to get a new Van de Graaff accelerator operating and to do the criticality calculations for a reactor designed to provide electrical power for satellites. The missile and satellite programs were very new, and I imagined there were lots of problems requiring nuclear expertise. Among these problems were the effects of high altitude nuclear detonations on incoming missiles, radar, and radio communications. Neutron irradiation of crews of high flying aircraft was also of concern as well as the more distant possibility of nuclear propulsion of aircraft and rockets. There was no suspicion whatever that space contained a particle radiation hazard whose measurement and interpretation would consume much of my professional life.

Shortly after I joined Lockheed, the Nuclear Physics Group obtained a study contract from the Air Force Special Weapons Center (AFSWC, now the Air Force Philips Laboratory, Albuquerque). The contract called for the investigation of a number of effects from high altitude nuclear detonations. The radio blackout from the ionization of air and the thermal effects from optical radiation were to be evaluated. We were also to do conceptual studies of instrumentation to measure these effects in future high altitude nuclear tests. My own contribution to this work was in designing detectors to measure neutron and gamma-ray fluxes.

2. THE ARGUS EXPERIMENTS

Early in 1958 our Air Force contract monitors at Albuquerque informed us of the proposal by Nicholas Christophilos of the Lawrence Radiation Laboratory at Livermore to inject relativistic electrons into the earth's magnetic field by a high altitude nuclear detonation. (A

Discovery of the Magnetosphere
History of Geophysics Volume 7

collection of papers describing the project was published in the *Journal of Geophysical Research 64*, No. 8, 1959). By analogy with magnetic confinement mirror machines it was believed that the electrons would be trapped. The military uses of this concept were not well defined although it was asserted that synchrotron radiation from the electrons might be large enough to interfere with ground-based radar. Hazards to satellites and to man-in-space from the electron fluxes were also recognized. Many technical questions were raised about the Christophilos proposal, and a panel of experts evaluated the overall concept. A fundamental question was whether the electrons emitted by fission fragments would be trapped for a long time. Another question was the trapping efficiency, namely what fraction of the total number of beta particles released by the fission fragments would be trapped. Because the dynamics of a nuclear fireball at high altitude were poorly known, field experiments would be needed to answer these questions. The responsibility for these tests, designated as the ARGUS experiments, was assigned to ARPA (Advanced Research Projects Agency, now the Defense Advanced Research Projects Agency, or DARPA).

ARPA decided to investigate the Christofilos concept by detonating three low-yield nuclear devices at high altitude over the South Atlantic Ocean in the region of the magnetic anomaly. Since the available booster rockets had limited altitude capability, it was necessary to launch them where the surface geomagnetic field was weak to insure that some of the trapped electrons would mirror above the atmosphere and be in long-term orbits. Another benefit of a South Atlantic detonation site was its remoteness as the tests were highly classified.

In the midst of all this planning, Van Allen announced the discovery of the radiation belts based on observations with Explorer 1 and Explorer 3. The existence of a high flux of energetic particles proved that long-term trapping could occur. However, the injection efficiency possible with a nuclear device still needed confirmation by full scale space experiments using an actual nuclear detonation. To measure the trapped electrons injected in the experiments Van Allen was asked to provide instruments for two new satellites; the one that was launched successfully was designated Explorer 4.

Satellite launching was a chancy business in 1958 (as it is now) and some backup to these satellites was deemed desirable. In 1958 test flights of Atlas missiles took place regularly from Cape Canaveral (now Cape Kennedy). Our Air Force contract monitors at AFSWC conceived the idea of attaching a pod, which carried radiation detectors, to each Atlas and ejecting the pod after propellant burn-out. The pod would then follow a trajectory through the region where electrons from the bombs would be trapped. A series of four pods were to be launched on four successive Atlas vehicles with the hope that a measurement would be made within a few days after the nuclear detonation. This pod program had severe limitations since the schedule was governed by the ballistic missile test program rather than the need to measure artificial radiation belts. Furthermore, the artificial radiation belts might well form outside the Atlas trajectory. To overcome these objections AFSWC also proposed a sounding-rocket program in which about 20 sounding rockets would be launched from three stations distributed along the East Coast of the U. S. The trajectories of these rockets (with apogees of about 800 km) would cover an extensive latitude range, and, of course, the launch schedule could be set by the explosions in the South Atlantic.

The nuclear weapons effects department at AFSWC in 1958 was an unusually capable group of young Air Force officers. They were intelligent, well educated, and somewhat aggressive, but their centrifugal forces were channeled by the very charismatic Major Lew Allen, who later became AF Chief of Staff and after that Director of the Jet Propulsion Laboratory at Caltech.

There was a time period of about four months between the decision to go ahead and the actual nuclear detonations. The sounding rocket instrumentation program had to move fast and be completed quickly by people who really didn't know much about the work. Since Lockheed had a contract and a funding channel in place with AFSWC and had an experimental nuclear physics group with expertise in nuclear particle detectors, we were asked to design and build the instruments for the Atlas pods and the sounding rockets. The five-stage solid propellant rockets were assembled and launched by the Airolab Development Company in Pasadena. We hastened to build and instrument the pods and the sounding rocket payloads in a period of about 3 months. Because of the short time available the instrumentation was very crude. For the sounding rockets the detectors were eight Geiger-Muller counters of assorted sizes to give a large dynamic range. Energy sensitivity was obtained by surrounding the detectors with different amounts of shielding. The detector outputs were connected via a commutator to an amplifier and radio transmitter. The commutator sampled each detector in turn (.013 s on each detector) and sent the output pulses directly to the transmitter. Figure 1 is a schematic of the rocket instruments as seen looking down the vertical axis. In operation we could obtain "quick-look" data by sending the signal to a scope which triggered on the start of the commutation cycle. The scope then showed a series of pedestals, one for each commutator point, and the counter pulses would be superimposed on these pedestals. After some experience it was possible to estimate the number of counts on each pedestal and thus obtain the counting rate. Final data analysis would have to be done by playing tapes onto strip charts and manually counting the number of pulses on each pedestal.

We had no experience in space instrumentation but approached the problems with freshman physics techniques. We were told by the rocket people that the sounding rocket payloads would receive about 180 g's acceleration. We tested subsystems by letting them fall a measured distance onto a cushion of wax. From the length of the drop and the compression of the wax we could compute the acceleration experienced by the subsystem.

By this time Van Allen was world famous and his views on space instrumentation were considered infallible. For this reason the Air Force insisted that we get his endorsement of our rocket plans. I traveled to Iowa City in the spring of 1958 and spent several hours with him reviewing geometric factors and detector specifications. He was generous with his time, considering that he was Iowa City's most notable citizen and had many distractions and high level visitors. He approved the design and immediately called our sponsors and told them to proceed. My most vivid memory of that visit was of a phone call he received from some important General while I was in his office. As I recall, his exact words were "Yes, General, I would be happy to come to Washington to testify for your project next week. However, one of my students is taking his oral exams then and I have to be here to help him." From then on I looked on Van Allen as a voice of reason in a world gone mad.

The launch sites for the sounding rockets were located at Wallops Island, Cape Canaveral, and Ramey Air Force Base in Puerto Rico. The horizontal range of the five-stage rockets launched from these bases appeared adequate to cover the possible latitude location of the trapped electrons. By late summer all the instruments were delivered to the launch sites. For practice and to measure the pre-nuclear explosion background, we launched one rocket from each of the three sites. Two of these test rockets broke up after launch, giving us an uncomfortable feeling about the reliability of the missile configuration. As a precaution we added an additional ring of rivets to the flange joining the payload to the last rocket stage. This field modification was done with hand drills and rivet guns borrowed from one of the contractors at Cape Canaveral.

Meanwhile plans for launching the nuclear devices were progressing. The launcher was the Lockheed X-17, which was a three-stage solid propellant rocket used extensively to test re-entry nose cones for the Polaris missile. The launch

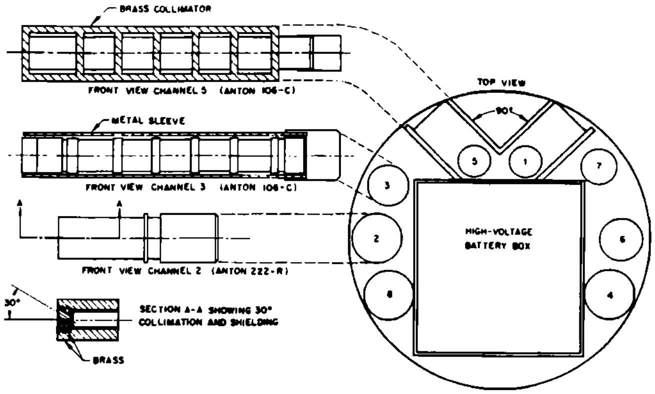


Figure 1. Instrument package containing Geiger-Muller counters used on the sounding rockets to measure Argus electrons. The eight GM tubes were mounted parallel to the rocket axis and surround the 1 KV high voltage battery. Counters 1 and 5 were collimated to view 90 degrees from each other to identify the signature of trapped electrons.

platform was the experimental missile launching ship, U.S.S. Norton Sound. I knew very little about this part of the operation at the time, but learned some of the details at meetings held to assess the results.

By late August we were ready for the real thing. On August 27 we received word that the first detonation had occurred. We then launched instrumented rockets from each of the three sites, the first rocket being launched from Cape Canaveral about 1 hour after the burst. During this stage of the experiment I was stationed at the telemetry center in Cape Canaveral looking at a scope which was connected to the telemetry receiver. The telemetry system had sufficient range to track the rockets launched at Wallops Island and Puerto Rico as well as the local launches at Cape Canaveral. On the scope were the familiar pedestals showing each position of the commutator, and as the rockets rose, pulses appeared on the pedestals to indicate counts from the Geiger counters. I watched intently during the ≈10-minute flights, focusing mostly on channel three which was an omnidirectional detector with a threshold near 1 Mev. Every 100 seconds I wrote down an estimate of the number of pulses on each pedestal, and at the end of the flight I plotted these numbers against time to see if a narrow band of trapped electrons had been crossed. After successful flights from Puerto Rico and Cape Canaveral the results were disappointing as only the normal background had registered. At that point we ceased launching and reported that no significant effects had been seen.

The second Argus test took place on August 30 and was more successful from our point of view. For this event 12 vehicles were launched from the three bases with 10

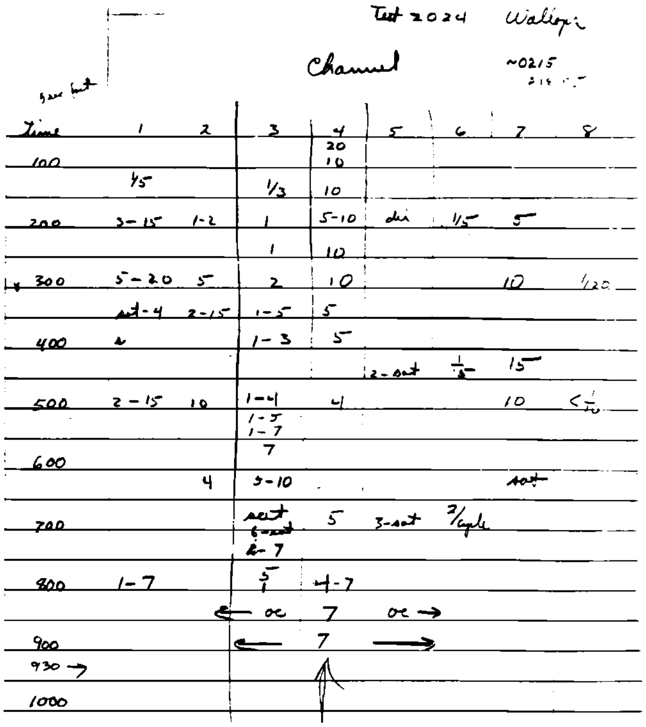


Figure 2. Page from a notebook used to record sounding rocket flight results. Note that channel 3 is saturated at 700 seconds, indicating that the detector had passed through a band of increased counting rate.

successes. Several of the rockets launched generally southeast from Wallops Island experienced short periods of increased counting rate near the ends of their trajectories. Figure 2 is a copy of the notebook page I used to record the "quick-look" data. Note that at about 700 sec, the number 3 pedestal was saturated. This band of electrons was sampled over a period of 18 hours by five rockets, and its decay could be seen. We all felt elated at that point. We learned later that Explorer 4 was making global measurements during this period and that the satellite data were much more comprehensive than the limited rocket results. However, the rocket instrumentation was more elaborate and gave the best determination of the energy spectrum of the electrons. Descriptions of the rocket experiment were published in the *Journal of Geophysical Research* (*Allen et al.*, 1959; *Cladis and Walt*, 1962). The first paper was a report of the overall rocket project, and the latter paper contained a more detailed analysis of the electron measurements during the second detonation, including post-flight calibration of a spare set of detectors.

