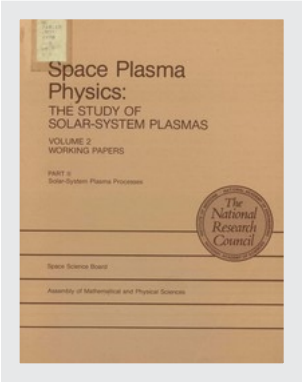


Space Plasma Physics--The Study of Solar-System Plasmas: Volume 2, Working Papers, Part 2, Solar-System Plasma Processes

DETAILS

525 pages | 8.5 x 11 | null
ISBN null | DOI 10.17226/12376

AUTHORS

Study Committee and Advocacy Panels, Space Science Board, Assembly of Mathematical and Physical Sciences, National Research Council

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the Councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the Committee responsible for the report were chosen for their special competences and with regard for appropriate balance.

This report has been reviewed by a group other than the authors according to procedures approved by a Report Review Committee consisting of members of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine.

Available from
Space Science Board
2101 Constitution Avenue
Washington, D.C. 20418

FOREWORD

The space age began exactly 20 years ago with the launch of Sputnik I and Explorer I. The Explorer spacecraft discovered regions of trapped radiation around the earth--the van Allen belts. This was the beginning of the study of particles and fields in space, or space plasma physics. A large part of the effort in the early years of the space program was devoted to the mapping of the magnetosphere, the measurements of time variations in particles and fields, and the exploration of the solar wind.

From these studies a sophisticated empirical knowledge of phenomena in space plasma physics has emerged. With the attainment of this observational maturity in the field, NASA funding for space plasma physics has declined as priorities have shifted to other exploratory ventures. The present study of space plasma physics was therefore requested by NASA in order to obtain guidance for future directions in the subject.

The study has involved a major effort on the part of a great many people working in space plasma physics. The Space Science Board formed a panel chaired by Stirling Colgate, composed for the most part of physicists expert in plasma physics but not especially knowledgeable about the space aspects of plasmas; the report of this panel constitutes the first part of Volume 1 of the report. The Committee on Space Physics of the Board was charged with the responsibility for soliciting technical review papers on a large number of topics in space plasma physics. These reviews are Volume 2 of the report; they constitute a most valuable resource for those working in the field. From these reviews, two advocacy panels prepared overview position papers that served as resource information for the Colgate panel and appear as Chapters 7 and 8 of Volume 1.

The Colgate panel has recommended that future research in space plasma physics should involve a much closer integration between theory and observation as is appropriate to the maturity of the field and in order to bring the research into closer contact with the mainstream of plasma physics research. The panel also concurred in the unified recommendations of the advocacy panels.

The Space Science Board is most grateful to the many people who have devoted so much time and effort to carrying out this study.

A. G. W. Cameron, Chairman
Space Science Board

PREFACE

This study was undertaken by a specially created committee of the Space Science Board of the National Research Council, at the request of the National Aeronautics and Space Administration, in order to identify the future objectives of research in space plasma physics.

The Space Science Board divided its study of plasmas in the solar system into four parts. The Study Committee on Space Plasma Physics, chaired by Stirling A. Colgate, consisted of six experts drawn from laboratory plasma physics and plasma astrophysics and the chairmen of three advocacy panels. Its function was to give an overall evaluation of the current status of solar-system plasma physics. The Study Committee on Space Plasma Physics was aided by three "advocacy" panels, whose members were practicing solar-system plasma physicists: a Panel on Plasma Physics of the Sun, chaired by E. N. Parker; a Panel on Solar System Magnetohydrodynamics, chaired by C. F. Kennel; and a Panel on Solar System Plasma Processes, chaired by L. J. Lanzerotti. The Kennel and Lanzerotti panels dealt with plasma phenomena beyond the solar corona. Solar-system magnetohydrodynamics treated large-scale plasma phenomena in the solar wind and at the planets. Solar-system plasma processes considered those microscopic plasma problems that emerged from the study of the objects considered in solar-system magnetohydrodynamics and also considered the impacts of these processes on terrestrial science and technology. The solar-physics panel treated macroscopic and microscopic plasma processes occurring on the sun together and included many topics from what is conventionally called solar astronomy.

Each member of the advocacy panels on solar-system magnetohydrodynamics and solar-system plasma processes wrote a scientific review article on his specialty. Using these review articles as working papers, these two advocacy panels met to compose overview reports that summarized the salient points of the working papers and made recommendations. The Committee on Space Physics of the Space Science Board, chaired by R. A. Helliwell, appointed outside reviewers for each of the scientific review articles and supervised the reviewing process. These working papers are the subject of this volume.

PARTICIPANTS AND CONTRIBUTORS

Study Committee

Stirling A. Colgate, Los Alamos Scientific Laboratory, Chairman
Harold Furth, Princeton Plasma Physics Laboratory
Jack R. Jokipii, University of Arizona
Charles F. Kennel, University of California, Los Angeles
Louis J. Lanzerotti, Bell Telephone Laboratories
Eugene N. Parker, University of Chicago
David Pines, University of Illinois
Marshall Rosenbluth, Institute for Advanced Study
Malvin Ruderman, Columbia University

Panel on Solar System Magnetohydrodynamics

Charles F. Kennel, University of California, Los Angeles, Chairman
Peter M. Banks, Utah State University
Aaron Barnes, NASA, Ames Research Center
Len A. Fisk, NASA, Goddard Space Flight Center
Thomas E. Holzer, High Altitude Observatory
Juan G. Roederer, University of Denver
George L. Siscoe, Dartmouth College

Panel on Solar System Plasma Processes

Louis J. Lanzerotti, Bell Laboratories, Chairman
Donald T. Farley, Cornell University
William C. Feldman, Los Alamos Scientific Laboratory
Robert W. Fredericks, TRW Systems Group
Eugene Greenstadt, TRW Systems Group
Gehard Haerendel, Institute for Estraterrestrischphysik
Lawrence R. Lyons, NOAA, Space Environment Laboratory
Francis W. Perkins, Princeton University
Stanley Shawhan, University of Iowa
Bengt U. O. Sonnerup, Dartmouth College
Peter A. Sturrock, Stanford University

1975 SSB Study Panel on Solar Physics

Eugene N. Parker, University of Chicago, Chairman
Jacques M. Beckers, Sacramento Peak Observatory
Arthur J. Hundhausen, High Altitude Observatory
Mukul R. Kundu, University of Maryland
Cecil E. Leith, National Center for Atmospheric Research
Robert Lin, University of California, Berkeley
Jeffrey Linsky, Joint Institute for Laboratory Astrophysics
Frank B. MacDonald, NASA, Goddard Space Flight Center
Robert Noyes, Smithsonian Astrophysical Observatory
Frank Q. Orrall, University of Hawaii
Laurence E. Peterson, University of California, San Diego
David M. Rust, American Science and Engineering
Peter Sturrock, Stanford University
Arthur B. C. Walker, Jr., Stanford University
Adrienne F. Timothy, NASA, Headquarters
Kenneth A. Janes, Boston University, Study Director

Study Director

Richard C. Hart, National Research Council

CONTENTS

PART II
Solar-System Plasma Processes

Kinetic Processes in the Solar Wind

William C. Feldman

University of California, Los Alamos Scientific Laboratory

Los Alamos, New Mexico 87545

1) Introduction

The solar wind at 1 astronomical unit (AU) is a fully ionized plasma, consisting primarily of electrons, protons, and alpha particles, which streams away from the sun at supersonic speeds. It is a nearby and accessible example of a cosmic plasma. Studies of the internal physical state of the solar wind are of scientific interest in their own right as well as for their relevance to several related physical and astrophysical disciplines. For example, such studies provide diagnostic information useful for placing constraints on theories of the coronal expansion. This information, is carried in part by the plasma ions and electrons and is evident in the characteristic shapes of particle velocity distributions at 1 AU. It is also evident in the hydromagnetic wave field which consists primarily of large amplitude Alfvén waves travelling away from the sun in the local solar wind rest frame.

A second reason for interest in the solar wind is that it is convenient for studying the state and development of plasma turbulence in an astrophysical setting. The type and amplitude of the turbulence determines the rates at which processes such as magnetic field reconnection and particle acceleration proceed. It also regulates the efficiency with which the plasma conducts heat and transports linear and angular momentum. Studies of the kinetic state of the solar wind as it expands away from the sun are therefore of use in obtaining a quantitative understanding of the manner and rate at which a cosmic plasma evolves into a turbulent state as well as of the physics of turbulent processes postulated to be occurring in other astrophysical plasmas.

A third reason for studying the internal state of the solar wind is that the interplanetary medium near 1 AU provides a laboratory in which several (in general nonlinear) collisionless processes can be observed to occur. It is therefore possible to use solar wind observations to quantitatively test theories which hope to describe (and eventually allow a control over) the behaviour of laboratory fusion plasmas. In particular, they are of use for providing quantitative information concerning 1) the thresholds, nonlinear saturation mechanisms, and asymptotic wave levels of various plasma instabilities, 2) details of nonlinear Vlasov-Maxwell equilibria possible in collisionless plasmas, and 3) rates of energy and momentum transport within generally noisy, collisionless plasmas. Such information is essential for achieving the stable confinement of fusion plasmas in laboratory devices. For example, the physics of kinetic heat flux regulating mechanisms is only poorly understood, yet is very important for controlling end losses in magnetic confinement devices and for producing spherically symmetric pellet coronae while limiting the preheating of pellet targets in inertially confined devices. As will be detailed below, similar mechanisms may be active in the solar wind and are currently being explored both experimentally and theoretically.

In addition, possible nonlinear equilibrium plasma configurations are largely unknown. However, a detailed knowledge of the various possible Vlasov-Maxwell equilibria is presently important to both laboratory and solar wind plasma research. Because of the large characteristic distances in interplanetary space, the solar wind convects through a density scale height in about 10^4

proton gyro periods. As a result, there is generally more than enough time for it to reach a steady state with regard to its internal configuration. For comparison, a laboratory fusion device with a 50 kG magnetic field must confine a deuterium plasma for about 4×10^4 gyro periods if it is to operate stably for 1 msec; a minimum time estimated for economic operation. These periods are long compared to typical growth times of micro-instabilities driven by non-Maxwellian particle configurations which can possibly evolve because both the solar wind and laboratory fusion plasmas are at times collisionless. The interplanetary medium near 1 AU therefore provides a laboratory for studying possible steady state configurations of a collisionless plasma. In contrast to the case in laboratory plasmas, very detailed, nonperturbing measurements are possible in the solar wind because at 1 AU, the Debye length (10 m) is larger than an entire spacecraft.

In this report we review some of the recent work on kinetic plasma processes active in the solar wind near 1 AU. The scope of this review will be confined to processes with wavelengths shorter than about 10 proton gyroradii which are driven by non-Maxwellian particle velocity distributions. The very important and interesting problem of the origin of these distributions is not considered. The reason for this decision is twofold. First, although many plausible explanations of the evolution of solar wind particle velocity distributions with distance from the sun have been given, and many more are possible, no plasma processes have been firmly established because of an absence of information about the internal state of the low and intermediate corona. Second a review of longer wavelength fluctuation phenomena of solar origin, which most likely strongly affects the internal state of the solar wind near 1 AU, is covered in an accompanying report.[14] A description of solar wind particle velocity distributions is given first in section 2. This description is followed in sections 3 through 5 by reviews of three of the

many different kinetic phenomena driven by non-equilibrium particle configurations which evolve in interplanetary space. In section 6 a brief description is given of a specific example of particle acceleration thought to occur in the solar wind near and beyond 1 AU. Suggestions for future expansion of this research are given in section 7.