The Argus tests did have important scientific fall-out. The narrow bands of radiation detected world wide allowed an experimental verification of McIlwain's L parameter (*McIlwain*, 1961) for organizing trapped radiation data in the earth's distorted field. The rapid decay of the Argus electron fluxes with negligible spread in latitude remained a stumbling block for radiation belt theory for years as it implied that the natural belt electrons had short lifetimes (a few weeks) and did not migrate in latitude. Many years later the Saint Louis University group (*Manson et al.*, 1968) took a closer look at the Explorer 4 data and detected a spreading of the electron bands in L as time progressed. These data are still the best evidence for radial diffusion of electrons in the region just above L=2.

The Atlas pod program, which had been intended to sample the Argus belts, was unsuccessful in that respect. Although several pods were carried in missile launches, the release mechanism failed to function and the pods were not deployed. Later the release mechanism was re-engineered by AFSWC and a successful flight through the inner belt was achieved. This pod carried a magnetic spectrometer which unambiguously detected trapped electrons, thus showing that the inner belt contained electrons as well as protons (*Holly and Johnson*, 1960).

Following the Argus tests the AFSWC wished to improve their capability for future experiments of this type. They arranged for a series of Javelin sounding rockets and contracted Lockheed to develop the instrumentation. The Lockheed group, which included Bill Imhof, John Cladis, Dick Johnson, Lloyd Chase and myself, built a series of magnetic and scintillation spectrometers. When launched on a Javelin in 1959, one of these magnetic spectrometers made the best measurement to that date of the trapped electron spectrum and angular distribution in the natural radiation belt. The results were reported at the COSPAR meeting in Nice, France (*Walt, et al.*, 1961) and in the *Journal of Geophysical Research* (*Cladis et al.*, 1961). This precise electron spectrum allowed us to exclude neutron decay as a source of the trapped electrons, although many scientists continued to believe in the neutron decay explanation as late as 1965. Figure 3 is a graph of the electron spectrum derived from the Javelin rocket measurements.

About 1960 some friction developed between Lockheed and AFSWC regarding scientific publication of the results of this work. The Air Force took the position that since they had organized and sponsored the research, they were entitled to the publication rights. The Lockheed scientists felt that their efforts had a creative, scientific component and that the publications should also reflect their contributions. Partially due to this impasse, the experimental space work at Lockheed was brought into the Air Force Center at Albuquerque and eventually transferred to AFCRL (Air Force Cambridge Research Laboratory, now the Air Force Philips Laboratory). After 1961 space research at Lockheed was sponsored primarily by the Office of Naval Research (ONR) and the Defense Atomic Support Agency (DASA), organizations which did not have competing research activities.

After the Argus tests were completed, various meetings were held to assess the results. One such meeting took place at ARPA shortly after the events but before we had done more than a cursory analysis of the rocket data. Major Lew Allen and Captain Roddy Walton of AFSWC and I represented the sounding rocket program. Because the data processing with the crude telemetry format was so time consuming we were only able to give an overall impression of the available data. It was clear at the meeting that Van Allen's satellite data were far superior in spatial and temporal coverage, although the larger number of detectors and varied shielding thicknesses gave us a better shot at the energy spectrum of the electrons. Discussions of the inventory of trapped electrons were lively although the data were very fragmentary. My recollection was that Christofilos argued forcefully for a high trapping efficiency in spite of the limitations of the data.

In the months that followed I puzzled over the rapid decay rate of the electrons. The initial results seemed to show a decay which went as $t^{-1.2}$, which is close to the approximate decay law for the emission of electrons by gross fission fragments. John Cladis and I toyed with the idea that the fission fragments were lodged somewhere along the field line and continuously emitted electrons which were quickly lost. If this were the case, the observed flux would decay at the $t^{-1.2}$ rate. However, further analysis convinced us that there were not enough fission fragments to produce the observed belts in this way.

Christofilos had made estimates of the trapping lifetime assuming the removal mechanism was collisions with atmospheric atoms. He reasoned that since the atmospheric density was greatest at the mirroring points of the electrons, most collisions took place there. Since a mirroring particle moves at 90° to the magnetic field, a pitch angle change in any direction will cause it to mirror at a lower altitude. Cumulative small angle scattering will therefore systematically lower the mirroring point and cause the particle to be lost. This approximation was further developed by Jack Welch and Bill Whitaker at AFSWC who did extensive computer calculations of loss rates and the evolution of the Argus belts. They found rather good agreement between the experimental results and their calculations. This theory was also used extensively by Wilmot Hess of NASA and various co-authors for many years.

Meanwhile Fred Singer, Bill MacDonald, and Bob

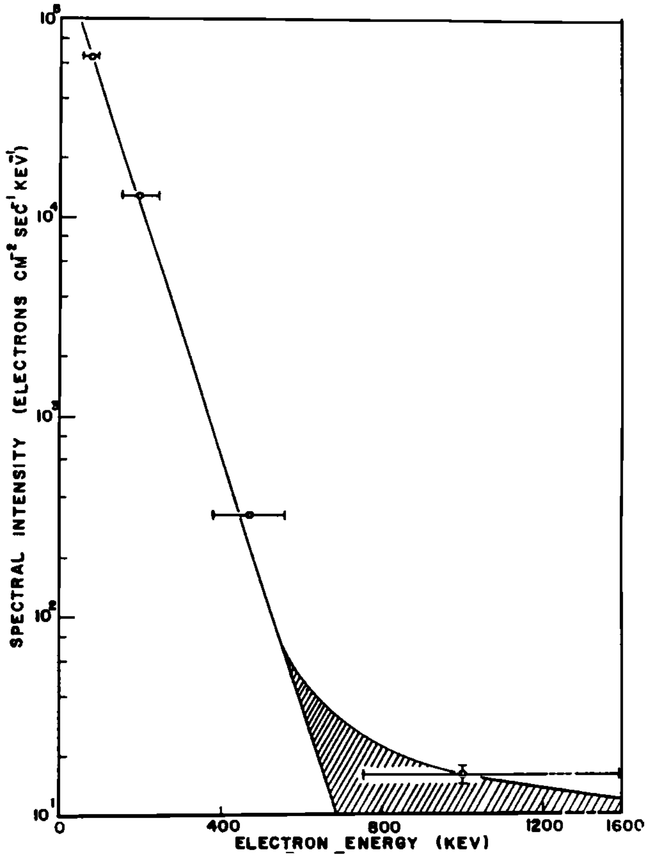


Figure 3. Trapped electron energy spectrum obtained with a magnetic spectrometer flown on a Javelin rocket in 1959. The shaded area indicates the uncertainty due to scattering of low energy electrons into the higher energy channels.

Wentworth at the University of Maryland had considered a diffusion approach to the atmospheric scattering problem using the Fokker-Planck formalism (Wentworth et al., 1959). MacDonald was an expert on Fokker-Planck theory having just completed a classic paper on the subject with Marshall Rosenbluth and Dave Judd while he was at the University of California at Berkeley. The Fokker-Planck approach showed that scattering collisions would raise or lower the mirroring points with nearly equal probability. However, the Wentworth, MacDonald, and Singer approach did not include a realistic magnetic field nor did it include energy loss in the collisions.

In the summer of 1960 Bill MacDonald was a consultant at the Lockheed Missiles and Space Company working on various topics in nuclear physics. We met by chance and quickly found a kindred interest in the trapped electron lifetime problem. By the end of the summer we had formulated the general case of magnetically trapped, relativistic electrons in a scattering atmosphere and submitted a paper to *Annals of Physics* (*MacDonald and Walt*, 1961). The diffusion formulation of the trapped electron phenomena led to a solution by separation of variables in which the pitch angle distribution could be represented as a superposition of normal modes. This solution was an approximation because the energy loss and scattering terms did not have the same dependence on pitch angle and could not be separated completely. Nevertheless, the approximate treatment was easy to use and was a major improvement on methods which neglected either scattering or energy loss. With these tools we quickly calculated the equilibrium spectrum of trapped electrons from decaying neutrons (a popular source mechanism in those days) and found it to be incompatible with our Javelin rocket measurements. (*Walt and MacDonald*, 1961) Paul Kellogg had earlier calculated the electron spectrum without including energy loss and had come to similar conclusions (*Kellogg*, 1960). We also showed that both energy loss and scattering for electrons occurred at comparable rates and that the monotonic lowering of mirroring points was incorrect (*Walt and MacDonald*, 1962)

In late 1961 I began to assemble data on atmospheric densities and to consider how to program the diffusion formalism for numerical integration of the equations as an initial value problem. By inserting the bomb source we could then calculate the evolution of the electron flux at any location as a function of time. Our intention was to apply this technique to the Argus experiments. At this point I enlisted the help of Bill Francis, a mathematician and computer programmer, who set up the partial differential equation in finite difference form. While this technique is standard now, in 1960 no such program was available, and constructing one was a formidable task. However, by 1962 we were prepared to calculate the evolution of a distribution of geomagnetically trapped electrons in the presence of atmospheric collisions. At this point the Starfish high altitude nuclear detonation occurred which changed the application of the work and made it an urgent project.

3. STARFISH

In 1962 the USSR conducted a series of high altitude nuclear weapons tests. In response to the need for information on the effects of nuclear weapon bursts at high altitude, the U.S. quickly planned its own series consisting of several explosions at various altitudes and yields. One of the experiments, called Starfish, was a megaton-range detonation at 400 km altitude over Johnston Island in the Pacific. Starfish was designed to ionize a large area of the upper atmosphere and to measure the degradation of radar sensitivity and the possible outage of communications systems. There was no interest or concern about trapped radiation except to plan the test so that there would be little or no long-term trapping. The field line passing through the detonation point of Starfish was $L \approx 1.12$, and it was believed that electrons injected on this field line would not survive a complete drift around the earth.

There was much uncertainty about the way the bomb debris, including the fission fragments would behave. It was thought the ionized atoms would expand against the magnetic field, forming a magnetic cavity. After a second or so this cavity would collapse and the fragments would move parallel to the field lines. Whether instabilities would accelerate the collapse or whether debris would escape the bubble was not known. At that time the magnetic confinement fusion program was plagued with various instabilities which shortened confinement times, and it was thought that similar conditions might affect nuclear detonations. To measure the motion of fission fragments after the detonation Frank Vaughn and I proposed a "gamma-ray scanner" as part of the Starfish program. This device was a highly collimated gamma-ray sensor which would be carried on a rocket launched several hundred kilometers from the Starfish detonation point. The sensor would scan the detonation region and by recording gamma rays would map the location of fission fragments during the bubble expansion and subsequent motion. However, we did not market this proposal aggressively enough, and the task was assigned to an inferior concept which gave no useful information.

Starfish was detonated July 9, 1962. As everyone now knows, large numbers of electrons were injected into the magnetosphere causing the premature death of several satellites. The extent of the belt and its expected lifetime were now crucial questions affecting the military and civilian space programs of all nations. To assess the situation NASA and DOD hurriedly convened a meeting at Goddard Space Flight Center on September 10-11, 1962. A multitude of people were there even though a "secret" security clearance was required. At that time Lockheed had several nuclear weapons effects contracts and was invited to send a representative. I was on vacation in the High Sierras when the meeting was announced and cut short my first vacation in years to attend. Other attendees had also truncated their vacations. The very large crowd heard reports by a number of people who had made measurements of the Starfish electron fluxes as well as the optical and radio phenomena accompanying the detonation. By far the most informative talks were Van Allen's presentation of data from INJUN 1 and Walt Brown's report of particle measurements with Telstar. Unfortunately, INJUN 1 had a

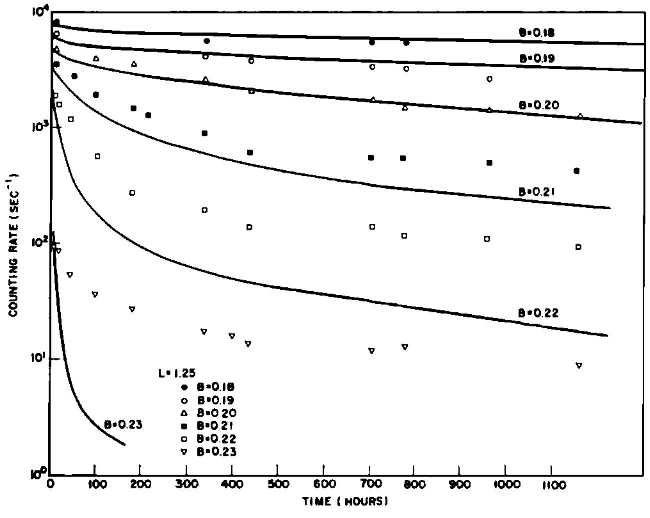


Figure 4. Time dependence of Starfish electron flux near L=1.185 at various values of magnetic field B. Symbols denote the counting rate of a detector on Injun 1, and the solid curves are theoretical values calculated from the initial distribution subject to atmospheric collisions.

low apogee and did not sample the entire extent of the belts, while Telstar was launched after the Starfish explosion and therefore had not measured the natural pre-Starfish radiation. Furthermore, both satellites carried detectors designed to measure the natural radiation rather than the high fluxes of relativistic electrons emitted by fission fragments.

Before leaving Palo Alto for the meeting I had time to do a computer run with our trapped radiation code for the low L shells of the Starfish test although I had to guess at the initial electron distribution. At the meeting I compared these results with Van Allen's observed decay rates and we were pleased at the agreement. We both immediately recognized the implications. If the losses of trapped electrons could be attributed to atmospheric collisions, then we could calculate the loss rates of the natural electron belts. Then knowing the flux in the natural belts we could calculate the characteristics of the electron source. In 1962 the only quantitative source which had been suggested for the natural belt was the decay of neutrons from cosmic ray albedo. As mentioned earlier this source of electrons had been discarded (by most of us) because of energy spectrum considerations. A quantitative knowledge of the source intensity would be important in testing future theories of the source mechanism.

With the DOD's keen interest in the lifetime of the Starfish belt, funding for my calculations followed immediately. With Les Newkirk and Bill Francis we assembled the best atmospheric density data available and with a multipole expansion of the earth's field we calculated the average atmosphere encountered by trapped electrons at L shells appropriate for Starfish. We even corrected for the variation of longitudinal drift velocity due to the distorted geomagnetic field, following a suggestion by Carl McIlwain. The upshot of all this number crunching was that we were able to show that the observed electron decay below L≈1.3 agreed with atmospheric scattering calculations, while above L≈1.3 the observed lifetimes were much too short (*Walt*, 1964). This work ended the controversy over whether a diffusion equation was needed or whether all collisions scattered electrons into lower mirroring altitudes. It also eliminated atmospheric collisions as the cause for decay of the Argus radiation belts. Thus, the search for other loss mechanisms above L≈1.3 began and still continues.

Figure 4 depicts as a function of time the counting rate of one of the INJUN detectors at about L=1.185 and at several values of magnetic field B. The solid lines are the expected counting rates based on the observed initial distribution of electrons and collisions with the atmosphere. Figure 5 (from *Van Allen*, 1964) illustrates electron lifetimes as a function of L as derived from Starfish and other nuclear experiments.

To explain the short lifetimes Jim Dungey (*Dungey*, 1963) suggested that the electrons were scattered by a gyroresonant interaction with circularly polarized whistler waves. This process was investigated by Roberts (1966) for field-aligned waves and found to be unsatisfactory. Some years later extensions of the theory to include obliquely propagating waves and plasmaspheric hiss were

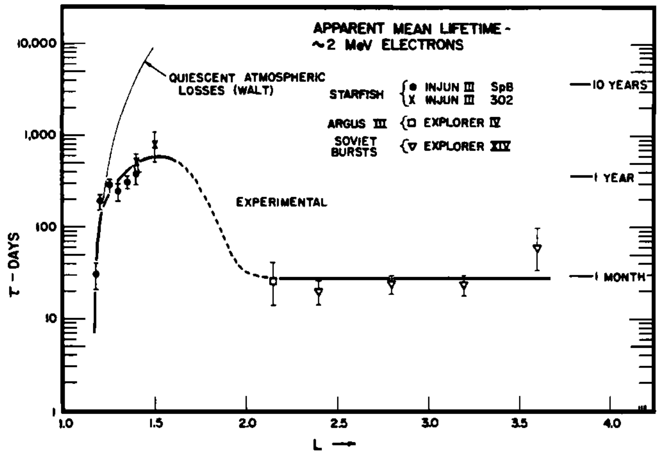


Figure 5. Electron lifetimes as a function of L. Note the break from atmospheric scattering theory near L=1.3 (from *Van Allen*, 1964).

done with more promising results (*Lyons et al.*, 1972), but better information on the earth's plasma wave environment is needed to be confident that the "other loss mechanisms" have indeed been found.

Meanwhile in the weeks after Starfish new calculations were done on injection efficiencies of nuclear weapons to make sure that additional belts would not be formed by the later detonations of the 1962 test series. In retrospect one is impressed by the speed with which these estimates were done as well as by the crudeness of the calculations. Actually the physics is very complex, and to this day the mechanisms by which Starfish injected electrons is controversial. Nevertheless, the other events of the 1962 test series, such as Bluegill, were at a sufficiently low altitude and reduced yield (compared to Starfish) that injection seemed unlikely and the test series was continued.

On the experimental side it was recognized that more and better experimental data on the Starfish belt were badly needed. The two principal satellites then in orbit were INJUN 1 and Telstar. As mentioned earlier INJUN 1 had a limited lifetime and an apogee of only 1010 km, well beneath much of the artificial belt. Telstar, which was in a better orbit for this purpose, had detectors intended to measure low energy electrons and an uncertain background. Both NASA and DOD independently pushed for additional measurements. The DOD reconnaissance satellites, which were being launched on a regular basis, were equipped on short notice with radiation detectors from Lockheed and Aerospace, although the low- altitude orbits were not very useful. More substantial contributions came from two new satellites, the DOD Starad (also known by the name 1962$\beta\gamma$) and NASA's Explorer 15, which were launched on 26 October, 1962 and 28 October, 1962, respectively.

Both of these satellites contained detectors with good energy discrimination above 1 Mev and could distinguish the fission electrons from the natural belt populations. They also had good angular resolution and could measure pitch angle distributions. Their high apogees carried them through the extended Starfish belts. Unfortunately, by the time they were launched, much of the Starfish belt above L=1.5 had decayed, so the early time evolution of the belt is still poorly known.

Instrumentation for these satellites was solicited from anyone who had on-the-shelf components available. Until that time the very limited space on prime scientific satellites was available only to a few veteran experimenters, most of whom had been involved in previous balloon and rocket experiments. Proposals from the rest of the community, even if they contained next-generation instrumentation, stood little chance of surviving the "old boy" reviewing network which heavily weighted previous space flight experience. The Starfish emergency opened up the field for many newcomers as instruments were accepted from LRL (Lawrence Radiation Laboratory, now Lawrence Livermore National Laboratory) Aerospace Corporation, Lockheed, and others. In addition to injecting electrons into the magnetosphere Starfish injected many highly qualified scientists into the space program. Some of the individuals who come to mind are Harry West of LLNL, Forest Mozer, who transited from Lockheed to Aerospace Corporation during this period, Charlie Roberts and Walt Brown of Bell Labs, Richard Kaufmann of AFSWC who is now at the University of New Hampshire, and Billy McCormac of DASA, who later organized a legendary series of international space physics symposia. Of the Lockheed Van de Graaff group almost half including myself, Dick Johnson, Dick Sharp, Ed Shelley and Bill Imhof eventually turned to full-time space research.

The Starfish event led to the best electron lifetime determinations for electrons trapped below about L=1.7 (*Van Allen*, 1964; *McIlwain*, 1963). Measurements of Starfish electrons at very low L-shells also led to values of the radial diffusion coefficient (*Farley*, 1969). One aspect of Starfish has not yet been adequately explained; namely the redistribution of inner belt protons of 55 MeV energy. This population had been monitored by Bob Filz and colleagues at AFCRL (*Filz and Holeman*, 1965) using nuclear emulsions carried on low altitude reconnaissance satellites and recovered along with the surveillance payloads. Immediately after Starfish the proton flux at low altitude in the South Atlantic anomaly was observed to be increased by a factor of about 10. Most of us attributed this result to a redistribution of mirroring points of inner belt protons by the electromagnetic fields of the nuclear detonation, but no quantitative proof has been forthcoming. In 1991 a shock wave from interplanetary space passed through the magnetosphere causing a major disruption to the energetic trapped protons as well as to untrapped solar flare protons (*Hudson et al.*, 1995). With this recent example in mind it would be interesting to calculate whether the shock wave produced by a high altitude nuclear explosion could also scatter 55 MeV protons in the inner belt and cause the observed changes in the trapped distribution.

4. HIGH ALTITUDE DETONATIONS BY THE USSR

On October 22, 28, and November 1, 1962 the USSR conducted a series of high altitude nuclear weapons tests which injected large numbers of fission electrons into trapped orbits. As luck would have it, STARAD was launched on October 26, and Explorer 15 was launched on

October 28, about 5 hours before the Soviet detonation of that day. Consequently the USSR detonation of October 28 is probably the best documented case of all the artificial radiation belts produced by nuclear explosions. The detonations of October 22 and 28 injected electrons over a wide region above the nominal L-value of the detonation points, L≈1.8. The explosion on November 1 created a narrow band near L=1.76, presumably because the explosion was at a lower altitude and the debris did not expand perpendicular to the geomagnetic field lines. This narrow band was ideal for studying radial diffusion. Walt Brown had instruments on Explorer 15 and was able to measure the steady increase with time in the width of this band (*Brown*, 1966). These data convinced me that radial diffusion was an important process even in quiet geomagnetic conditions.

Reports of these belts, their decay rates, energy spectra and pitch angle distributions were prominent at space science meetings for a decade. (See for example papers by *Van Allen, McIlwain, Brown, and West* in *McCormac*, 1966). My contracts with DASA were quickly modified to include study of the data from the USSR tests. Of considerable interest to the DOD were the altitudes and the yields in these detonations as well as the phenomena associated with the fireball expansion and electron transport. Newkirk and I spent some tedious hours integrating flux distributions to find the total number of electrons injected by the explosions. Lower limits to these fission yields could be obtained directly from the initial electron inventory in the belts. With some additional assumptions about the detonation altitudes, yields, and the uncertain hydrodynamic processes occurring after detonation, one could estimate the injection efficiency and thereby derive a value of the fission yield of the weapon. The electron injection process was (and still is) quite uncertain, although estimates were made and refined over the next decade.

With the sophisticated magnetic spectrometer of STARAD several anomalous features were found in the trapped electron distributions of the Russian detonation of October 28. Harry West (1966) of LRL noted that the electron spectrum at higher latitudes had more low-energy electrons than expected from a fission spectrum. A similar surplus of low energy electrons at high L had been suspected for Starfish, but experimental confirmation in that case was weak. With West's results in hand theorists immediately suggested causes for the low energy enhancement. These postulated mechanisms included (1) betatron deceleration in the expanding magnetic bubble, (2) energy loss as electrons reflected from the sides of the expanding bubble, and (3) a flute instability which caused the magnetic tube containing the fission electrons to move to higher L shells while conserving the first two adiabatic invariants. It is not known if any of these speculations is the correct explanation.