Certain topics have been omitted because it would be premature to present them at their present stage of development. For example, the subjects of viscosity and angular momentum transport which are, at least in part, affected by processes considered in this review have not been included per se although they have important astrophysical application. In particular they are essential for understanding mechanisms 1) of angular momentum shedding in star formation, 2) of stellar spin down and 3) of interactions between binary stars which lead to substantial accretion and intense X-ray emission. Furthermore, although studies of interplanetary shocks and magnetic reconnection are of interest and have broad application, they are not reviewed here. Both subjects can be studied in far greater detail within the near earth environment and are reviewed in companion reports.[38,72]

2) Particle Velocity Distributions in the Solar Wind Near 1 AU

a) Electrons

During quiet conditions, interplanetary electrons of solar origin are observed over a broad range of energies spanning the interval $0 \leqslant E \leqslant 100$ keV.[59,51] In order to quantify the free energy carried by solar wind electrons it is convenient to subdivide the entire energy range into three distinct subranges. For typical quiet conditions the lowest subrange extends from zero to about 60 eV, the intermediate range extends from about 60 eV to several keV, and the highest range covers the interval between several keV up to about 100 keV.[51,23]

A cut through a typical solar wind electron velocity distribution along the magnetic field direction, $\hat{B}$, is shown in Fig. 1.[24,59,58] The measured distribution, f_e, is plotted using solid dots and the lines represent two essentially bi-Maxwellian functions which fit the data best at low and intermediate electron energies respectively. The energy beyond which f_e rises above the bi-Maxwellian function, f_c, that characterizes f_e well at low energies, marks the boundary between the low and intermediate electron-energy subranges. Although for this example, the hotter bi-Maxwellian is seen to represent the measured velocity distribution, $f_H = f_e - f_c$, in the intermediate velocity range quite well, it is not a unique representation. A power law in energy, exponential in speed or some other analytic form may equally well describe f_H. At times, however, the bi-Maxwellian fits to f_H are definitely not acceptable. When the solar wind bulk speed is high, measured electron distributions are more strongly beamed (f_H exhibits a large thermal anisotropy) along the magnetic field direction than can be accommodated by a single bi-Maxwellian function.[25] In fact recent, more detailed measurements of solar wind electron velocity distributions, indicate a better characterization of f_H in terms of a superposition of two separate components; a nearly isotropic hot component and a strongly beamed component travelling away from the sun but along $\hat{B}$.[57] Although this more complex characterization of f_H is only required by two-dimensional measurements such as shown in Fig. 1 in the high speed solar wind, it is more general and therefore probably more useful for understanding the development of the internal state of solar wind electrons during most flow conditions.

Within the energy range spanning the interval between several keV and 100 keV during quiet conditions, interplanetary electron spectral intensities (electrons cm^{-2} s^{-1} $sr^{-1}keV^{-1}$) follow a power law dependence, $dj/dE \propto E^{-\delta}$ with spectral index, $\delta \cong 3.5$.[51] Although velocity distributions within this

energy range fit smoothly onto those measured at intermediate energies, the highest energy distributions are more isotropic.

The above characterization of f_e allows a qualitative discussion of how measured electron velocity distributions deviate from simple bi-Maxwellian functions. A quantitative discussion using two of the many possible ways of characterizing electron distributions is given elsewhere.[23,27] Under appropriate circumstances, these deviations can become sources of free energy leading to instability. The salient deviations are as follows. Electrons in the intermediate energy range are always hotter than those in the low energy range. Although with present two-dimensional measurements it appears possible to characterize the intermediate range electrons in the low speed solar wind in terms of single bi-Maxwellian, bi-Lorentzian, or other similarly shaped functions, this is not the case in the high speed solar wind. At high speeds, measured electron distributions are more complex and require a description consisting of at least two separate components as mentioned earlier. Furthermore, hot and cold electron components move relative to one another along $\hat{B}$ in such a way that the net electron particle flux is zero in the frame of reference moving with the ions. In general, the hot component travels away from the sun faster than does the cold component.

Although electrons in the highest energy range are generally distributed as a power law, occasionally a transient peak appears which moves from high to low energy with time (see e.g. Ref. 52). Peaks in the electron flux distributions near 1 AU have been observed as high as 100 keV[52] and as low as ~6 keV,[39] These secondary peaks are generally associated with interplanetary Type III

radio bursts and are thought to arise because of the transient nature of the source at the sun.[10,81,54] A one-dimensional schematic illustration of this process is given in Fig. 2. For this purpose it is assumed that at time τ_o, energy is deposited in a localized region near the base of the corona. The resultant electron heating produces an enhanced high energy extension on the coronal electron velocity distribution as shown in the top panel of the figure. At some location in interplanetary space a distance ΔR from the heating region, outwardly travelling electrons which were energized at time τ_o, will be observed at time τ_i in the velocity interval above some lower limit, V_i. In the absence of significant scattering, velocity dispersion will therefore produce a peaked spectrum of heated electrons with a sharply cutoff lower velocity limit which depends on ΔR and τ_i as $V_i = \Delta R/(\tau_i - \tau_o)$. The total electron velocity distribution at ΔR will then consist of the sum of ambient low energy electrons and heated high energy electrons. A transient secondary electron peak should therefore appear at the speed $V_i(\tau_i)$ which decreases monotonically with increasing time as illustrated in the lower three panels of Fig. 2.

b) Protons

Proton velocity distributions, f_p, measured in the solar wind near 1 AU range from isotropic Maxwellian to velocity resolved double streaming configurations.[21] Most of the time, and especially during high speed flow

conditions, f_p cannot be described by models consisting of simple extensions to a bi-Maxwellian shape which include a third velocity moment.[78,21] In fact a visual survey of measured proton distributions suggest that much of the time, they can be best described in terms of two unresolved, relatively drifting components, $f_p = f_M + f_B$. Here the subscripts M and B refer to main and beam proton components respectively. Such a description is demonstrated in Fig. 3 for two representative proton distributions measured in the high speed solar wind.[24] In the top two panels of Fig. 3, the squares give the measured velocity distribution integrated over velocities perpendicular to $\hat{B}$, the solid triangles (circles) give the fit to $f_M(f_B)$ in terms of bi-Maxwellian functions, and the open triangles give $f_M + f_B$. (Equally good fits are obtained if bi-Lorentzian instead of bi-Maxwellian functions are used.[24]) Two dimensional contours of both velocity distributions are drawn in the bottom half of the figure.

Quantitative characteristics of the model fits typically obtained during high speed flows are given elsewhere.[24] Of importance here, is a qualitative discussion of the nature and extent of the free energy carried by solar wind protons near 1 AU. This energy is carried mainly by a secondary proton beam convecting along $\hat{B}$, relative to and faster than a main proton component. This motion is observed in association with a distortion of the velocity distribution of the main component such that its perpendicular temperature, $T_{\perp M}$, is larger than its parallel temperature, $T_{\parallel M}$.[21,12] In addition the beam is generally hotter than the main component and only weakly anisotropic. Although in a gross sense $T_{\parallel B} \geq T_{\perp B}$, most two dimensional contour diagrams show evidence that the beam as well as the main

component is being heated perpendicular to $\hat{B}$. For example both contour diagrams in Fig. 3 show a pronounced perpendicular bowing at high proton energies.

c) Alpha Particles

Characteristics of alpha-particle velocity distributions measured in the solar wind near 1 AU are known in far less detail than corresponding proton distributions because they are, on the average, only a 5% constituent. Although striking examples of double streaming configurations have been observed on rare occasions[7] most often alpha distributions appear as a single component convecting relative to and in general faster than the protons.[67,42,63,8,84] During high speed flows, the drift speed of the alpha particles is such that they lie roughly, but slightly more than midway between the proton main and beam components in velocity space.[24] In addition solar wind alpha particles are typically four times hotter than the protons near 1 AU.[67,27]

3) The Regulation of Solar Wind Heat Flux Near 1 AU

Electron heat conduction provides a major means of energy transport in collisionless plasmas. A thorough knowledge of all possible heat flux regulating mechanisms is therefore essential for understanding the behaviour of physical systems containing hot plasmas. Several processes capable of reducing the flux of heat transported through the solar wind have been suggested as applicable within 1 AU of the sun. Most of these processes are either non kinetic in nature or rely critically on assumed plasma conditions which may or may not apply to the inner solar wind. Since in the interest of brevity, the scope of this report is confined to reviewing only those

kinetic processes about which detailed plasma measurements are available, many of these suggested heat flux regulating mechanisms will not be covered. A more comprehensive review can be found elsewhere.[43]

It now appears likely that at least one of many possible kinetic heat flux regulating mechanisms is active at times in the solar wind near 1 AU. Solar wind plasma and field data can therefore be used to obtain a detailed and quantitative understanding of this and perhaps other such mechanisms. In the following paragraphs the strongest evidence indicating effective heat flux regulation near 1 AU in the solar wind is first presented. This is followed by a review and critical discussion of possible mechanisms proposed to explain the measurements.

As mentioned in Section 2a, below several keV energy, solar wind electron velocity distributions can be separated into two relatively convecting, distinct components, f_c and f_H. (Note though, that the hot component, f_H, may often be complex and consist of two distinct parts; one which is relatively isotropic and one which is strongly beamed.[57]) Heat transport in the solar wind near 1 AU appears to result mainly from the bulk motion of the hot electrons relative to the plasma frame of reference, ΔV_H. Simultaneously, the cold electrons move opposite to ΔV_H with relative drift speed, ΔV_c, in such a way that the net electrical current is zero within experimental uncertainties.[23] In other words, the solar wind heat flux, Q_e, is observed to be proportional to both ΔV_c and ΔV_H. Consequently, any kinetic mechanism capable of limiting either ΔV_c or ΔV_H will also limit Q_e.

Recent evidence strongly suggests that ΔV_c and/or ΔV_H are limited by the local Alfvén speed, V_A, near 1 AU. Not only do variations in ΔV_H and ΔV_c follow variations in V_A, but the average magnitudes of ΔV_c and V_A are closely equal.[26] A remarkably clean example of correlated variations between ΔV_c and V_A occurring on a fine time scale is shown in Fig. 4. Even though the ratio, $\Delta V_c/V_A$, is not constant throughout this particular 12 hour period, inspection of the figure leaves little doubt that variations in ΔV_c and V_A are indeed related. An example indicating a correlation between ΔV_c and V_A over a longer time period is reproduced in Fig. 5. These examples as well as others have been interpreted[26] to suggest that the Aflvén speed is at times a prime regulating factor of the solar wind heat flux through the relation $Q_e \propto \Delta V_c \propto V_A$.

Theoretical analyses of heat conduction in the solar wind[30,69,66,33,34] have shown that if ΔV_c is ever larger than about V_A, one of several microinstabilities may develop depending on the values of various plasma parameters. Assuming model velocity distributions consisting of two relatively convecting electron bi-Maxwellians and one proton bi-Maxwellian, the most important instabilities involve the Alfvén, magnetosonic and whistler modes. Whereas the Alfvén mode is driven unstable by the relative motion between the cold electrons and the ions, both the magnetosonic and whistler modes are driven unstable by the relative motion between the hot electrons and the ions.[33,34]

The following grossly simplified overview of solar wind heat conduction near 1 AU can be synthesized from current measurements and ideas as follows. Near 1 AU, heat is carried primarily by electrons with energy greater than about 60 eV, which move away from the sun relative to

the solar wind rest frame. This motion is accompanied by a return current of cold electrons in order to prevent a secular charge build up on the sun. As the entire plasma expands away from the sun, solar wind electrons traverse regions of ever decreasing Alfvén speed so that at some location, ΔV_c and/or ΔV_H become sufficiently large compared to V_A that one or more of the heat flux instabilities become active. The effect of these instabilities must be to reduce continuously both ΔV_c and ΔV_H in order to maintain a marginally stable state. If the instability interacts strongly only with the cold electrons thus reducing ΔV_c, then the interplanetary electrostatic potential must increase sufficiently to reduce ΔV_H (and hence the heat flux) in order to maintain a zero net electrical current.[69,83,23,82] It is expected that this reduction is effected both by a direct deceleration of all weakly interacting electrons and a trapping of a small fraction of previously unbound, intermediate energy electrons. However, it is also possible that the instability interacts strongly with the hot electrons thus reducing both ΔV_H and the heat flux directly.[2,35] In either case the flow of heat is continuously regulated by the ever decreasing magnitude of the Alfvén speed.