There was an obvious military need to be able to predict the consequences of a high altitude detonation at any location and yield. Several mechanisms were postulated to explain the large number of electrons found at higher latitudes than the detonation points. These processes include, flute instabilities resulting in outward convection of flux tubes, charge exchange of fission debris with the neutral atmosphere thereby allowing the fission fragments to cross field lines, and jetting of the debris across field lines as the cloud moved through the weak magnetic field at the equator. In the decades after the tests these processes have been tested by particle simulation codes. Unfortunately, the adjustable, unknown factors in these processes allow the codes to fit observations in several ways, leaving our understanding of this phenomena in an unsatisfactory state.

Analysis of the Starfish and USSR tests of 1962 continued for over a decade. The work eventually culminated in a computer program to predict the radiation belts formed by a nuclear detonation at any location and yield. The code was based on first order physics of the debris expansion, but treated the trapped particle motion in detail. It included the drift of the electrons around the earth in a realistic geomagnetic field and the subsequent diffusion and loss of the electrons. The ephemeris of existing satellites could be included so that the radiation striking any satellite could be calculated as a function of time after detonation and the damage to the satellite estimated.

Although this weapon prediction work was strictly applicable to nuclear weapons detonations, the physical processes included in the calculation were well-recognized geophysical effects present in the natural belts. These processes include pitch angle scattering (strong diffusion in the case of very high radiation belt fluxes), L-shell diffusion, flute instabilities in the initial magnetic tube containing the debris, atmospheric scattering losses, and many other factors. Thus, the work in understanding and predicting weapon phenomena led to a more quantitative understanding of the source and loss mechanisms of the natural radiation belts.

EPILOGUE

In the decade from 1958 to 1968, the scientists with whom I was associated at Lockheed contributed in several ways to the understanding of trapped radiation in the earth's magnetosphere. It is of interest to note that all of this

Van Allen radiation belt *J. Geophys. Res., 67*, 5013, 1962.
Walt, M., L. F. Chase, J. B. Cladis, W. L. Imhof, and D. J. Knecht, Energy spectra and altitude dependence of electrons trapped in the earth's magnetic field, in *Space Research,* edited by Kallmann-Bijl, pp910, North Holland Publishing Company, Amsterdam, 1960.
Wentworth, R. C., W. M. MacDonald, and S. F. Singer, Lifetime of geomagnetically trapped particles determined by coulomb scattering, *Phys. Fluids 2*, 499, 1959.
West, H. I., Some observations of the trapped electrons produced by the Russian high-altitude nuclear detonation of October 28, 1962. in *Radiation trapped in the earth's magnetic field*, edited by B. M. McCormac, D. Reidel Publishing Co., Dordrecht, Holland, 1966

A Brief History of Research at Minnesota Related to the Magnetosphere - 1957-1970

John R. Winckler

School of Physics and Astronomy, University of Minnesota, Minneapolis, Minnesota

Research in cosmic ray physics leading up to the International Geophysical Year (IGY) is described. Discoveries during the IGY created interest in magnetic storms, aurora and Solar activity, and led to studies of the energetic particle population of the Magnetosphere using space vehicles. Of particular interest was the injection of radiation belt particles during substorms, most effectively observed at the geostationary orbit. The instrumentation and the roles played by various investigators are described.

1. INTRODUCTION

Our account begins in the early 1950's, when the research interests in what might be called "cosmic physics" at the University of Minnesota were concentrated in studies of the primary cosmic radiation. This continued the earlier work of Freier, Lofgren, Ney and Oppenheimer [*Freier et al., 1948*], who had participated in the discovery of heavy nuclei in the primaries. The essential experimental tool was the fixed volume, high altitude, constant level type plastic "Skyhook" balloon which could reach altitudes of 30 km ($<$ 10 gm cm^{-2} residual stopping power of the atmosphere above the balloon). Work on these balloons was begun at Minnesota earlier by Jean Piccard, Otto Winzen and others and they were fabricated initially by General Mills Corporation Instrument Division under Office of Naval Research (ONR) sponsorship. In the period 1952-55 an important technical program in the Physics Department of the University of Minnesota, sponsored by the joint military services, was directed at understanding and improving these high altitude plastic vehicles. The program was directed by Charles Critchfield, Edward Ney and John Winckler. Two important improvements in balloon technology produced by this program were the "natural shape" ballon configuration, in which the internal pressure of the balloon lifting gas was sustained entirely by the network of load-bearing meridional tapes, and the circumferential stresses were very low or zero, and the "duct" appendix, which permitted the balloon to valve its excess lifting gas upon reaching full volume at ceiling altitude, without the premature admixture of air, which frequently occurred in balloons using an open bottom appendix, and which often decreased the floating altitude.

2. RESEARCH AT MINNESOTA BEFORE THE INTERNATIONAL GEOPHYSICAL YEAR

The improvement in high altitude plastic balloons played an important role in continuing cosmic physics research, and investigations of the cosmic ray primaries using nuclear emulsions, cloud chambers, electronic detectors such as Cerenkov and scintillation counters, Geiger counters and other devices continued in parallel with the balloon development. Physics graduate students associated with this early effort included Frank MacDonald, John Naugle and Kinsey Anderson.

An interesting highlight of this period was the annual meetings of the "Midwest Cosmic Ray Colloquium" which met at some "Big Ten" universities such as Wisconsin, Michigan, Nebraska, Minnesota, and the University of Chicago, as well as Canadian participation from the

Discovery of the Magnetosphere
History of Geophysics Volume 7

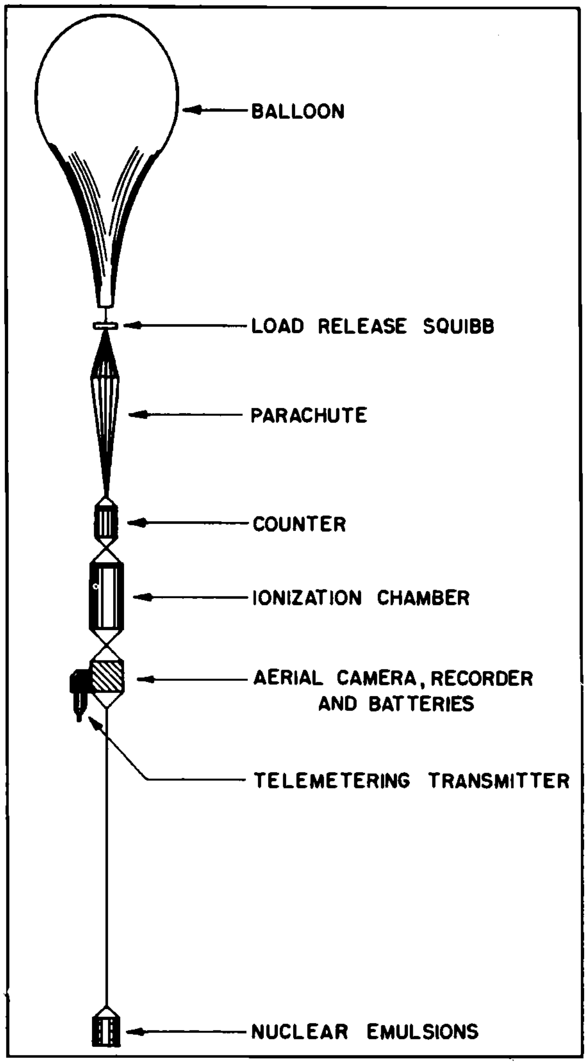


Figure 1. IGY flight train. After about 24 hours of flight, the load was released by firing the squibb electrically from the recorder package. The entire load train then dropped to the ground by parachute for recovery.

Canadian NRC. The meetings featured lectures by Enrico Fermi and others on cosmic ray origin theories and on high energy interactions, and by John Simpson, Hugh Carmichael, James Van Allen and many others on the subject of primary cosmic rays, Solar flares and Solar cosmic ray events, the modulation of primaries by Solar cycle effects and "Forbush decreases", and later the Earth's radiation belts. This period marked a transition when high energy particle interaction physics was rapidly being taken over by large accelerators, but the astro- and geo-physical aspects of cosmic rays continued to be a major part of these colloquia.

3. THE INTERNATIONAL GEOPHYSICAL YEAR DISCOVERY OF AURORAL X-RAYS

The advent of the International Geophysical Year (IGY) on July 1, 1957 brought a new emphasis to high-altitude research at Minnesota. Previously, the background of experience of the group was in high-energy particle physics, although the geomagnetic effects on cosmic rays was a familiar topic, and the "Stormer" theory of the entry of primary cosmic ray particles into the Earths magnetic field was well understood, including the latitude and "east-west" effects on the momentum spectrum of the incident primaries [*Winckler et al.*, 1950]. Ney and Winckler proposed a high-altitude program to "monitor" the primary cosmic rays during the period of sunspot maximum of the IGY (Solar Cycle No. 19). This program was supported by the ONR and is described in an IGY monograph [*Ney and Winckler*, 1958]. The program ultimately consisted of 85 high altitude balloon flights carrying a standardized payload train. (Figure 1) This train comprised a Neher-Millikan type integrating ionization chamber and a small Geiger counter to record individual particle counts. These data as well as the pressure altitude were telemetered to ground, and were also recorded on the film of a small 35 mm aerial camera which photographed the ground during daylight hours. The flight train also included a small vertically-oriented stack of nuclear emulsions encapsulated in a sealed can, for recording the tracks of primary cosmic rays. These emulsions were processed and examined by Phyllis Freier and Edward Ney (see 5. Solar Cosmic Rays, below). The aerial camera films were analyzed by Homer Mantis [*Mantis*, 1958] and yielded many examples of wind fields and turbulence structure at 30 km altitude. Graduate students involved in this program included Lawrence Peterson, Roger Arnoldy and Robert Hoffman (Figure 2).

Most of the IGY balloon flights were launched from an abandoned airfield near Minneapolis where the telemetry receiving station was located. The flights were terminated by a timer which separated the balloon from the flight train, which then descended on a parachute. Many of the flight trains were recovered and could be reused. Some even crossed the Atlantic in the high winter jet stream. In summer the flights drifted slowly westward over the Dakotas.

The first full-scale flight was launched on the first day of the IGY, 1 July 1957 Universal Time (the evening of June 30 in Minneapolis). Shortly after the flight reached floating altitude at about 30 km (8 gm cm^{-2} atmospheric depth) excited voices could be heard from inside the telemetry building where the flight records were viewed in real time. Substantial, irregular increases were observed in the ion

increases, and provided insight into Magnetosphere processes during magnetic storms.

Energetic x-ray bursts directly from the Sun in coincidence with Solar flares were also discovered and frequently detected by the IGY balloons [*Peterson and Winckler,* 1959]. These bursts created ionospheric effects, but had little direct interaction with the Magnetosphere. Large numbers of these bursts were later monitored in space on the OGO satellite series (see 7. The Sun and the Radiation Belts, below).

Under a continuation of the balloon program sponsored by the National Science Foundation (NSF) in 1959, a new light-weight radiation monitoring system using smaller balloons was devised, capable of being launched in an unrestricted way under Federal Aeronautics Administration (FAA) rules, even in overcast cloud conditions. With this "continuous monitoring" system more than 500 balloon flights were made in the post-IGY period during 1958-59, yielding data on auroral x-rays, Solar x-rays and Solar Protons, geomagnetic storm effects, and even radioactive layers in the atmosphere from Soviet nuclear tests [*Mantis and Winckler,* 1960].

6. EARLY SPACE EXPERIMENTS

The space age began with the electrifying news that the Soviet Union had launched the first "Sputnik" orbiting payload on October 4, 1957. The next day at Minnesota a group gathered around a radio receiver set up in one of the cosmic ray research labs in the Physics Department and heard the clear beeping signal from Sputnik 1 for about 10 minutes during a pass over the area. The ultimate launching of the US Explorer 1 on February 1, 1958, further alerted the group to the new possibilities ahead. But the announcement by James Van Allen at the 1958 joint spring meeting of the American Physical Society and National Academy of Sciences of the detection of intense belts of trapped protons in the inner Magnetosphere using Explorer 1 was the real start of the era of the Magnetosphere. Eventually the Minnesota group was offered the opportunity to provide a radiation detector for the NASA Pioneer 5 and then the Explorer 6 payloads. The instruments were miniaturized versions of the IGY balloon ion chambers. Although the devices operated as planned, the interpretation turned out to be difficult as the ion chambers were buried inside the payload. Particularly for Explorer 6 whose orbit passed repeatedly through the radiation belts an attempt was made to interpret the readings as x-rays produced by radiation belt electrons bombarding the outer skin of the instrument package. The electron fluxes thus obtained were extremely high, and were therefore suspect. In a conversation between Winckler and the Russian A. E. Chudakoff, it was revealed that the Soviet Sputnik 3 mission had detected very energetic trapped electrons, in the MeV energy range. This provided the explanation for the high electron fluxes using the x-ray interpretation. These high energy eletrons penetrated directly into the ion chamber and gave a high direct count rate, although their fluxes were low at the high energy tail of the electron spectrum. The ion chamber, however, continued to be useful for the detection of Solar x-rays which reached the Earth with the speed of light following a Solar flare. On the NASA OGO 1 and 3 missions, described in the next section, a Xenon-filled ion chamber was mounted on a boom outside the payload skin, and successfully recorded thousands of Solar x-ray bursts when the satellite was "outside" the radiation belts. [*Arnoldy et al.,* 1968]. Their detection now survives in the GOES "space weather" monitoring satellites at geostationary orbit.

7. THE SUN AND THE RADIATION BELTS

Following the early space experiments, it was decided that what was needed was a good electron spectrometer that could correctly measure the electron spectra throughout the radiation belts. Karl Pfitzer was assigned this project for his PhD thesis, and provided the instrument for the NASA OGO 1 and 3 missions (OGO=Orbiting Geophysical Observatory), whose orbits spent some time in interplanetary space, but cut through the "outer" and the "inner" belts of trapped electrons [*Pfitzer et al.,* 1966]. This spectrometer covered the energy range from 50 keV to 4 meV, and was well shielded to decrease the background counts in the detector due to high-energy x-rays from energetic electrons bombarding the exterior. Energy selection was by a magnetic deflection system. One major finding was that following a large Solar flare and great magnetic storm there was a large increase in the electron fluxes in the "outer" region. These electrons diffused inward over a period of several months, and at the same time increased their energy until finally the "inner" zone lying just above the atmosphere in equatorial regions also increased. Thus the origin of both the outer and inner regions could be linked to solar-terrestrial magnetic activity. The region between the belts, known as the "slot", was observed to fill during the inward diffusion, but afterwards to empty to a low level. This process was later linked to "whistler" wave activity in this region of space, which was claimed to be capable of scattering slot region electrons into the loss cone in the quasi-stable period between magnetic storms. The importance of these findings

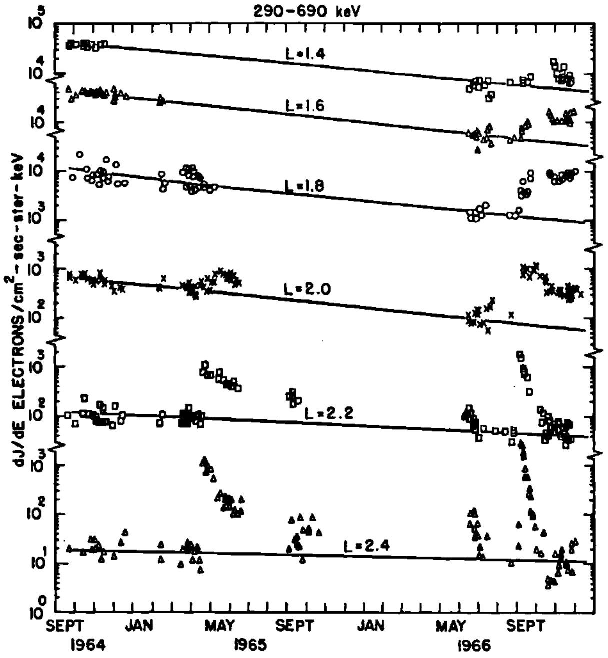


Figure 5. Long-term time history of the inner radiation belt for electrons of energy between 290 to 690 keV, for various McIlwain L-parameter values, and for particles moving with equatorial pitch angle of 90°. From the OGO 1 and OGO 3 missions. Note the major injections in March and September, following large Solar flare events. The September injection penetrated to the lowest "L" values. Note also the steady decay of the background flux of "Starfish" electrons shown by the solid lines at each "L" value.

was that prior to this work the inner zone regions were attributed to the decay of neutrons of cosmic ray origin, actually a source far too weak to account for the inner zone intensities. The injection of electrons associated with large Solar flares and magnetic storms is shown in Figure 5, from [*Pfitzer and Winckler,* 1968]. A theory concerning the acceleration of radiation belt particles by a violation of the "third invariant" of motion was described by [*Kellogg,* 1959].

It was a matter of great interest that during the history of the OGO 1 and OGO 3 missions an artificial belt of trapped electrons in the inner zone known to have been created by the "Starfish" high-altitude fusion bomb explosion was observed by the Pfitzer spectrometer to decay slowly. The natural lifetime of the inner zone electrons was thus determined to be 300-400 days. Note the baseline decay of the "Starfish" electrons in Figure 5.

8. GEOSTATIONARY ORBIT OBSERVATIONS

A further important step in the analysis of the radiation belts was provided by the ATS 1 (Applications Technology Satellite 1) mission, with the satellite injected into the geostationary orbit at 6.6 RE (Earth Radii) from Earth center on December 6, 1966. The ATS 1 mission was successful, and was happily located at 165° west longitude, at the intersection of the geographic and magnetic equators. The Minnesota group provided a small but effective electron spectrometer instrumented by Thomas Lezniak as his PhD thesis (summarized in [*Lezniak et al.*, 1968]). The great advantage of the ATS 1 orbit configuration was its fixed position above a point on the Earth's surface, so that with the 24 hour rotation of the Earth it moved slowly through the outer Magnetosphere. Something approaching the true time variation of the electron spectra and the magnetic environment could thus be sampled from point to point, and the disadvantage of the usual satellite which moved at high speed through the belts removed. The first major discovery was to show that the injection and acceleration of electrons into the outer zone occurred during an event called a "Magnetospheric Substorm." An example of this process during a substorm is shown in Figure 6, from [Parks and Winckler, 1968]. The "substorm" concept, originated by Sidney Chapman and Syun-Ichi Akasofu, Kinsey Anderson, and others [*Akasofu,* 1968], was brought to Minnesota by George Parks in an extended post-doctoral visit during 1965-68. These substorms typically involved a sudden brightening of the aurora on the night side of the Earth, a negative excursion of the nighttime horizontal component of the Earth's magnetic field accomapanying an intensification of the auroral electojet, a sudden collapse of the extended nighttime Magnetosphere tail-like magnetic field configuration to a more dipole-like shape, and other features. A number of these substorms accompanied the "main phase" of a magnetic storm when the equatorial ring current referred to earlier had developed. In fact, it was recognized that each sub-storm made a contribution to the main-phase ring by the injection of energetic protons which drifted westward around the Earth (see diagram page 6 of [*Akasofu,* 1968]). At the same time energetic electrons were injected which drifted eastward. In one study involving both the OGO and ATS1 satellites these bunches of drifting electrons were timed and identified at two positions in the Magnetosphere [*Pfitzer and Winckler,* 1969]. The ATS1 mission was also very useful in following the long-term variations in the outer zone. For example in 1965-66, during a Solar minimum period when geomagnetic storms were weak and infrequent, the outer

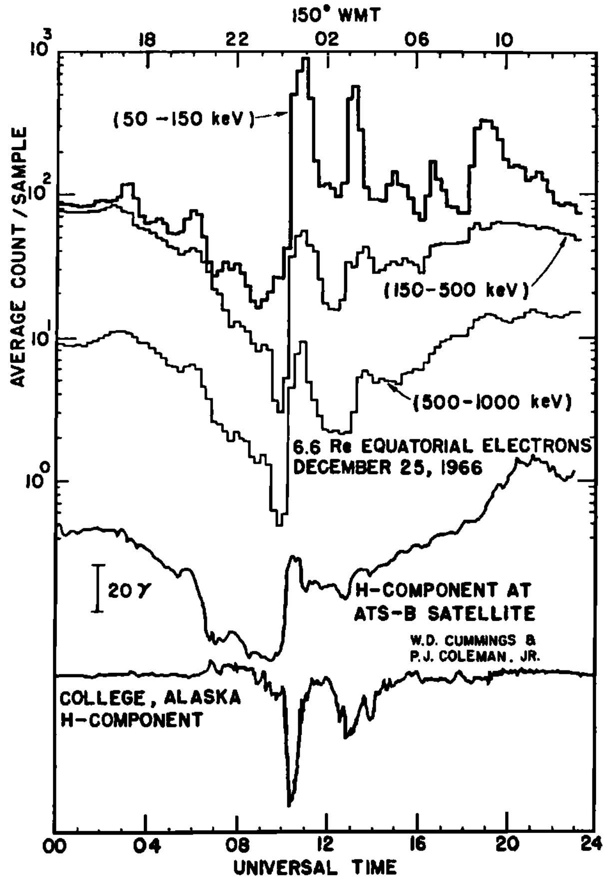


Figure 6. The time dependence of 50-1000 keV trapped electrons at 6.6 R(e) in the magnetic equatorial plane for a complete 24 - hour orbit of the ATS-1 (or ATS-B) satellite during an active magnetic period, The most crucial feature shown is the 10-100-fold increase in electron flux in the 50-150 keV energy interval which occurred at the time of the magnetic "bay" onset at 09 UT (i.e., a substorm onset) showing the acceleration ond injection of new trapped electrons into the outer radiation belt region. Further increases corresponded to continuing substorm activity. Note also the recovery of the magnetic field H-component at 6.6 R(e), as the night sector Magtnetospheric magnetic field changed from more "tail-like" to more "dipole-like," also at substorm onset.

zone steadily decayed to a very low level, showing the Solar control over this feature of the Magnetosphere (J. R. Winckler, private communication).

One key observation gave direct evidence about the auroral x-rays discovered during the IGY. Data from the ATS 1 elecron spectrometer was time correlated with auroral x-rays measured by George Parks in a balloon flight near the ground intercept of magnetic field lines connecting the geostationary orbit with the ionosphere. During an auroral substorm the balloon x-ray fluxes and the electron intensity at geostationary orbit both increased together, with very similar time profiles, as shown in Figure 7 [*Parks and Winckler,* 1968]. Now it was clear that the electrons responsible for auroral x-rays were accelerated and injected during auroral sub-storms. Finally, the mystery of the origin of the auroral x-rays was largely understood.

The ATS 1 mission, because of its unique rotaional scan over spatial directions, was able to show a layered structure of the Magnetopause (in terms of 50-150 keV electrons). Disjoint regions, separate from the closed boundary, were clearly evident during "Magnetopause Crossings" when, due to a transient very high velocity and energy density of the Solar wind, the Magnetopause moved from its normal position near 10 Earth radii to inside the 6.6 Earth radii distance of the Geostationary orbit. [*Lezniak and Winckler,* 1968].

9. OTHER MAGNETOSPHERIC CONCEPTS

The decade 1957-1967 saw a rapid growth in theoretical and experimental knowledge of the Magnetosphere. As an extension of the Chapman-Ferraro concept of the Magnetopause boundary, and since the Solar wind stream is supersonic (in the sense that the flow velocity exceeds the Alfven wave speed in the medium) one would expect that a shock wave would form analagous to the shock wave before a supersonic projectile in air. Paul Kellogg at Minnesota was one of the first to describe this in the literature [*Kellogg,* 1962]. This standing "bow shock" was identified experimentally on many space missions, but the

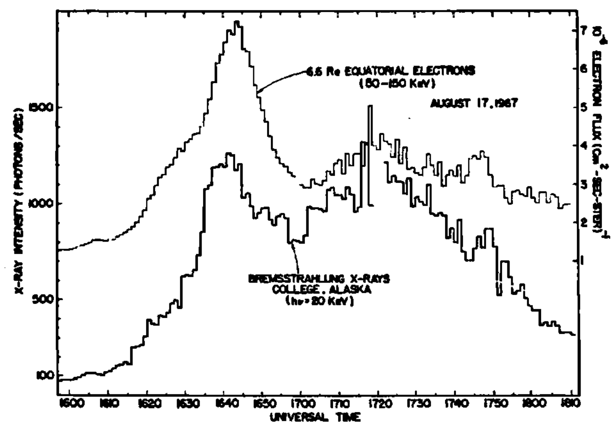


Figure 7. Showing the close correlation of the 50-150 keV trapped electron increase at 6.6 R(e) with the X-ray flux, due to precipitating electrons on the same lines of force in the auroral zone, during a sub-storm. This explains the origin of the Auroral X-rays discovered during the IGY.