Because its phase speed is sufficiently high that solar wind ions cannot interact strongly with its oscillating fields, only the heat flux driven whistler mode has been identified tentatively in solar wind data. This identification has been made primarily in the low speed solar wind when the hot component anisotropy is generally small and f_H can be adequately characterized as a single convecting component. Regulation by whistler waves is suggested because ΔV_c appears to be correlated with V_A only when the temperature anisotropy of the hot electrons is low; a result predicted by the linear theory of the whistler heat flux instability.[34,2] This identification does not preclude the possibility that other

instabilities are also effective during these and other flow conditions, in regulating the flow of heat in interplanetary space. In fact this latter possibility is suggested by observations of correlated $\Delta V_c - V_A$ variations in the high speed solar wind as shown in Fig. 5.

Although the above overview seems reasonable and is consistent with current knowledge about solar wind electrons near 1 AU, a few words of caution are in order. The shapes of both electron and ion velocity distributions are generally more complex than those assumed by theoretical analyses of heat flux regulating mechanisms. Since many of these mechanisms involve resonant instabilities which depend sensitively on the shapes of velocity distributions in the resonant velocity range, the physics of heat flux regulation may be considerably more complicated than the simple picture drawn above. For example, it is not known how effective the heat flux driven whistler instability will be in limiting ΔV_c and ΔV_H. Quasilinear theory predicts that growing whistler waves should only modify electron velocity distributions in a very small region of velocity space.[35] Since the free energy available to these waves from such a region is small, in the absence of other active processes stabilization would then be expected to occur at a very low wave level. However, if low level stabilization actually occurs, then sufficient time is not available for the whistler heat flux instability to effectively limit ΔV_H and hence the heat flux. This result is contrary to observations. A further indication of the inadequacy of a simple minded picture of interplanetary heat flux regulation comes from measured shapes of solar wind electron distributions. In interplanetary space, electrons

above about 2 thermal speeds are generally collisionless (see e.g. Ref. 27). It is therefore difficult to understand 1) why cold electron distributions are so nearly isotropic and Maxwellian out to about 2.5 thermal speeds and 2) why the hot electrons travelling back towards the sun exist at all and appear so isotropic, unless other processes are simultaneously active near 1 AU. We are thus led to speculate that the effectiveness of the whistler and other instabilities in limiting Q_e requires either 1) the simultaneous action of other unrelated instabilities 2) generally noisy or turbulent plasma conditions driven by waves of solar origin, or 3) an inhomogeneous medium in which nonlocal processes are fundamentally important. Finally, even if sufficient time is available for the whistler or other instabilities to be effective in limiting ΔV_H near 1 AU, a full understanding of heat flux regulation requires clarification of the processes which determine the densities and temperatures of f_H in relation to those of f_c.[26] It is likely that a basic understanding of the physics of heat conduction in the solar wind requires a fundamentally nonlinear and inhomogeneous theory in which the particle distributions maintain an equilibrium with self-consistent wave fields and continuously adjust to changing plasma conditions as the solar wind convects away from the sun.

4) Type III Radio Emission

Both the theory and observation of interplanetary Type III radio emission have been reviewed recently.[18,53,71,84] These bursts consist of radio waves which drift in frequency from high values above 100 MHz to low values near about 10 kHz. It is now reasonably certain that they are excited at twice the local plasma frequency by solar generated electron beams with peak energy between 5 and 100 keV.[4,5,31,52,47] These observations pose two basic theoretical problems. It is necessary 1) to understand the plasma mechanism which is responsible for converting the energy carried by an electron beam in a low density plasma into energy carried by plasma oscillations without disrupting the beam and 2) to understand the coupling mechanism which converts plasma waves to electromagnetic radio waves. Such an understanding is essential for proper data interpretation in much of radio astrophysics.

The basic beam-plasma mechanism responsible for the production of longitudinal electrostatic plasma waves has been known for some time (see e.g. Refs. 71 and 53). However subsequent interactions between the waves and the exciter beam, as well as among the waves, which lead both to beam disruption and electromagnetic radio waves, are not very well understood. For example the most important nonlinear mechanisms which couple a plasma wave to ion density fluctuations and to other plasma waves do not scatter the pump radiation out of resonance with the initial electron beam.[13] This fact has two important consequences. First the insufficiently scattered plasma radiation will react back on the electron distribution in such a way as to remove the free energy which drives plasma waves initially unstable. A marginally stable state will therefore be quickly established. Second,

excitation of electromagnetic radiation at twice the local plasma frequency is thought to require two plasma waves travelling in opposite directions. However, a recent three-dimensional calculation shows that the fastest growing wave-wave scattering process does not produce oppositely propagating plasma waves[13] as was thought previously.[64]

A resolution of the beam disruption problem has been suggested in terms of spatially bounded and time dependent electron streams as illustrated in Fig. 2.[10,81,54] If true, then plasma waves will be driven unstable in isolated locations in interplanetary space wherever secondary peaks in local electron velocity distributions happen to form. It is then expected that at the various sites of unstable wave growth, quasilinear relaxation processes will act quickly to establish marginally stable states leading to termination of local wave growth. This cycle will then be repeated at other locations in interplanetary space giving the appearance of a continuous and slow removal of energy from the solar generated electron streams. Recently, evidence consistent with this picture has been found using in situ plasma wave measurements.[39,40] In particular, electron plasma oscillations with sufficient amplitudes to account for the intensity of observed radio waves are observed (although only rarely) in association with solar electron streams. Furthermore when such oscillations are measured they are observed to be intense and intermittent. However, due to as yet unexplained effects, these plasma ascillations have not been observed during the time of rising intensity of electromagnetic waves.

At present a theoretical resolution of the second difficulty has no widespread acceptance although current research in this area is quite active. There appears to be no generally accepted mechanism capable of generating oppositely propagating electrostatic plasma waves which subsequently couple

to form electromagnetic radio waves at twice the local plasma frequency.

5) Ion Beam Regulation

In the past, much theoretical effort (see e.g. reviews in Refs. 68 and 43) has been expended towards understanding kinetic processes effective in limiting the overall thermal anisotropy of solar wind protons.[46,62] These studies assumed that solar wind proton velocity distributions could be adequately characterized by bi-Maxwellian functions. However, recent, more detailed measurements of solar wind proton velocity distributions have shown that the overall thermal anisotropy of solar wind protons is intimately associated with two component proton configurations.[21] Indeed, most of the time, proton distributions in the solar wind are complex and better described in terms of two relatively convecting components then in terms of a simple bi-Maxwellian (see discussion in section 2b and Ref. 24). An understanding of the radial evolution of the internal state of solar wind protons therefore seems to require at the least, a thorough knowledge of ion beam driven instabilities in addition to simple anisotropy driven instabilities. Even though a complete treatment of proton distribution shapes must include interactions with the entire solar wind wave field independent of origin, for reasons given in section 1 the following discussion concentrates only on kinetic ion beam regulation mechanisms active in the high speed solar wind near 1 AU. A comprehensive review of other processes which may be effective in determining the internal state of solar wind protons near 1 AU is given elsewhere.[68,43,14]

The initial approach taken towards understanding ion beam regulation relevant to solar wind plasma conditions, arbitrarily separates the problem into two parts. The first assumes various non-equilibrium two component ion configurations in the absence of significant wave fields. The stability limits of various plasma modes are then determined.[60,61,48,66] Next, the second order effects of the growing waves on the initially unstable,

spatially averaged velocity distributions are calculated (see e.g. Chapter 12 of Ref. 16 as well as references cited therein, and Ref. 36). This approach provides a description of the evolution of particle velocity distributions and fluctuating fields only during the initial stages of instability. In particular it determines neither the nonlinear saturation mechanisms nor the Vlasov-Maxwell equilibria which are eventually established. A knowledge of the final stationary state requires an exact solution of the nonlinear Vlasov-Maxwell equations (see, e.g. Ref. 1 and references cited therein, as well as Refs. 15 and 3).

The solar wind provides an excellent laboratory for studying, in quantitative and nonlinear detail, ion beam regulation in a collisionless, high β plasma (β is the ratio of the particle pressure to the background magnetic field pressure). An overview of ion beam interactions in the high speed solar wind based on current measurements and ideas is synthesized in the following paragraphs.

For reasons not yet fully understood interpenetrating proton streams are observed in interplanetary space. Since the solar wind expansion sets up conditions such that V_A decreases with increasing distance above the coronal base, the relative velocity between interpenetrating proton streams should at some distance become comparable to V_A. At this point, one of three possible modes will be driven unstable depending on the parallel β of the main proton component, β_M.[60,61] If $\beta_M \lesssim 0.35$, then an obliquely propagating ion cyclotron instability will have the lowest threshold. This instability becomes more field aligned as $T_{\perp M}/T_{\parallel M}$ increases. However if $0.35 \lesssim \beta_M \lesssim 0.45$, an oblique magnetosonic instability has the lowest threshold and above $\beta_M \cong 0.45$ a field aligned magnetosonic instability has the lowest threshold.

The next step in the evolution of solar wind proton velocity distributions is, at present, not understood. Multidimensional theories have not yet been developed but are needed to determine the modifications of particle velocity distributions subject to the action of obliquely propagating, ion-cyclotron and magnetosonic waves. However, since for these distributions, free energy is carried by virtue of the relative streaming between beam and main proton components, it is expected that ion beam driven instabilities should cause perpendicular heating at the expense of relative convection energy. In other words, one expects a faster moving ion species to slow down, a slower moving species to speed up and $(T_\perp/T_\parallel)$ of all ion components to increase. In particular the damping of ion-cyclotron waves should accelerate alpha-particles in the solar wind up towards the phase speed of the wave.[75]

It is not possible to determine the marginally stable states of proton beam driven instabilities from quasilinear theory. Instead, such a determination requires a fully nonlinear analysis. Since nonlinear analyses are very difficult, measured solar wind ion velocity distributions have been consulted for guidance. This procedure is expected to provide the desired solutions because as noted earlier, the time required for the solar wind to expand through a density scale height is long compared to typical e-folding growth times of ion-beam driven instabilities. Indeed, statistical studies of shapes of solar wind proton velocity distributions have indicated the existence of a preferred class of configurations that can be summarized by an empirical closure relation.[20] A subsequent, more detailed investigation has led to the identification of a class of ion velocity configurations which are

exact, stationary solutions of the nonlinear Vlasov equation and Maxwell equations.[3] These solutions support a nonlinear ion-cyclotron wave and remove the strong ion-cyclotron wave instability predicted by linear theory[48,50,61,37] which was inconsistent with the persistence of observed proton and alpha-particle velocity distributions in the high speed solar wind.[12,24] The theoretical and experimental proton velocity distributions consist of two interpenetrating beams convecting relative to one another along the average magnetic field direction, $\hat{B}_o$. The higher density main component travels slower than the nonlinear Alfvén wave and has a thermal speed perpendicular to $\hat{B}_o$ which is larger than that parallel to $\hat{B}_o$. The lower density beam component travels faster than the wave and has a thermal speed along $\hat{B}_o$ which is larger than that perpendicular to $\hat{B}_o$. Simultaneously, the alpha particles travel faster than the total proton rest frame in such a way that their speed is close to but slower (faster) than the wave phase speed if $(T_\perp/T_\parallel)_\alpha$ is greater (less) than one. A schematic two-dimensional representation of the proton part of this configuration is shown in Fig. 6.

6) In Situ Acceleration of Energetic Particles

Viewed from a distance, the heliosphere must appear like a source of energetic particles. Most of these particles are injected with high energy into interplanetary space by acceleration mechanisms active in the lower solar atmosphere. However, several different experimental observations indicate that some of these particles are also accelerated by mechanisms active in interplanetary space (see e.g. Refs. 70, 6, 77, 49, 55 and 56).

All such mechanisms are of special interest because they provide clues as to the origin of galactic cosmic rays which carry a non-negligible fraction of the total energy density present in interstellar space.[73] Furthermore one of our windows in astrophysics is in the form of high energy particles. An understanding of particle acceleration mechanisms will therefore help in understanding the physical conditions in the neighborhood of those astrophysical objects which are the major sources of these particles.

One of the many possible acceleration mechanisms which, by inference, must operate only in interplanetary space is presented in the next several paragraphs. This mechanism involves the acceleration of interstellar neutral atoms which penetrate into the heliosphere and become singly ionized by either photoionization or charge exchange collisions. Subsequent to ionization, a small fraction of these particles get accelerated by the solar wind up to energies of about 10 MeV per nucleon.

Details of the interaction between interstellar neutral atoms and the sun have been reviewed elsewhere and so will not be repeated here (see reviews in Refs. 9, 74, 17 and 44). Instead, we will concentrate on the evolution of singly ionized interstellar atoms after their injection into the solar wind. Particular attention is given to He^+ because it is most abundant and should be representative of the heavier ions such as N^+, O^+ and Ne^+.

The evolution of He^+ velocity distributions subsequent to photoionization is not known. The anisotropic distribution shapes which should evolve in the absence of wave-particle interactions[19] have been shown to be unstable.[79,41] Since initial time scales for wave growth are short compared to the expansion time scale, it was postulated that He^+ ions become

quickly incorporated into the solar wind.[80] However a nonlinear analysis has not been made so that wave saturation levels and consequent energy and pitch angle diffusion times are not known.[76] It is therefore theoretically possible that in addition to the above suggestion, that the He^+ ions 2) evolve adiabatically[19,76], 3) become isotropized but not thermalized by subsequent wave-particle interactions[76], or 4) are accelerated by transit time damping of a spectrum of fast mode waves present in the outer solar wind.[28,29]

Only a meager amount of experimental information is presently available to settle this question. Significant concentrations of an ion identified as He^+ were first observed in two solar wind ion spectra.[11] However a subsequent more comprehensive search of solar wind heavy ion spectra yielded only upper limits which were an order of magnitude lower than that expected if interstellar He^+ ions are quickly thermalized.[22] Thus although it is possible that occasional solar wind conditions can lead to complete assimilation, most often He^+ velocity distributions must remain diffuse. The first possible evolution sequence mentioned above can therefore not be dominant. On the other hand an anomolous component of energetic ions has been observed in interplanetary space[45,32,55]; and plausibly interpreted[28,29] in terms of the acceleration of a small fraction (10^{-4}) of the initially low energy heavy ion population (including He^+) of interstellar origin. These observations seem to favor alternative 4. However, since the number fraction of the ions actually observed with high energy is exceedingly small, a firm decision cannot be made at present. It is likely that the bulk of the ions evolve to some nonlinear stationary state characterized by a high energy

extension which is observable to energies up to 10 MeV per nucleon. Future observations which are capable of a unique identification of singly ionized heavy ion species will be necessary to settle this question.

7) Summary and Suggested Future Research

Velocity distributions of solar wind electrons and ions are observed to be very complex near 1 AU. Although they depend ultimately on details of the internal state of the plasma in the low corona, they seem to respond to variations in local plasma conditions indicating that often, enough time is available that nonlinear stationary states are locally established. Perturbations from these states excite a variety of instabilities in isolated locations which saturate quickly and nonlinearly to produce new stationary states. Three examples of this process were illustrated in Sections 3 through 5. Sufficient data are not yet available to determine how velocity distributions of suprathermal ions of interstellar origin evolve in interplanetary space but they may evolve similarly.

The physics of these and related problems not only have intrinsic interest but they have wide application. 1) An understanding of heat conduction regulating processes is important to both astrophysics and laboratory fusion research because both involve hot, fully ionized, collisionless plasmas. Not only may heat conduction be a major source of heat loss from the outer atmospheres of the sun and other stars, but also from plasmas magnetically confined in laboratory fusion devices as well as coronae surrounding pellet targets in inertially confined fusion devices. 2) A detailed and quantitative understanding of radio emission mechanisms is

essential for interpreting radio astronomical measurements in terms of conditions and processes occurring within some astrophysical objects. 3) An understanding of the evolution and eventual stationary states of ion beams traversing hot tenuous plasmas is of particular interest to fusion research in that use of such beams has been suggested as one way of heating the ambient ions in two component Tokamak devices. In addition, a categorization of nonlinear Vlasov-Maxwell equilibria is of general interest to laboratory fusion research because of the necessity here of long time confinement for economical operation. 4) An understanding and assessment of the importance of all mechanisms leading to particle acceleration is important to the field of high energy astrophysics because information concerning conditions within some astrophysical objects comes in the form of high energy particles. In addition, galactic cosmic rays carry a significant amount of the total energy density which permeates interstellar space and so must be important both to the nature and dynamical evolution of the state of the intersellar gas.

The successes achieved during the past five years in solar wind plasma physics point the way, in part, to future efforts. Specific suggested expansions of past work relevant to the topics discussed in this report include the following. 1) Measurement in three-dimensions of solar wind electron velocity distributions should be made on a fine velocity grid for all radial distances inside of 1 AU. In particular, measurements at radial distances as close to the sun as possible as well as over the polar regions will be important for understanding the dynamics of the coronal expansion. 2) The heat flux driven whistler instability should be simulated on the

computer in order to understand its saturation mechanism as well as determine the final nonlinear stationary electron state. 3) A search through existing solar wind data should be made for the signatures of other kinetic processes capable of regulating the flow of heat in interplanetary space. Particular attention should be given to interactions with waves of solar origin as well as with those produced by ion driven instabilities. 4) High time resolution measurements of the amplitude and k-vector orientation of plasma waves should be made in interplanetary space (preferably at two radial distances) to understand the mechanism of plasma wave excitation and conversion to produce Type III radio emissions. Time constants as short as 1 msec may be necessary to detect the short plasma wave bursts expected during the rising portion of these emissions. Simultaneous measurements of suprathermal electron velocity distributions should be made in three dimensions with as high a time resolution (snaphot time) as is practicable. It would be of interest to make these measurements simultaneously using two spacecraft with one as close to the sun as possible in order to study the beam evolution and understand the nonlinear saturation and electrostatic to electromagnetic wave conversion mechanisms. 5) High spectral resolution measurements of solar wind ion velocity distributions in conjunction with ion gyroradius and smaller scale wave fields should be extended throughout the inner regions of interplanetary space in order to determine the origin and radial evolution of double ion streams. It is also important to assess the importance of such streams in the transfer of energy from the sun to interplanetary space. In this regard measurements both as close to the sun as is practicable as well as over the solar poles will be of special interest. 6) Theoretical work is needed in order to understand wave-wave

coupling mechanisms and their resulting coupling rates, for the general case of obliquely propagating electromagnetic waves. It would then be possible to understand the development of, and assess the importance of, this type of turbulence in the solar wind. In particular it is not known whether the magnetic irregularities so prominently observed in interplanetary space constitute a turbulent spectrum or just plasma noise. It is also not known at what wavelengths energy is fed into the system. In addition, a theoretical effort is needed to extend present work on nonlinear Vlasov-Maxwell equilibria to include the general case of a spectrum of obliquely propagating electromagnetic and electrostatic waves. 7) Detailed measurements in three dimensions should be made of the velocity distributions of as many of the solar wind minor ionic constituents as possible in order to probe the nature of the solar wind wave field. 8) Measurements of ion velocity distributions which are capable of a unique identification of singly ionized suprathermal heavy ion species are needed in order to understand the processes which result in the acceleration of initially neutral atoms, injected into the solar wind from the local interstellar medium.

In conclusion, the study of kinetic plasma processes active in the solar wind is a growing field of research which, in addition to its own intrinsic interest, has application to a wide range of laboratory fusion and astrophysics research problems. Recent progress, reviewed in the above sections, has demonstrated this fact. Such progress points the way towards a continuing joint experimental and theoretical effort in understanding some very complex, collective processes, which can occur in any collisionless plasma whether it is produced in a laboratory fusion device or in some astrophysical setting.

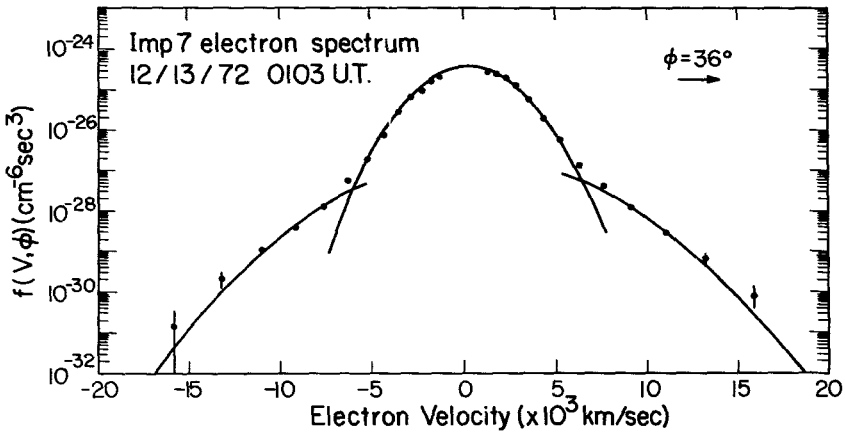


FIGURE 1 A cut through a solar wind electron velocity distribution along the magnetic field direction. The two solid parabolas are the two bi-Maxwellian functions which best fit the low and intermediate energy electrons respectively. The vertical lines through the data points indicate the statistical plus digitization errors where they are larger than the plotting symbols.

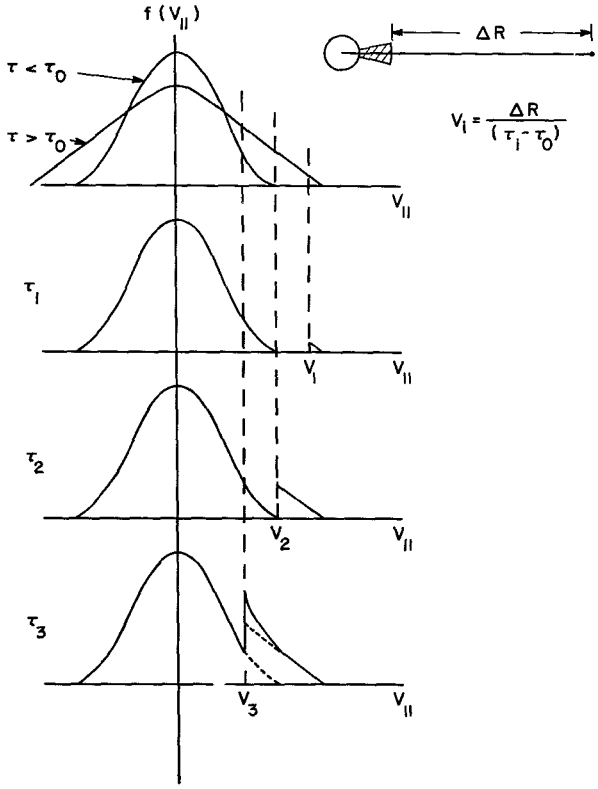


FIGURE 2 A one-dimensional schematic illustration of how a secondary electron peak in interplanetary space can result from a sudden and localized deposition of energy in the corona. The upper panel illustrates the effect of heating at time τ_0 on coronal electron distributions and the lower three panels trace the subsequent evolution of interplanetary electron distributions under the assumption of scatter free propagation of the heated electrons.