Pfitzer, K. A., S. R. Kane, and J. R. Winckler, The Spectra and Intensity of Electrons in the Radiation Belts, *Space Research VI,* 702-713, 1966.

Pfitzer, K. A. and J. R. Winckler, Experimental Observation of a Large Addition to the Electron Inner Radiation Belt after a Solar Flare Event, *J. Geophys. Res. 73,* 5792-5797, 1968.

Pfitzer, K. A. and J. R. Winckler, Intensity Correlations and Substorm Drift Effects in the Outer Radiation Belt Measured with the OGO 3 and ATS 1 Satellites, *J. Geophys.Res., 74,* 5005-5018, 1969.

Sison, J. A., Cosmic Radiation Neutron Intensity Monitor, Ann. Intern. Geophys. Yr. 4, 351, 1957.

Van Allen, J. A., G. H. Ludwig, E. C. Ray, and C. E. McIlwain, Some Preliminary Reports of Experiments in Satellites 1958 Alpha and Gamma, *Trans. Am. Geophys. Union 39,* 767, 1958. Also, *National Academy of Sciences, Washington, D. C., IGY Satellite Rept. Sec. 3,* 73-92, 1958.

Winckler, J. R., and L. Peterson, Large Auroral Effect on Cosmic-Ray Detectors Observed at 8 g/cm^2 Atmospheric Depth, *Phys. Rev. 108,* 903-904, 1957.

Winckler, J. R., T. Stix, K. Dwight, and R. Sabin, A Directional and Latitude Survey of Cosmic Rays at High Altitude, *Phys. Rev. 79,* 656-669, 1950.

Winckler, J. R., P. D. Bhavsar, and L. Peterson, The Time Variations of Solar Cosmic Rays during July 1959 at Minneapolis, *J. Geophys. Res., 66,* 995-1022, 1961.

Winckler, J. R. and P. D. Bahvsar, Low Energy Solar Cosmic Rays and the Geomagnetic Storm of May 12, 1959 *J. Geophys. Res., 65,* 2637-2655, 1960.

John R. Winckler, 2012 Irving Ave. S., Minneapolis MN 55405. E-Mail, winck001@gold.tc.umn.edu

Present Knowledge of the Magnetosphere and Outstanding Remaining Problems

D. N. Baker

Laboratory for Atmospheric and Space Physics, University of Colorado

The Earth's environs remain the most readily accessible prototype for understanding remote planetary and astrophysical magnetospheric systems. The near-Earth region has proven to be a plasma physics laboratory rich in phenomena of cosmic significance: magnetic reconnection, particle acceleration, wave-particle interactions, and electrical currents of huge spatial scale are all seen to occur. Even after four decades, we continue to learn new things about the physical processes occurring within the magnetospheric system. This paper summarizes briefly the structure and dynamics of the magnetosphere as it is presently understood. It also discusses some of the key outstanding problems that remain to be solved.

1. INTRODUCTION

The Earth's magnetosphere has provided fertile ground for exploration since its discovery nearly 40 years ago. Following the first pioneering of the terrestrial plasma physical domain, today more general perspectives have come to apply. A general definition is that a *magnetosphere* is a relatively self-contained region in space whose global topology is organized by the magnetic field associated with the parent (compact) object. A number of the planets (Mercury, Earth, Jupiter, Saturn, Uranus, and Neptune) are known to have intrinsic magnetic fields and have magnetospheric regions around them. These differ from one another in very significant, and interesting, ways. Indeed, the Sun itself may be viewed as having a magnetosphere (the "heliosphere") within which all of the planetary magnetospheres are embedded. It is the continual outflow of hot coronal gas from the Sun which gives rise to the heliosphere and this supersonic gas flow is called the solar wind. The solar wind compresses, distorts, and confines the planetary magnetospheres and imparts to them much (if not most) of the energy that is dissipated and/or radiated away.

In addition to planetary and solar-system scale magnetospheres, there also appear to be magnetospheres of galactic proportions. In each case – and on all scales – there are analogous features as well as distinctly different characteristics. By far the most thoroughly explored magnetosphere is that of the Earth and it is this physical system which forms the basis for our best understanding.

The intrinsic magnetic field of the Earth arises from a complex dynamo action in the molten outer core of the Earth and may be well-represented for many purposes by an Earth-centered dipole. This field extends far into space and serves to deflect the on-rushing solar wind plasma. The stand-off distance at the subsolar point is highly variable (depending on solar wind pressure), but is commonly about 10 R_E (1 R_E = 1 Earth radius = 6375 km). The flowing solar wind applies stresses to the outer reaches of the Earth's intrinsic field and sets up a system of currents in the boundary regions. The forces due to these currents act to distort the magnetic field and field lines are dragged downstream to form a very elongated magnetotail.

The solar wind flows continually over, around, and into the Earth's magnetosphere and in so doing it continually imparts mass, momentum, and energy to the system. This transfer, however, occurs with great variability. When the change in energy is high, the magnetosphere moves far out of its equilibrium "ground state". The energy change then implies dissipation, either continuously or sporadically. It has been found, in fact, that the dissipation of solar wind energy imparted to the magnetosphere occurs in a quite sporadic way and the sudden occurrence of this magnetospheric dissipation is a major feature of the collection of physical processes that are called "geomagnetic activity".

Discovery of the Magnetosphere
History of Geophysics Volume 7

In this paper, we will discuss a few selected topics in magnetospheric structure and dynamics. The general approach will be to take cognizance of the earliest view of a particular magnetospheric feature at the time of its discovery. This will be followed by a presentation of the "modern" (present-day) view of this magnetospheric topic. In most cases we will also try to assess where a particular research area is heading in the future. Given, of course, the vast complexity of the magnetosphere, only a limited number of topics can be treated: We restrict ourselves here to the general themes addressed earlier in this monograph by the pioneers of magnetospheric research.

2. RADIATION BELT STRUCTURE

Within the magnetospheric cavity there exists a limited region where the motion of energetic particles is confined by the Earth's magnetic field. This region comprises the Earth's radiation belts. The radiation belts contain electrons, protons, helium, carbon, oxygen, and other ions with energies from less than 1 keV to hundreds of MeV. Particles with energies below about 200 keV provide the principal energy density and form the extraterrestrial ring current. Confinement (or trapping) of these particles results from the dipolar-like topology of the geomagnetic field which is characterized by magnetic field lines that converge at high latitudes toward the poles resulting in a relative minimum magnetic field strength region in the vicinity of the geomagnetic equator.

The discovery of the Earth's radiation belts by Van Allen and co-workers was the first major discovery of the space age. Figure 1a shows a diagram of the radiation belts taken from the work of *Van Allen et al.* [1959]. In this earliest work, it was deduced that a band or torus of energetic charged radiation was present, but the exact composition and spectral distribution was rather unclear. It would await more sophisticated and extensive observations to flesh out the view of radiation belt structure.

A more modern characterization of the radiation belts is presented in Figure 1b. It shows that the belts consist dominantly of electrons (outer zone) and protons (inner zone) with energies of hundreds to thousands of keV. In the past few years, new measurements from EXOS-C, the Solar, Anomalous, and Magnetospheric Particle Explorer (SAMPEX), and the Combined Release and Radiation Effects Satellite (CRRES) mission have provided substantial new insights into the radiation belt structure [see *Baker*, 1995, and refs. therein]. As an example, EXOS experimenters have presented long-term average particle fluxes as a function of geographic latitude, geographic longitude, and altitude. These maps, along with other analyses based on the same data set, show that there are significant changes compared with earlier models of the inner zone. The CRRES mission was dedicated in substantial part to mapping the structure and dynamics of the radiation belts using modern detection instruments. The mission showed the basic structure of the outer zone electron belt, but also revealed the highly dynamic nature of this population.

In the realm of high-energy ion measurements, SAMPEX has recently located a new radiation belt surrounding the Earth. This belt traps material from the nearby interstellar medium. The newly discovered belt, which dips closest to Earth over the Southern Atlantic, is embedded in the lower of the two previously-known Van Allen belts (see black bands in Figure 1b). SAMPEX pinpointed the new belt, the existence of which was first predicted nearly two decades ago, and is now measuring its composition and monitoring its intensity variation.

The new belt consists of trapped heavy ions, including nitrogen, oxygen, and neon which are part of the "anomalous" cosmic ray component. Components of the interstellar gas that are electrically neutral can penetrate the heliosphere. Some of these neutral interstellar atoms are singly ionized by solar UV radiation, and are then accelerated to cosmic ray energies at the solar wind "termination shock" at the fringes of our solar system. Such singly-charged cosmic rays striking the Earth's atmosphere may lose the remaining electrons and become trapped in the Earth's magnetic field. Once inside the new belt, these ions may be trapped for weeks before leaking out into space or into the atmosphere. As a result, the amount of matter inside the new belt increases and decreases dramatically [see *Baker*, 1995 and references therein].

3. RADIATION BELT DYNAMICS

It is of great interest and importance in magnetospheric physics to understand basic particle acceleration and loss mechanisms and to determine the variation time scales for high-energy particles throughout the Earth's radiation belts. It has been noted by several authors that the relativistic electrons may provide a significant coupling mechanism between the magnetosphere and the Earth's middle atmosphere [*Baker et al.*, 1987; *Callis et al.* 1991]. Relativistic electrons in the Earth's magnetosphere are also of considerable practical importance because of their deleterious effect on spacecraft subsystems.

An early view by some researchers (fostered by diagrams such as Figure 1a and 1b), was that the radiation belts were rather fixed and steady in their characteristics. On the other hand, even in the early days, many scientists

the wave-particle interaction related to these relativistic electrons [*Imhof et al.*, 1991].

The precipitation of energetic electrons (energies of a few hundred keV and above) was one of the first areas of exploration in the Earth's magnetosphere [*Forbush et al.*, 1962]. Studies continued, governed as always by flight opportunities, until the present day [*Blake et al.*, 1996]. This research has shown energetic electron precipitation to be continuous at some level, highly complex and dynamic, and the major sink for magnetospheric electrons.

The great temporal variability in the observed electron fluxes, caused both by actual flux changes and by the rapid motion of a low-altitude satellite through differing magnetospheric regions, has been one of the major difficulties in gaining an understanding of the underlying physics. One useful approach in dealing with this problem is to increase the sampling rate of the spacecraft instrumentation. However, this leads to statistical problems unless there is a corresponding increase in the geometric factor of the sensor. The SAMPEX mission contains instrumentation that combines high-rate sampling of the data with a very large geometric factor, giving a new opportunity to study the precipitation of energetic electrons. The HILT (Heavy Ion Large Telescope) is providing exciting new data on the magnetospheric locations and temporal fluctuations of the electron precipitation. Local time and magnetic activity dependencies have also been found.

Enhanced electron fluctuations can be identified mainly at the high latitude portion of the outer radiation belt. During such crossings, bands of precipitation with a time scale of 10~30s are prominent near the high latitude edge of the precipitation region. The flux level of these "precipitation bands" often exceeds that of the main part of the radiation belt.

Another precipitation type prominent in the SAMPEX data is the shorter time scale bursts down to a few hundred milliseconds. These microbursts are usually distributed in the high latitude portion of the outer radiation belt. Unlike the precipitation bands, microbursts die out during the interval between orbits. Microbursts in both hemispheres are frequently observed. Many microburst sequences show a sharp increase in flux followed by fluctuations with a more slowly decaying overall amplitude.