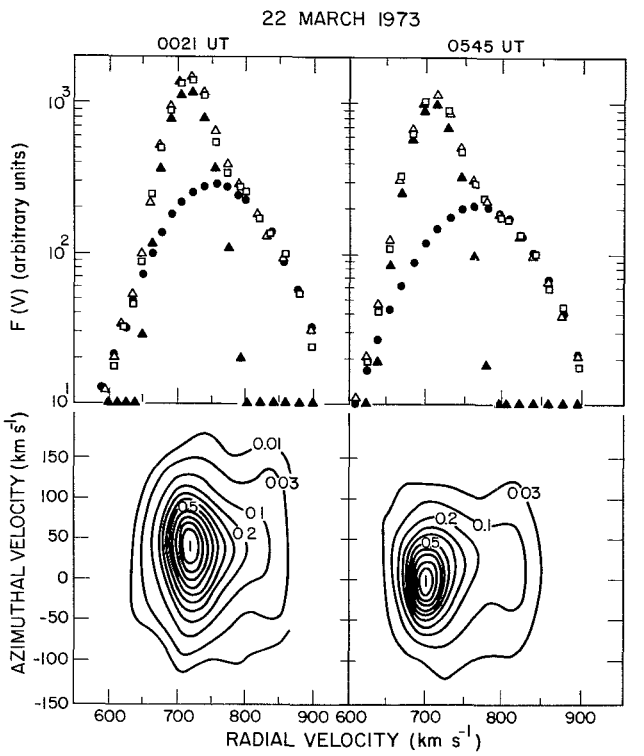


FIGURE 3 A plot of two typical proton velocity distributions measured during high speed solar wind flow conditions. The lower panels show the two dimensional contour diagrams of the measured distributions and the upper panels compare the best fit relatively convecting two component bi-Maxwellian model with the radially projected data. The beam fits are represented by the solid circles, the main component fits are given by the solid triangles, the open triangles give the sum of beam and main component fits, and the measured distribution is plotted by the open squares.

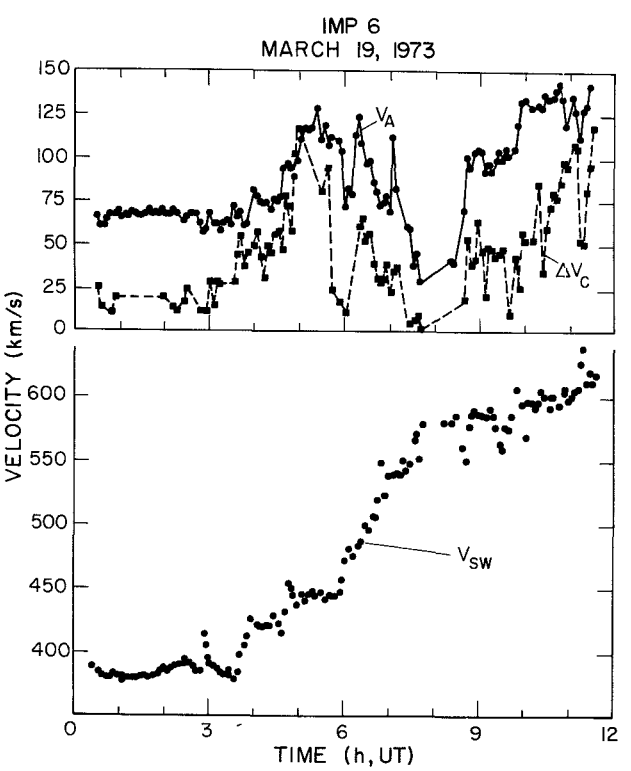


FIGURE 4 A plot of the Alfvén speed, V_A, relative cold component to total electron population bulk speed difference, ΔV_c, and solar wind speed, V_{sw}, during a 12 hour period centered on the maximum speed gradient region at the front edge of a high speed stream.

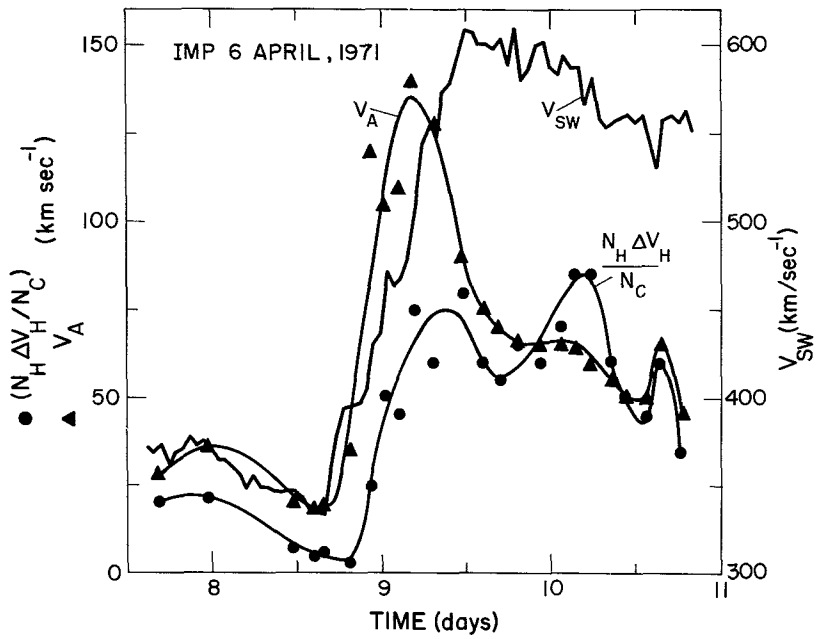


FIGURE 5 Hourly averages of the same parameters plotted in Figure 4 showing their variation through a clean example of a high speed stream. The curve labelled $N_H \Delta V_H / N_c$ in this figure corresponds to that for ΔV_c in Figure 4.

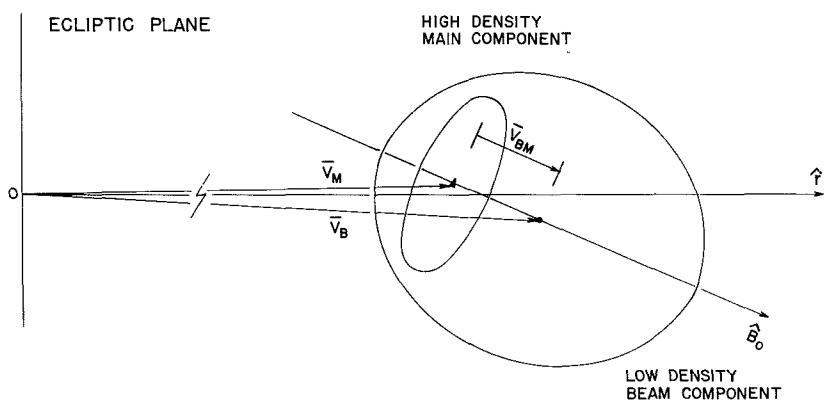


FIGURE 6 A schematic 2-dimensional representation of proton velocity distributions measured during high speed solar wind flow conditions. The $\hat{r}$ direction points radially away from the sun and $\hat{B}_o$ is the direction of the background magnetic field. The vector, $\bar{V}_M(\bar{V}_B)$ is the bulk velocity of the main (beam) proton component relative to a reference frame stationary with respect to the sun and $\bar{V}_{BM}$ is their relative streaming velocity. The ellipses represent velocity contours of the same constant fraction of the peaks of each of the proton components. Since in actuality, the main component peak is substantially higher than the beam component peak (see Fig. 3), the two elliptical contours drawn here cannot be simply combined to yield a contour of the total velocity distribution. The symmetry axes of both ellipses in three dimensions are aligned along $\hat{B}_o$ so that a 3-D contour of the main component is an oblate ellipsoid with a thermal speed perpendicular to $\hat{B}_o$ larger than that parallel to $\hat{B}_o$ (door knob shaped) and that of the beam component is a prolate ellipsoid with a thermal speed parallel to $\hat{B}_o$ larger than that perpendicular to $\hat{B}_o$ (cigar shaped).

Acknowledgments

I wish to thank B. Abraham-Shrauner, S. P. Gary, M. L. Goldstein, J. T. Gosling, D. A. Gurnett and D. F. Smith for many useful discussions concerning parts of this review. This work was performed under the auspices of the U.S. Energy Research and Development Administration.

1) Abraham-Shrauner, B., Exact stationary wave solutions of the nonlinear Vlasov equations, Phys. Fluids, 11, 1162, 1968.

2) Abraham-Shrauner, B., and W. C. Feldman, Whistler heat flux instability in the solar wind with bi-Lorentizian velocity distribution functions, J. Geophys. Res., to be publ., 1977.

3) Abraham-Shrauner, B., and W. C. Feldman, Nonlinear Alfvén waves in high speed solar wind streams, J. Geophys. Res., to be publ., 1977.

4) Alvarez, H., F. T. Haddock, and R. P. Lin, Evidence for electron excitation of type III radio burst emission, Solar Phys., 26, 468, 1972.

5) Alvarez, H., R. P. Lin, and S. J. Bame, Fast solar electrons, interplanetary plasma and km-wave type III radio bursts observed from the IMP-6 spacecraft, Solar Phys., 44, 485, 1975.

6) Armstrong, T. P., S. M. Krimigis, and K. W. Behannon, Proton fluxes at 300 keV assocated with propagating interplanetary shock waves, J. Geophys. Res., 75, 5980, 1970.

7) Asbridge, J. R., S. J. Bame and W. C. Feldman, Abundance differences in solar wind double streams, Solar Phys., 37, 451, 1974.

8) Asbridge, J. R., S. J. Bame, W. C. Feldman, and M. D. Montgomery, Helium and hydrogen velocity differences in the solar wind, J. Geophys. Res., 81, 2719, 1976.

9) Axford, W. I., The interaction of the solar wind with the interstellar medium, NASA SP-308, p. 609, 1972.

10) Baldwin, D. E., Electron overtaking as a cause of instability and type

III solar radio bursts, Physics Letters, 12, 202, 1974.

11) Bame, S. J., A. J. Hundhausen, J. R. Asbridge and I. B. Strong, Solar wind ion composition, Phys. Rev. Letters, 20, 393, 1968.

12) Bame, S. J., J. R. Asbridge, W. C. Feldman, S. P. Gary, and M. D. Montgomery, Evidence for local ion heating in solar wind high speed streams, Geophysical Res. Letters, 2, 373, 1975.

13) Bardwell, S., and M. V. Goldman, Three-dimensional langmuir wave instabilities in Type III solar radio bursts, Astrophys. J., 209, 912, 1976.

14) Barnes, A., Turbulence and dissipation in the solar wind, this Volume, 1977.

15) Channell, P. J., Exact Vlasov-Maxwell equilibria with sheared magnetic fields, Phys. Fluids, 19, 1541, 1976.

16) Davidson, R. C., "Methods in Nonlinear Plasma Theory," Academic Press, New York, 1972.

17) Fahr, H, J., The extraterrestrial UV-background and the nearby interstellar medium, Space Sci. Rev., 15, 483, 1974.

18) Fainberg, J., and R. G. Stone, Satellite observations of type III solar radio bursts at low frequencies, Space Sci. Rev., 16, 145, 1974.

19) Feldman, W. C., D. Goldstein and F. Scherb, Annual variation of interplanetary He^+ velocity distribution at 1 AU, J. Geophys. Res., 77, 5389, 1972.

20) Feldman, W. C., J. R. Asbridge, S. J. Bame, and H. R. Lewis, Empirical closure relation for the Vlasov moment equations, Phys. Rev. Letters, 30, 271, 1973.

21) Feldman, W. C., J. R. Asbridge, S. J. Bame, and M. D. Montgomery, Interpenetrating solar wind streams, Rev. Geophys. Space Phys., 12, 715, 1974.

22) Feldman, W. C., J. R. Asbridge, and S. J. Bame, Upper limits for the solar wind He^+ content at 1 AU, J. Geophys. Res., 79, 1808, 1974.

23) Feldman, W. C., J. R. Asbridge, S. J. Bame, M. D. Montgomery, and S. P. Gary, Solar wind electrons, J. Geophys. Res., 80, 4181, 1975.

24) Feldman, W. C., B. Abraham-Shrauner, J. R. Asbridge, and S. J. Bame, The internal plasma state of the high speed solar wind at 1 AU, in "Physics of Solar Planetary Environments", D. J. Williams, ed., AGU, p. 413, 1976.

25) Feldman, W. C., J. R. Asbridge, S. J. Bame, S. P. Gary, and M. D. Montgomery Electron parameter correlations in high speed streams and heat flux instabilities, J. Geophys. Res., 81, 2377, 1976.

26) Feldman, W. C., J. R. Asbridge, S. J. Bame, S. P. Gary, M. D. Montgomery, and S. M. Zink, Evidence for the regulation of solar wind heat flux at 1 AU, J. Geophys. Res., 81, 5207, 1976.

27) Feldman, W. C., J. R. Asbridge, S. J. Bame, and J. T. Gosling, Plasma and magnetic fields from the sun, in "The Solar Output and Its Variations," edited by O. R. White, J. A. Eddy, and D. Heath, University of Colorado Press, 1977.

28) Fisk, L. A., B. Kozlovsky, and R. Ramaty, An interpretation of the observed oxygen and nitrogen enhancements in low-energy cosmic rays, Astrophys. J., 190, L35, 1974.

29) Fisk, L. A., The acceleration of energetic particles in the

interplanetary medium by transit time damping, J. Geophys. Res., 81, 4633, 1976.

30) Forslund, D. W., Instabilities associated with heat conduction in the solar wind and their consequences, J. Geophys. Res., 75, 17, 1970.

31) Frank, L. A., and D. A. Gurnett, Direct observations of low-energy solar electrons associated with a type III solar radio burst, Solar Phys., 27, 446, 1972.

32) Garcia-Munoz, M., G. M. Mason, and J. A. Simpson, A new test for solar modulation theory: The 1972 May-July low energy galactic cosmic-ray proton and helium spectra, Astrophys. J., 182, L81, 1973.

33) Gary, S. P., W. C. Feldman, D. W. Forslund, and M. D. Montgomery, Electron heat flux instabilities in the solar wind, Geophys. Res. Letters, 2, 79, 1975.

34) Gary, S. P., W. C. Feldman, D. W. Forslund, and M. D. Montgomery, Heat flux instabilities in the solar wind, J. Geophys. Res., 80, 4197, 1975.

35) Gary, S. P., and W. C. Feldman, Solar wind heat flux regulation by the whistler instability, J. Geophys. Res., to be publ., 1977.

36) Gary, S. P., and W. C. Feldman, A second order theory for $\bar{k} \| \bar{B}_o$ plasma instabilities, Phys. Fluids, submitted for publication, 1976.

37) Gary, S. P., M. D. Montgomery, W. C. Feldman, and D. W. Forslund, Proton temperature anisotropy instabilities in the solar wind, J. Geophys. Res., 81, 1241, 1976.

38) Greenstadt, E. W. and R. W. Fredricks, Shock Systems in Collisionless space plasmas, this volume, 1977.

39) Gurnett, D. A., and L. A. Frank, Electron plasma oscillations associated

with Type III radio emission and solar electrons, Solar Phys., 45, 477, 1975.

40) Gurnett, D. A., and R. R. Anderson, Electron plasma oscillations associated with Type III radio bursts, Science, to be publ., 1976.

41) Hartle, R. E., and C. S. Wu, Effects of electrostatic instabilities on Planetary and interstellar ions in the solar wind, J. Geophys. Res., 78, 5802, 1973.

42) Hirshberg, J., J. R. Asbridge, and D. E. Robbins, The helium component of solar wind velocity streams, J. Geophys. Res., 79, 934, 1974.

43) Hollweg, J. V., Waves and instabilities in the solar wind, Rev. Geophys. Space Phys., 13, 263, 1975.

44) Holzer, T. E., The solar wind and related astrophysical phenomena, this volume, 1977.

45) Hovestadt, D., O. Vollmer, G. Gloeckler, and C. Y. Fan, Differential energy spectra of low-energy (< 8.5 MeV per nucleon) heavy cosmic rays during solar quiet times, Phys. Rev. Letters, 31, 650, 1973.

46) Hundhausen, A. J., S. J. Bame, J. R. Asbridge and S. J. Sydoriak, Solar-wind proton properties, Vela 3 observations from July 1965 to June 1967, J. Geophys. Res., 75, 4643, 1970.

47) Kaiser, M. L., The solar elongation distribution of low-frequency radio bursts, Solar Phys., 45, 181, 1975.

48) Lakhina, G. S., and B. Buti, Stability of solar wind double-ion streams, J. Geophys. Res., 81, 2135, 1976.

49) Lanzerotti, L. J., Observations of solar particle propagation, in "Correlated Interplanetary and Magnetospheric Observations", D. E. Page,

ed., D. Reidel Publ. Co., Dordrecht, Netherlands, p. 345, 1974.

50) Lemons, D. S., and S. P. Gary, Temperature anisotropy instabilities in a plasma of two ion components, J. Plasma Phys., 15, 83, 1976.

51) Lin, R. P., K. A. Anderson, and T. L. Cline, Detection of interplanetary electrons from 18 keV to 1.8 MeV during solar quiet times, Phys. Rev. Letters, 29, 1035, 1972.

52) Lin, R. P., L. G. Evans, and J. Fainberg, Simultaneous observations of fast solar electrons and type III radio burst emission near 1 AU, Astrophys. Letters, 14, 191, 1973.

53) Lin, R. P., Non-relativistic solar electrons, Space Science Rev., 16, 189, 1974.

54) Magelssen, G. R., Nonrelativistic electron stream propagation in the solar atmosphere and solar wind: Type III bursts, NCAR cooperative thesis # 37, University of Colorado, 1976.

55) McDonald, F. B., B. J. Teegarden, J. H. Trainor, and W. R. Webber, The anomolous abundance of cosmic-ray nitrogen and oxygen nuclei at low energies, Astrophsy. J., 187, L105, 1974.

56) McDonald, F. B., B. J. Teegarden, J. H. Trainor, T. T. von Rosenvinge, and W. R. Webber, The interplanetary acceleration of energetic nucleons, Astrophys. J., 203, L149, 1976.

57) Miggenrieder, H., M. D. Montgomery, W. G. Pilipp, H. Rosenbauer, and R. Schwenn, Preliminary results of the electron measurements in the solar wind by the Helios plasma experiment, EOS, Trans. AGU, 57, 999, 1976.

58) Montgomery, M. D., Thermal energy transport in the solar wind, in "Cosmic Plasma Physics" edited by K. Schindler, p. 61, Plenum, New York,

1972.

59) Montgomery, M. D., S. J. Bame, and A. J. Hundhausen, Solar wind electrons Vela 4 measurements, J. Geophys. Res., 73, 4999, 1968.

60) Montgomery, M. D., S. P. Gary, D. W. Forslund, and W. C. Feldman, Electromagnetic ion-beam instabilities in the solar wind, Phys. Rev. Letters, 35, 667, 1975.

61) Montgomery, M. D., S. P. Gary, W. C. Feldman, and D. W. Forslund, Electromagnetic instabilities driven by unequal proton beams in the solar wind, J. Geophys. Res., 81, 2743, 1976.

62) Ness, N. F., A. J. Hundhausen, and S. J. Bame, Observations of the interplanetary medium: Vela 3 and IMP 3, 1965-1967, J. Geophys. Res., 76, 6643, 1971.

63) Neugebauer, M., The role of coulomb collisions in limiting differential flow and temperature differences in the solar wind, J. Geophys. Res., 81, 78, 1976.

64) Papadopolous, K., M. L. Goldstein, and R. A. Smith, Stabilization of electron streams in Type III solar radio bursts, Astrophys. J., 190, 175, 1974.

65) Perkins, F., Heat conductivity, plasma instabilities, and radio star scintillations in the solar wind, Astrophys. J., 179, 637, 1973.

66) Perkins, F. W., Ion streaming instabilities: electromagnetic and electrostatic, Phys. Fluids, 19, 1012, 1976.

67) Robbins, D. E., A. J. Hundhausen, and S. J. Bame, Helium in the solar wind, J. Geophys. Res., 75, 1178, 1970.

68) Scarf, F. L., Microscopic structure of the solar wind, Space Science

Rev., 11, 234, 1970.

69) Schulz, M., and A. Eviatar, Electron-temperature asymmetry and the structure of the solar wind, Cosmic Electrodynamics, 2, 402, 1972.

70) Singer, S., The association of energetic storm particles with interplanetary shock waves, in "Intercorrelated Satellite Observations Related to Solar Events," edited by V. Manno, D. Reidel, Dordrecht, Netherlands, 1970.

71) Smith, D. F., Type III radio bursts and their interpretation, Space Science Rev., 16, 91, 1974.

72) Sonnerup, B. U. Ö., Magnetic Field Reconnection, this volume, 1977.

73) Spitzer, L., Jr., "Diffuse Matter in Space," p. 47, Interscience, New York, 1968.

74) Thomas, G. E., Interplanetary gas of non-solar origin, Rev. Geophys. and Space Phys., 13, 1063, 1975.

75) Hollweg, J. V., and J. M. Turner, Acceleration of solar wind He^{++}: 3. Effects of resonant and non-resonant interactions with transverse waves, submitted to J. Geophys. Res., 1977.

76) Vasyliunas, V. M., and G. L. Siscoe, On the flux and the energy spectrum of interstellar ions in the solar system, J. Geophys. Res., 81, 1247, 1976.

77) Vernov, S. N., A. E. Chudakov, P. V. Vakulov, E. V. Gorchakov, N. N. Kontor, Y. I. Logatchev, G. P. Lyubimov, N. V. Pereslegina, and G. A. Tunofeev, Propagation of solar and galactic cosmic rays of low energies in interplanetary medium, in "Intercorrelated satellite observation related to solar events," edited by V. Manno and D. Page, p. 53, D.

Reidel, Dordrecht, Netherlands, 1970.

78) Whang, Y. C., Higher-moment equations and the distribution function of the solar wind plasma, J. Geophys. Res., 76, 7503, 1971.

79) Wu, C. S., and R. C. Davidson, Electromagnetic instabilities produced by neutral particle ionization in interplanetary space, J. Geophys. Res., 77, 5399, 1972.

80) Wu, C. S., R. E. Hartle, and K. W. Ogilvie, Interaction of singly charged interstellar helium ions with the solar wind, J. Geophys. Res., 78, 306, 1973.

81) Zaitsev, V. V., N. A. Mityakov, and V. O. Rapoport, A dynamic theory of type III solar radio bursts, Solar Phys., 24, 444, 1972.

82) Eviatar, A., and M. Schulz, Quasi-exospheric heat flux of solar-wind electrons, Astrophysics and Space Science, 39, 65, 1976.

83) Hollweg, J. V., Collisionless electron heat conduction in the solar wind, J. Geophys. Res., 81, 1649, 1976.

84) Ogilvie, K. W., Differences between the bulk speeds of hydrogen and helium in the solar wind, J. Geophys. Res., 80, 1335, 1975.