Thus, broad areas of strong precipitation, extending ~2-3° in latitude, frequently are observed near the high-latitude boundary of the outer zone. These features can persist for hours and are seen in conjugate locations. The transient form of strong precipitation, microbursts, often are seen lasting for less than a second, indicating that microbursts sometimes occur in a very localized region; the narrow temporal structure is a consequence of the spacecraft orbital velocity. In other cases, where the spatial size is greater, the temporal evolution of the microburst can be followed. These observations clearly indicate that outer-zone electron precipitation frequently results from a strong scattering process, and not by weak diffusion of stably trapped electrons into the drift loss cone.

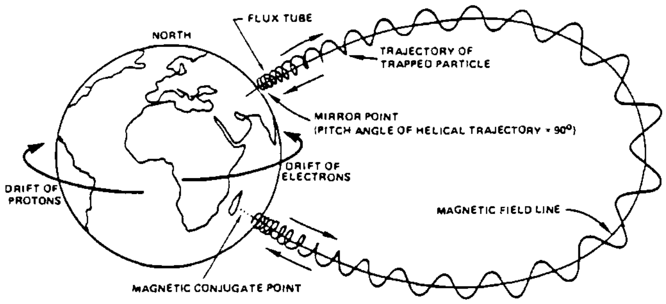
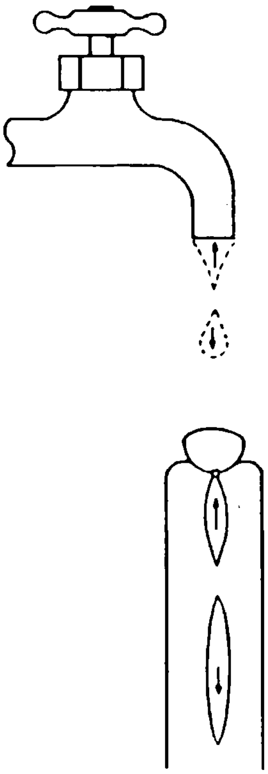


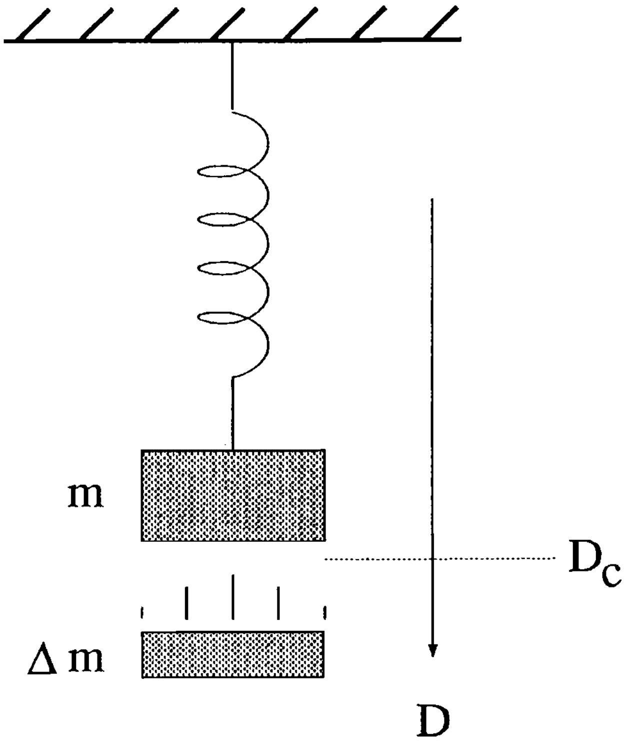
Figure 3. A schematic diagram showing energetic particle motion in the dipolar region of Earth's magnetosphere.

The modern era has ushered in remarkable new detection capabilities for studying wave-particle interactions and their manifestations near the atmospheric loss core. As noted above, the large area and high time resolution capability of SAMPEX has provided several new insights. An even more recent innovation is the Source/loss-cone Energetic Particle Spectrometer (SEPS) aboard the POLAR spacecraft. The SEPS consists of two separate telescopes that measure electrons and ions in the vicinity of the magnetic field direction. Given the front/back viewing apertures, the SEPS sensors can monitor both the atmospheric loss cone portion of the particle distribution function as well as the oppositely directed "source" cone region [*Blake et al.*, 1995]. Figure 4 shows an example from the SEPS investigation for Day 103 (April 12) of 1996. The deep loss-cone structure along the magnetic field line (in this case shown for alpha particles) is quite striking. SEPS sets the stage – along with measurements as from SAMPEX – for a new era of sensitivity, angular, and temporal resolution in studying wave-particle effects and pitch-angle scattering.

5. SOLAR WIND COUPLING AND MAGNETOSPHERIC CONVECTION

It had been recognized for many decades prior to the dawn of the Space Age that substantial energy must be supplied in some way to account for the dissipation which occurs in the auroral oval and in the formation and decay of the ring current. Other authors [*Akasofu*, this volume; *Dungey*, this volume] have recounted various aspects of this problem. It was recognized rather early that the solar





$h(s)$ is often called the lag-time and it expresses the response of the output at time t to the summed effects of the input at all times prior to t. Generally, *Bargatze et al.* found a peak in the filters at lag-time ~20 min showing that there is frequently a response in electrojet activity to solar wind activity some 20 min earlier; *Bargatze et al.* attributed this peak to the directly-driven magnetospheric response. There was a second peak at lag-time ~1 hr that was most evident for the moderate activity filters; this peak was attributed to the unloading magnetospheric response.

The *Bargatze et al.* [1985] LPFs showed that it is primarily the loading-unloading cycle in the magnetosphere that has a nonlinear response to the solar wind. *Hones* [1979] had earlier drawn an analogy between "plasmoid" formation and the dripping of a faucet. This analogy is illustrated here in Figure 6a. The basic idea is that a piece of the plasma sheet is pinched off during geomagnetic activity to form a plasma structure, i.e., the plasmoid. *Baker et al.* [1990] adapted a dripping faucet analogue model devised by *Shaw* [1984] and began a study of the nonlinear magnetospheric response using the methods of modern nonlinear dynamics. Shaw's model is illustrated in Figure 6b; it consists of a variable mass hanging on a spring. The displacement downward from the unstressed spring position is measured by the variable D. To model drop formation the mass M is increased with time at a constant rate

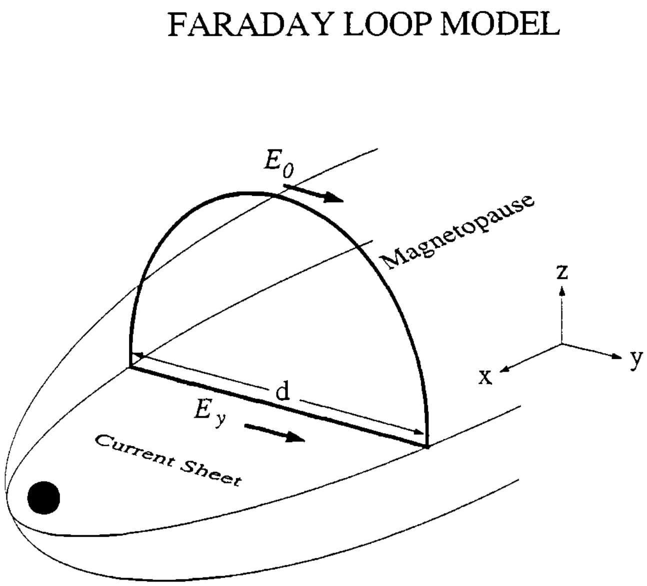


Figure 6. (a) A dripping faucet analogy with plasmoid formation in Earth's magnetotail [from *Hones*, 1979]. (b) The mechanical analogue of plasmoid formation [*Shaw*, 1984; *Baker et al.*, 1990]. (c) The Faraday Loop analogue model of substorm dynamics [*Klimas et al.*, 1992].

M_L until D reaches a critical displacement D_c. Then a piece of the mass ΔM is dropped so that M decreases discontinuously.

Baker et al. [1990] modified the dripping faucet analogue by adding friction and by changing the mode of mass loss. The friction was added to model dissipative processes in the magnetosphere and the mass loss was made continuous rather than instantaneous in order to model magnetotail field line merging during plasmoid formation.

It was demonstrated that this simple analogue model reproduces many of the dynamical features of a carefully controlled dripping faucet. The dripping faucet, or course, involves a water drop containing, effectively, an infinite number of degrees of freedom. *Shaw* [1984] showed that essentially only 3 of those degrees of freedom are involved in the dripping process. The water drop is a high-dimensional physical system whose dynamics is organized or low-dimensional.

Following the very simple analogue model of *Baker et al.* [1990], the Faraday loop model or FLM, [*Klimas et al.* 1992] is a time-dependent representation of magnetotail convection with a superposed loading-unloading cycle. The net flux content of a tail lobe is monitored in terms of flux accretion through the magnetopause and flux loss through Earthward convection in the plasma sheet and/or plasmoid release (see Figure 6c). It is assumed, as a basic part of the model, that if the loading rate through the magnetopause is sufficiently high, then Earthward convection loss cannot balance energy gain and the growth phase of a substorm results as the net flux content of the lobe grows. It is further assumed that if this imbalance persists, then at some point a critical point in the tail is reached when the flux content becomes too large and an unloading event occurs. The existence of this critical point and the consequent unloading is imposed on the model in a manner similar to the release of a portion of the mass in the dripping faucet model [*Baker et al.*, 1990].

A Faraday loop which encircles one of the tail lobes (Figure 6c) is used to relate changes in the magnetic flux in the lobe to electric potentials around the loop, which are then expressed in terms of solar wind parameters and cross-tail currents. The dynamic variables of the model are the cross-tail electric field measured in the current sheet, the flux content of a lobe, and a quantity that depends on the flaring angle and the diameter of the tail. The plasma sheet convection evolves within a magnetotail shape that varies in response to the dynamical evolution of the convection. To relate the FLM output to measured electrojet index data, an elementary mapping of the cross-tail electric field to the westward electrojet strength has been added to the model. From the first simple beginnings of the *Dungey* [1961] model, the modern analogue models are able to replicate geomagnetic time series remarkably well [*Klimas et al.*, 1996].

6. STORMS AND SUBSTORMS

A magnetospheric substorm, in an idealized sense, results when the interplanetary magnetic field (IMF) turns southward and thereby enhances dayside magnetic reconnection. This increases the transfer of energy from the solar wind to the magnetosphere. Part of the increased energy imparted to the magnetosphere is lost quickly in the form of current flow and Joule dissipation in the polar ionosphere. Most of the energy imparted to the magnetosphere, however, is added to the magnetotail where the magnetic flux may often increase by ≥30% during a substorm growth phase. The substorm expansion onset represents the explosive release of accumulated tail energy much like a solar flare represents sudden release of magnetic energy on the sun's surface. After 30-60 minutes of explosive auroral activity, the system begins recovery back to relative quiescence.

Akasofu [1964] was the first to develop the idea of the auroral substorm. His phenomenological sequence – a classic synthesis of extremely complex observational data – is shown here as Figure 7. When the expansion and recovery phases that Akasofu discussed are combined with the growth phase first described by *McPherron* [1970], the substorm sequence is complete. *Akasofu* [this volume] describes the historical aspects of substorm research.