SHOCK SYSTEMS IN COLLISIONLESS SPACE PLASMAS

by

E. W. Greenstadt and R. W. Fredricks

Space Sciences Staff
TRW Defense & Space Systems Group

INTRODUCTION

A gas in motion must find a way of flowing around an impenetrable object in its path. Gas particles flowing directly into the object are reflected from it, i.e., they bounce off the object's surface, and in turn collide with other particles farther away, about as far as the mean free path; these collide with others, and so on, until a sufficient amount of the gas is deflected to form a flow pattern around the object. This successive colliding of particles is essentially the transmission of sound through the gas and the transmission travels at the natural sound speed in the gas, "warning" the gas that it is approaching an obstacle.

But what if the gas is supersonic? That is, what if it flows toward the object at high speed, faster than sound can travel backward to warn of the oncoming barrier? In that case the gas particles or, equivalently, the sound waves, pile up ahead of the object, forming a new object, a "shock" wave which is not impenetrable, but through which the gas can travel. On passing through the shock, however, the gas is slowed down and heated, i.e., the particles are scattered by collisions, so that between the shock and the original object the gas is no longer supersonic and flow deflection can take place as it would have subsonically.

But what if the gas is so tenuous as to be collisionless? And what if the gas is a plasma consisting almost entirely of ionized hydrogen, which is really two gases, one of protons, one of electrons? And what if the plasma is magnetized so that particle motion is related to the field and there is not one sound wave velocity but many types of waves with different velocities dependent on frequency and on direction in the plasma with respect to the magnetic field?

Then there is formed a collisionless, plasma shock: a new, "penetrable object" fashioned in a complex way by the interactions of charged particles, magnetic gradients, and wave electric fields.

The problems posed in the preceding paragraphs are far from academic. The gas they describe is the solar wind that occupies the entire known solar system. Moreover, spacecraft instruments have disclosed a collisionless plasma shock in front of every one of the five planets visited so far, and still more of them are continually found traveling through the solar wind away from solar flares and ahead of certain plasma streams originated in the sun. The existence of others has been inferred in the solar atmosphere. The first importance of collisionless shocks derives therefore from their ubiquity in the solar system, where few interactions with or within the solar wind or solar atmosphere can be conceptualized or comprehended without including an appropriate shock "surface."

Shocks are also of great value for study of their physical construction, however, which incorporates some of the fundamental processes common to plasmas. The paramount interest in plasmas in general, whether for understanding or application, in the laboratory, in near-earth environment, or in astrophysical extrapolations, lies in their responses to nonequilibrium, particularly nonlinear, conditions. The collisionless shock provides a nonlinear perturbation that "exercises" the gas flowing through it with many of the wave-particle calisthenics plasmas everywhere are expected to perform. Moreover, the collisionless shock <u>in space</u> performs its exercises on such a large dimension that one or more spacecraft can make detailed measurements within its nonequilibrium scale length. This gives opportunity for really authoritative observations, almost like the idealized concept il-

lustrated in the motion picture "Fantastic Voyage" in which a crew pilots a microscopic craft among the cells of the human anatomy.

Ordinarily, the discovery of so interesting an entity as a natural collisionless shock would be expected to inspire intensive study under the controlled conditions of the laboratory. There have indeed been laboratory-produced shocks and much has been learned from them. In this case, however, nature offers such an excellent observation site that much of the follow-up laboratory work is better done by spacecraft. There are of course advantages and disadvantages inherent in both laboratory and in spacecraft measurements. The advantage of doing shock experiments in a laboratory is the obvious one of control on conditions and consequent repeatability of the experimental measurements. Unhappily, finite dimensions and plasma lifetimes place severe constraints on the ability to measure the full evolutionary characteristics of the shock. Also, density of the plasma customarily is so high that shock scale lengths are smaller than the dimension of the diagnostic probes used, thus obviating acquisition of space-resolved fine structures.

The greatest advantage of spacecraft measurements in either the Earth's bow shock or interplanetary shocks is potential resolution of the fine structure by probes of dimensions much smaller than shock scale lengths. The offsetting disadvantage has been the restriction to single-point measurements of passing structures whose motion relative to the probe is usually unknown, so that ambiguous Doppler-shifted frequencies and scale lengths result. Further complications arise from unknown or poorly-known temporal changes in upstream plasma parameters, with the result that characteristics of magnetohydrodynamic (MHD) fluid elements in the shock cannot be connected to the real characteristics of the same fluid elements upstream at an earlier time.

The single-point restriction has been partially overcome by taking advantage of unplanned but advantageous conjunctions and separations of existing spacecraft to answer some of the large-scale (macroscopic) questions of shock structure. The main benefit has been characterization of the unshocked fluid element. Fortuitous conjunctions have not provided separation distances suitable for isolating microscopic phenomena. Even so, some inferences have been made about the modes of wave-particle interaction responsible for rearranging the MHD properties of the plasma within the shock transition layer.

The "exercises" through which a perturbed plasma restores itself to equilibrium are a marvel of complex instabilities, meaning electric and electromagnetic wavemodes that grow and decay in such a way that their fields can deflect the ions and electrons of the plasma and establish new average MHD parameters of the gas without particles colliding mechanically or electrically. The growth, maintenance, and decay of the wavemodes depend on the numerous quantities it takes to characterize an ionized MHD gas. A few of these are the effective "temperatures" $T_{\parallel}$ and $T_{\perp}$ parallel and perpendicular to the magnetic field $\underset{\sim}{B}$; their ratio $T_{\parallel}/T_{\perp}$; the ratio of electron to ion temperature T_e/T_i; the average density $N \equiv N_i = N_e$; the ratio of thermal to magnetic energy density, $\beta = 8\pi N\kappa(T_i+T_e)/B^2$; and the magnetosonic Mach number of the flowing gas, $M = V/\sqrt{C_A^2 + C_S^2}$, where C_A and C_S are the Alfvén and sonic wave velocities. While a few inferences have been justified regarding the instabilities occurring within some shock profiles under some upstream conditions, the complex of possible interactions has been but partially explored, and a full accounting of shock processes corresponding to all parameter sets has yet to be posted, let alone explained theoretically.

The following sections present the status of our knowledge of shock structure from a plasma physics point of view, based on direct observations in space. We do not cover the formation or propagation of extraterrestrial shocks via the fluid approximation and we do not intend or attempt a critical review of the details, arguments, or contributions that have yielded the picture we present. Our objective is a timely, reasonably documented picture of the collisionless shock and our methods of observing it in space, in which image of the present and future of the subject can be discerned without spectacles. We begin by describing more technically the objects of our attention, and follow with a succinct history of the subject. We then describe what is known of shock structures in space on both large and small scales by synopsizing the documented features of the Earth's bow shock system. The ensuing discussion touches on the properties of shocks elsewhere in the solar system and on the status of shock theory and then describes in some detail an example of how properties of the bow shock system are investigated with paired satellite measurements. We close with some recommendations for the future.

BRIEF HISTORY OF COLLISIONLESS SHOCK EXPERIENCE

"Collisionless shock" is a title applied, by analogy with the familiar irreversible shock in neutral gases, to the case of a plasma sufficiently tenuous such that any classically calculated mean-free-path for particle-particle (coulomb) collisions is much greater than the characteristic length of field, density, or velocity jumps usually associated with "shock fronts." Such collisionless shocks can occur in a bewildering variety in both the laboratory and in space plasmas. This is because the structures of such plasma shocks can depend critically on the parameters of the plasma, e.g., magnitude of flow velocity, directions and magnitudes of intrinsic magnetic fields, details of upstream ion and electron velocity distributions, plasma betas, and of course on the size and configuration of either obstacles to the flow or of sources of driver plasmas. In this review we deal only with hydromagnetic shocks; that is, we talk only about collisionless plasmas in which intrinsic magnetic fields are present and have an influence on shock structures that cannot be neglected.

There are two types of collisionless shocks of interest here. One is a bow shock produced when an obstacle is placed in a "supersonic" plasma flow. The obstacle may be a solid body, an intrinsic magnetosphere, or even a planetary ionosphere. The term "supersonic" may refer to the ratio of flow velocity to some characteristic velocity (sound, ion-acoustic, Alfvén, magnetosonic, etc.) with which information about the presence of the obstacle would be transmitted upstream.

A second type of collisionless shock is produced by a secondary plasma (i.e., a driver plasma) injected into a primary plasma of different parameters). The secondary plasma acts as a piston to produce shock waves in advance of it. Typical examples are interplanetary shocks produced by solar flare plasmas injected into the evaporating solar corona (solar wind). Such shocks, if they are configured to illuminate the earth, may command our attention by producing sudden commencements and sometimes magnetic storms which affect radio communications, power distribution networks, and even telephone networks, not to mention often deleterious effects in communications satellites.

Apparently the first theoretical work on MHD shocks was done by De Hoffman and Teller[14] who derived the MHD analog of the Rankine-Hugoniot relations for gas-dynamic shocks. Montgomery[60] investigated the development of MHD shocks from large amplitude Alfvén waves.

Papers by Zhigulev[103] and Zhigulev and Romishevskii[104] appear to be the first serious conjectures that the continuous solar wind plasma, upon encountering the intrinsic magnetic field of the earth, should produce a standing bow shock front. No dissipation or entropy-changing mechanism was specified in any detail.

General theoretical work on MHD collisionless shocks was published by Fishman, Kantrowitz and Petschek[25], in which non-linear wave-wave interactions were postulated as a dissipation mechanism. In 1962, Axford[7] and Kellogg[55] published independent papers postulating a bow shock around the earth due to a solar wind plasma flowing, at a speed greater than the Alfvén speed, by earth's magnetospheric obstacle.

In 1961, Auer, Hurwitz and Kilb[5] published a computer study of large-amplitude compressional waves in collisionless, magnetized plasma, and produced shock-like solutions. They attributed dissipation to particle orbit-crossings.

The geophysically oriented theoretical work between about 1962 and 1968 proceeded largely along the lines of the "gas dynamic analog," in which the collisionless MHD bow shock was studied grossly by reducing the MHD jump conditions for $\underset{\sim}{B}$, $\underset{\sim}{v}$, ρ and pressure to a set of jump conditions analogous to those for shocks in gases. An excellent bibliography of such efforts is contained in the review by Spreiter and Alksne[98]. While such gas-dynamic analogs enjoyed great popularity and some success in predicting gross configurational characteristics of earth's bow shock, they were incapable of predicting any detailed structure, and in fact gave misleading information on parameters of the plasma flow behind (earthward of) the shock. Dissipation mechanisms remained a mystery, as did even gross structural properties of the real MHD shock.

Meanwhile, other plasma physicists concerned themselves with theoretical details of MHD collisionless shocks, for various reasons. Collisionless shocks were proposed to heat CTR plasmas, so that joint interest was exhibited by both laboratory and space physicists. Kellogg[56], noting that magnetic field gradients inherent in the soliton solutions of Adlam and Allen[1] corresponded to currents which should produce unstable plasma oscillations within the gradient (Langmuir waves, Buneman waves and ion-acoustic waves), postulated that such instabilities would produce turbulent wave-particle interactions and consequent heating, a dissipation mechanism.

Sagdeev[85] pursued this same reasoning. Tidman[99,100] proposed both ion-electron and ion-ion counterstreaming instabilities in magnetic field gradients as the dissipation mechanism and presented a partly quantitative but largely qualitative model for earth's bow shock.

In 1968 and 1969, theoreticians were presented for the first time with two experiments which indicated that microinstabilities were indeed present in shock structures, and were somehow important to dissipative mechanisms. The first of these almost simultaneous experimental results came from the plasma wave detector aboard NASA's Orbiting Geophysical Observatory (OGO-5). The second came from experiments in the TARANTULA device at Culham Laboratories in England under the direction of J. Paul.