In a broader sense, since the pioneering initial work on substorms, exploration of the Earth's space environment has revealed a dynamic and complex system of interacting plasmas, magnetic fields, and electrical currents. This domain formed by solar wind plasma impacting the magnetized Earth is called "geospace". The near-Earth space environment has, over the years, been explored and studied as a system of independent component parts – the interplanetary region, the magnetosphere, the ionosphere, and the upper atmosphere. From early explorations, it was known that geospace is a complex system composed of highly interactive parts. While previous programs advanced the understanding of these geospace components individually, an understanding of geospace as a whole has clearly required a planned program of simultaneous space and groundbased observations and theoretical studies keyed to a global assessment of the production, transfer, storage, and dissipation of energy throughout the solar-terrestrial system. Prior understanding of the various geospace components plus the availability of advanced instrumentation are now poised to allow, for the first time, the definition and planning of a

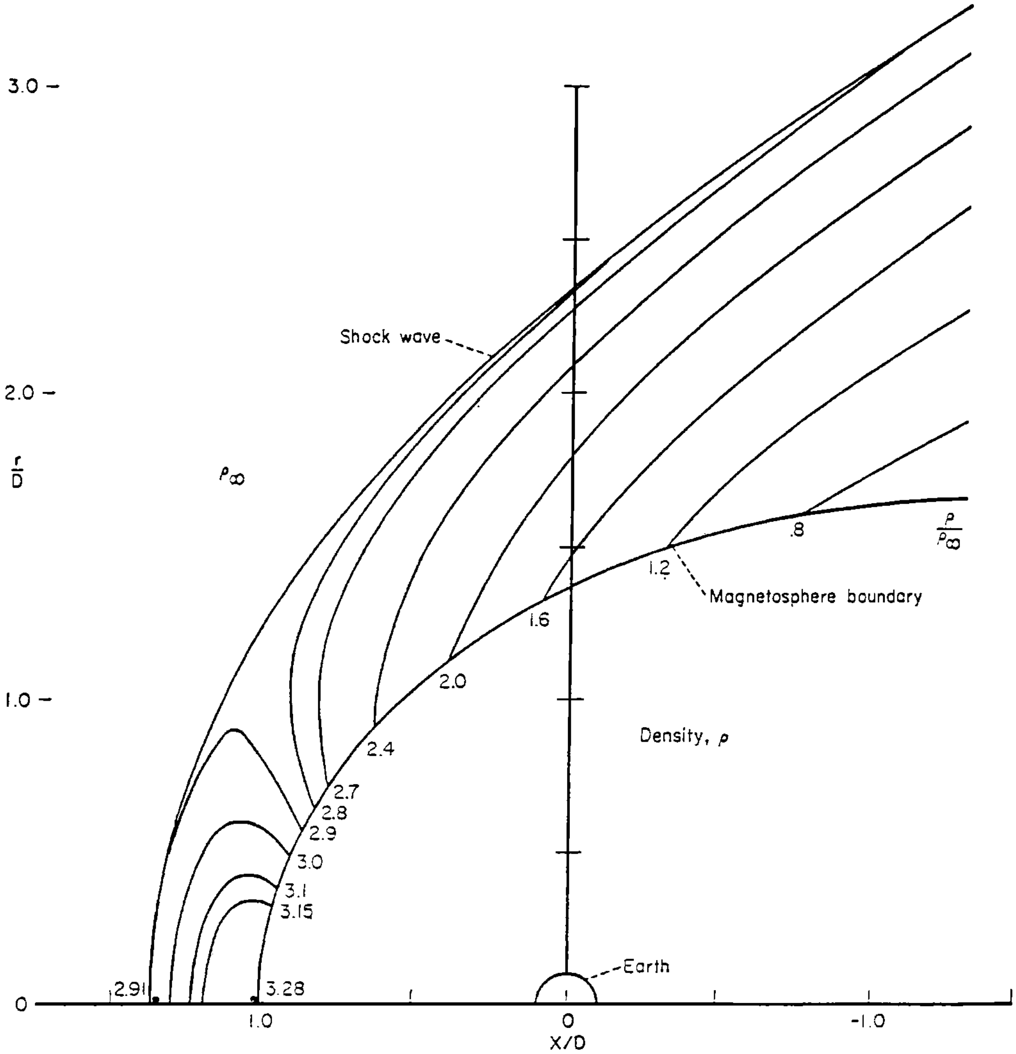


Figure 7. The sequence of auroral development during a magnetospheric substorm [*Akasofu*, 1964].

comprehensive, quantitative study of the solar-terrestrial energy chain.

Thus, the Global Geospace Science (GGS) component within the International Solar Terrestrial Physics (ISTP) program is reaching full stride. As shown, for example, in Figure 8, there are new global auroral imagers on the POLAR spacecraft that have the potential to revolutionize our understanding of magnetospheric substorms through frequent, high-resolution pictures of the aurora. ISTP represents a major advance in space plasma physics. It utilizes a large team of experimenters and modelers who should dramatically push back the frontiers of substorm and storm studies.

7. GLOBAL MODELING

As discussed previously in this volume [see *Spreiter*, this volume], the first use of large-scale numerical modeling codes provided a breakthrough in the understanding of magnetospheric properties. With such codes, it was possible to get something approaching a self-consistent treatment of the magnetosphere's shape and size as well as the collisionless shock wave standing upstream in the solar wind flow. Figure 9 is an illustration from *Spreiter et al.* [1966] showing an MHD simulation of plasma flow around the magnetosphere.

The power and speed of computers have increased dra-

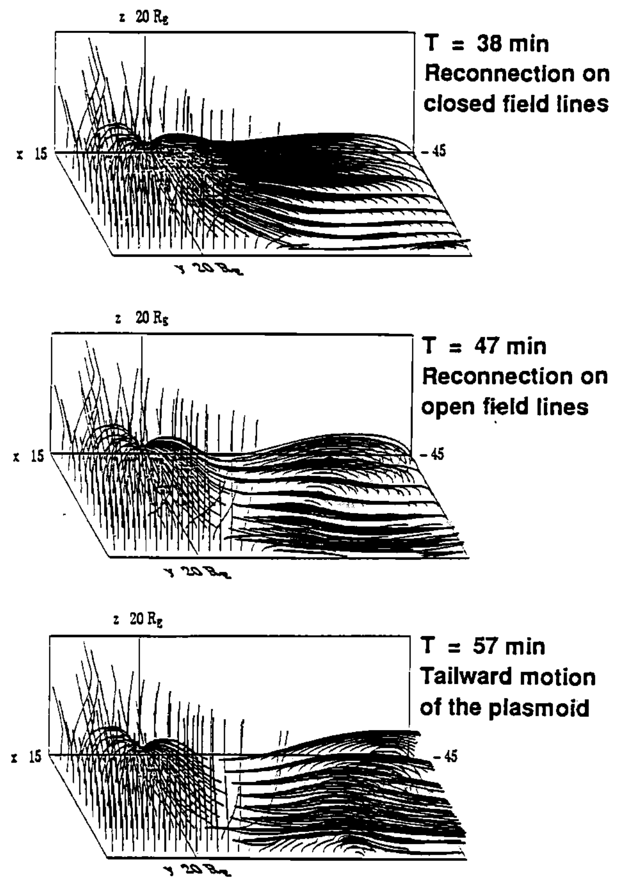


Figure 8. A modern global auroral image taken with the Visible Imaging System (VIS) onboard the POLAR spacecraft. The image, taken in UV wavelengths, shows an auroral intensification indicative of substorm onset at 0713 UT on 15 May 1996 [courtesy of L. A. Frank].

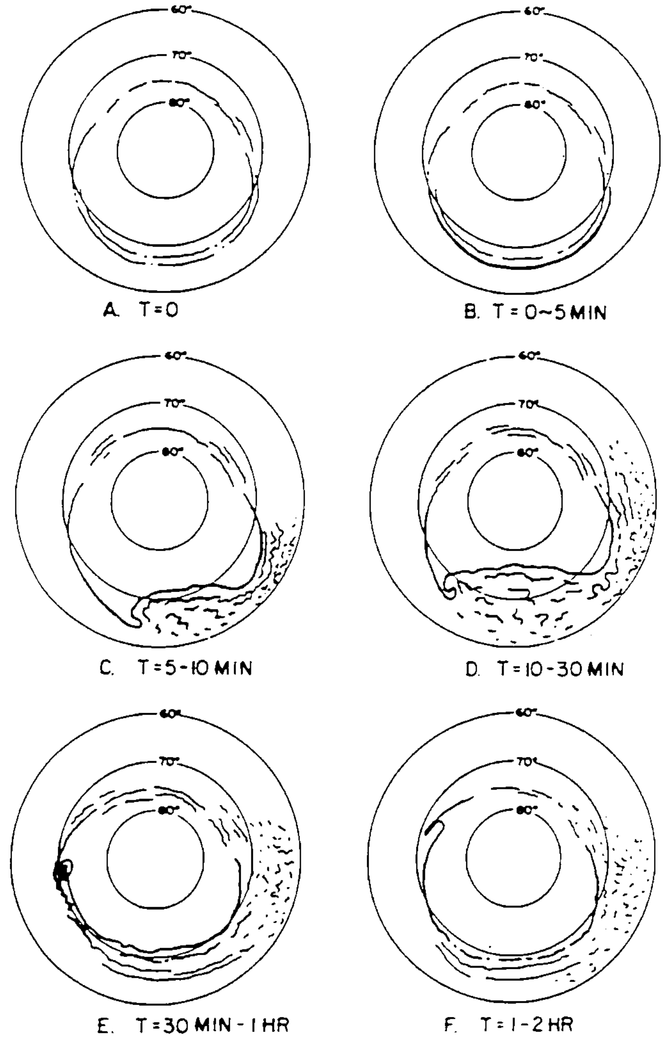


Figure 9. An early gas-dynamic simulation model of solar wind-magnetosphere interaction [*Spreiter et al.*, 1966].

matically since the 1960s. Today it is possible to simulate the 3-dimensional magnetosphere-solar wind interaction with considerable faithfulness. Global and large-scale fluid simulations of magnetospheric dynamics and of the solar wind-magnetosphere interactions consistently show that the response of the magnetotail to external energy input (southward IMF) is the formation and tailward ejection of an isolated plasma structure, a plasmoid [e.g., *Walker et al.*, 1993]. This has been true for a wide variety of simulation codes and external parameters used in the simulation runs. This is very consistent with modern observational results.

Enhanced energy input is seen in the MHD simulations as an increase in the near-Earth cross-tail current, in agreement with direct observations. The details of the numerical models affect the stability of the tail, but eventually in all simulations the tail is driven to an unstable state, and global reconfiguration of the magnetic field pattern follows. For example, in the simulation by *Walker et al.* [1993], even during a period that would be characterized as the late growth phase, a quasi-stable magnetic neutral line forms in the near-Earth tail (Figure 10). During this period, slow reconnection begins on closed field lines, but this does not cause a large-scale field reconfiguration. When reconnection proceeds to open field lines, a plasmoid is severed and begins to move away from the Earth. As discussed in prior sections, formation of a neutral line during the late growth phase had been earlier suggested by several observers based on observational results obtained in the near-Earth magnetotail.

Klimas, A. J., D. Vassiliadis, D. N. Baker, and D. A. Roberts, Organized nonlinear dynamics of the magnetosphere, *J. Geophys. Res., 101*, 13,089-13,114, 1996.

McPherron, R. L., Growth phase of magnetospheric substorms, *J. Geophys. Res., 75*, 5592-5599, 1970.

Petschek, H. E., Magnetic field annihilation, in *AAS-NASA Symp. Phys. of Solar Flares*, W. N. Hess, ed., NASA SP-50, p. 425, 1964.

Shaw, R., *The dripping facet as a model chaotic system*, The Science Frontier Express Series, Aerial Press, Santa Cruz, CA, 1984.

Spreiter, J. R., Historical development of magnetospheric modeling, this volume.

Spreiter, J. R., A. L. Summers, and A. Y. Alksne, Hydromagnetic flow around the magnetosphere, *Planet. Space Sci., 14*, 223-253, 1966.

Van Allen, J. A., High-energy particles in the Earth's magnetic field, this volume.

Van Allen, J. A., C. E. McIlwain, and G. H. Ludwig, Radiation observations with Satellite 1958 ε, *J. Geophys. Res., 64*, 271, 1959.

Walker, R. J., T. Ogino, J. Raeder, and M. Ashour-Abdalla, A global magnetohydrodynamic simulation of the magnetosphere when the interplanetary magnetic field is southward: The onset of magnetotail reconnection, *J. Geophys. Res., 98*, 17235-17250, 1993.

D. N. Baker, Laboratory for Atmospheric and Space Physics, Campus Box 590, University of Colorado, Boulder, CO 80303-0590