The OGO-5 results, published by Fredricks et al.[34,35] showed clearly that strong electrostatic waves were present in the larger magnetic field gradients within the structure of earth's bow shock. Because the probe size was small compared to local Debye lengths as well as gradient scale lengths, these experiments were quite conclusive. However, they suffered from some defects, such as inability to measure field polarization, and unknown Doppler shifting caused by relative motion between the spacecraft and the flowing solar wind. Also, it is clear that the shock structures themselves had intrinsic motion. Although Fredricks et al.[34] were forced to make some unverifiable interpretations of shock thickness to model their shock, the presence of electrostatic wave turbulence was unequivocally demonstrated.

The laboratory experiments by Paul et al.[76,77] and Doughney et al.[13] also clearly showed electrostatic waves in magnetic field gradients in laboratory shocks. Measurements were carried out by laser scattering diagnostics.

Unfortunately, the laser scattering came from a volume of dimensions comparable to the scale length of the magnetic gradient representing the shock front. This did not allow adequate spatial resolution in the experiment to determine unambiguously the exact electrostatic mode responsible for the observed turbulence.

In spite of such drawbacks, these two results stimulated a significant theoretical thrust to investigate just what electrostatic instabilities could be produced by the current systems which produce the detailed magnetic field profiles in MHD collisionless shocks. This was further reinforced by the publication of a more comprehensive study of selected bow shock structures made by instruments aboard OGO-5, and has been confirmed by instruments aboard subsequent flights, such as NASA's IMP-6,7,8 series, ESA's HEOS-1,2 and the USSR's Prognoz series. However, with respect to resolving exactly the details of the plasma wave modes responsible for shock dissipation mechanisms, little substantial has been added to the OGO-5 results. The reason for this is that in order to advance the state of knowledge further, coordinated measurements between two, or among several, satellites at known separation are required. A number of studies of such measurements have been conducted with relatively little impact on dissipation modes, but with appreciable success in defining the overall properties of shock structural forms. The next section summarizes the results to date.

EARTH'S BOW SHOCK SYSTEM

Shock Macrostructure

Figure 1 is a conceptual composite of the overall characteristics of the earth's bow shock system as we now perceive it in an imagined steady-state condition. The shock is defined in the figure by the signature of the magnetic field magnitude carried through it by the solar wind. The nominal shock is not planar, but curved, so that its relationship to a uniform solar wind flow with velocity $\underset{\sim}{V}_{SW}$, carrying uniform field $\underset{\sim}{B}_{SW}$ at an average 45° angle to $\underset{\sim}{V}_{SW}$, is necessarily asymmetric. That is, the local normal of the nominal shock varies continuously from point to point in its orientation with respect to both the uniform upwind (unshocked) flow and the uniform field.

The nomenclature attached to the shock in the figure is defined by the orientation of the local normal to the upstream field, with the guidance of theory and data from laboratory and space observations: perpendicular is taken to mean that the local orientation of $\underset{\sim}{n}$ to $\underset{\sim}{B}_{SW}$, represented by angle θ_{nB}, is within $\arctan (M_e/M_i)^{1/2} \approx 1.8$ of 90°. This criterion is adopted directly from laboratory and theoretical nomenclature[76,101] and is associated with certain specialized shock properties which have not yet had comprehensive verification in space. The orthogonal geometry at the left side of the shock in the figure, where $\theta_{nB} = 0°$, is called parallel, in agreement with theoretical designation[58].

Every geometrical condition other than perpendicular was, before measurements in space, simply called oblique. All the remaining properties,

designations, and distinctions in Figure 1, excepting the standing whistlers, have therefore resulted from spacecraft observations.

When $\theta_{nB} \lesssim 80°$, but $\gtrsim 50°$, standing whistlers are observed upstream from the bow shock, as in the laboratory[80], at least when $\beta << 1$ and the Mach number is low; i.e., when $M \lesssim 3$[22,50]. When θ_{nB} falls below about 50°, the character of the shock is altered radically. Empirical results from satellite observations have therefore divided the shock into two broad categories, quasi-perpendicular and quasi-parallel. Quasi-perpendicular shocks have a monotonic, sawtooth, or wave-step magnetic profile, and resemble the perpendicular shock in being sharply defined, although the ramp through which the field rises to its downstream level is thicker than in the very specialized perpendicular case. Quasi-parallel shocks have multigradient magnetic profiles, are thicker yet, and do not feature a clearly-definable boundary between upstream and downstream field or plasma.

The entire range of θ_{nB} has been surveyed statistically by many shock passages with many satellites. The asterisks in Figure 1, however, mark those more restricted parts of the range which have been observed at high-resolution with comprehensive, but still incomplete instrumentation and with two or more spacecraft at once.

An enumeration of the qualifications to be attached to the very simplified Figure 1 serves as a virtual prescription for the next decade's study of shock structure in space. Before discussing these, the proton velocity distributions and the foreshock need to be mentioned.

The proton distributions are represented qualitatively by the small three-dimensional sketches scattered around Figure 1. A simple, relatively cool anisotropic solar wind is shown at lower right, with its long axis parallel to $\underset{\sim}{B}_{SW}$, since on average, $T_{\parallel}/T_{\perp} \approx 2$ in many solar streams. A modified proton distribution, currently under study and still incompletely defined, is attributed to the solar wind flowing through the foreshock at lower left, where a second group of high velocity return protons, reflected and possibly energized by the shock, is traveling away from it. The "foreshock," to be discussed in greater detail later, is a region attached to, indeed part of, the quasi-parallel structure in which protons (and electrons) reflected from the shock gradients into the solar wind interact with the incoming, unshocked plasma to generate magnetic waves of appreciable amplitude ($\approx$ 1/2 B_{SW} peak-to-peak). The long dashed line represents the foreshock boundary, which follows the return-proton, guiding-center path resulting from the vector sum of the parallel component of ejection velocity and the drift velocity given by $\underset{\sim}{V}_{SW} \times \underset{\sim}{B}_{SW}$.

The flat-topped, heated, bimodal distribution found behind the typical high-β($\approx$ 1), high-M($\approx$ 5) quasi-perpendicular bow shock[61] is shown at upper right, while the sketch at upper left is intended to represent the multimodal, undefined, but predominantly solar wind-like distribution in the quasi-parallel structure, of which fragmentary cross sections have been obtained by the plasma detectors of several spacecraft.

The question marks in Figure 1 accompany those characteristics of the depicted bow shock that represent reasonable extrapolations from measurement but are by no means established facts. The incompletely-determined issues are, explicitly, left to right:

1&2. What is a typical instantaneous velocity distribution within the quasi-parallel, or parallel, structure? Are return protons separable as a group? So far no plasma probe in the shock or magnetosheath has had the temporal or angular resolution to develop a reliable spectrum in velocity space. One source of difficulty is the variability of the field direction, because any anisotropy of the particles would be constantly changing direction in phase with the large amplitude waves in the field, at periods of a few seconds. The best that has been established so far is that there is an average energy distribution in the parallel structure that's distinguishable from either solar wind or quasi-perpendicular magnetosheath plasma[49]

3&4. What, really, is the magnetic structure of the parallel shock? We know the multigradient profile becomes 2 R_E or more thick, but does it really extend far upstream? Or downstream as far as the magnetopause? The principal difficulty here is that the IMF is seldom in a constant direction for very long, and the establishment of an extensive wave region must take a finite time. We haven't had, knowingly, many opportunities to observe the parallel, or quasi-parallel, structure in steady state condition, let alone to do so with adequate instrumentation.

5. What is the true three-dimensional proton spectrum behind the quasi-perpendicular structure? The post-shock proton distribution sketched at upper right is shown as essentially isotropic, but the true shape is undetermined. Isotropy is improbable, particularly for the second peak of the bimodal distribution, which is most likely a high energy component in the ambient field direction. The direction is not always well-defined, however, and observational distributions remain to be documented.

We may now return to the qualifications attached to Figure 1. To begin with, the sketches represent somewhat mixed conditions, for illustrative purposes. The whole magnetic profile is drawn for essentially laminar conditions, i.e., for β and M both low, but the post-shock proton distribution, upper right, is most appropriate for a moderate $\beta(\lesssim 1)$, supercritical $M(\gtrsim 3)$ solar wind. In contrast, the distribution corresponding to the laminar case would show rather little thermalization and no secondary peak [30], which is why it hasn't been drawn. Further, when β is very high ($\gtrsim 8$), the monotonic nature of the quasi-perpendicular ramp is destroyed, at least on a fine time scale of seconds[31]. For high Mach numbers, $M \gtrsim 5$, the magnetic ramp is also more irregular than displayed here, but the demarkation between upstream and downstream states is never in doubt as it always is in quasi-parallel cases.

Mention of β and M provides the basis for the generalized qualification that the shock system exists in a multiparameter plasma in which every combination of values affects shock structure on either the macro or microscopic scale, or both. In addition, there are numerous physical properties of the plasma that simply do not appear in Figure 1, all of which are of great interest: electron spectra, T_e/T_i, $T_{\parallel\pm}/T_{\perp\pm}$, α-particle contributions, <E> wave spectra, <B> wave spectra, are some of these. Finally, the most important geometrical and temporal qualifications are not pictured at all: the figure is drawn to represent the shock's structure in a single plane defined by $\underset{\sim}{B}_{SW}$ and $\underset{\sim}{V}_{SW}$, but all shocks in space are doubly curved in three-dimensional space. Moreover, every parameter characterizing the solar wind varies either continuously or sporadically so any picture of a shock in the

solar wind is soon replaced by another. Indeed shocks often interact with each other as, for example when a fast-stream shock, tipped at some angle, sweeps past the earth, colliding and passing through the curved bow shock, first on one side, then on the other. No hint of such a dynamic situation is indicated in the static, two-dimensional, limited-parameter drawing offered here.

General Scheme

The gross effects of M and β on shock structure are summarized in a classification scheme devised by Dobrowolny and Formisano[17]. Some examples were illustrated by Greenstadt[43] and Formisano[29], and some cases have been described in specialized reports. The classification is elaborated in Table 1, with citations of detailed descriptions where available.

In addition to the entries in the table, some perpendicular, or very nearly perpendicular, cases have been displayed in detail by Rodriguez and Gurnett[81,82]. Their examples would be summarized according to the same scheme as follows:

Parameter Values	Plasma Conditions	Name of Structure	Features
$\beta \approx 1$, $M_A > 3$	Warm plasma, supercritical mach number	PERPENDICULAR, TURBULENT	Sharp field gradient followed by appreciable multigradient fluctuation. Proton distributions unspecified.

For quasi-parallel geometry, the classifications according to parameter have not been clearly differentiated by observation, largely because of the difficulty in knowing whether any given spacecraft passage has encountered

a steady-state or a transient view of the shock. Only one study has produced generalizable details, as noted. The other cases are isolated incidents, possibly of partially developed structures. Wherever no detailed study has been conducted, the parenthesized citation of the Asilomar report (43) provides at least an example of the corresponding case.

In the table, the dividing value $M \approx 3$ is the experimentally-determined value approximating the critical Mach number known in laboratory shock studies[76]. Under q-parallel geometry, and for very high β, strictly local combinations of parameters may play an important role in defining the processes taking place in local gradients.

Microturbulence

An overall description of the wave properties that have been inferred from a combination of statistical and case-history studies in the bow shock is sketched in Figure 2. Electromagnetic noise power P_B, shown here in terms of magnetic field, increases with β to levels roughly an order of magnitude or more higher when $\beta > 10$ than when $\beta \lesssim .1$. Average electric wave strength $\langle E \rangle$ is seen at well over an order of magnitude higher at $M \approx 10$ than at $M \approx 1$, and peak fields in critical portions of quasi-perpendicular shocks may reach two orders higher than typical values in laminar shocks.

The general outlines of Figure 2 result from numerous observations, published and unpublished, of varying degrees of depth. The asterisks and triangles mark the parameter combinations of particular examples that have been exhibited in one or more of the references in Table 1. Most of these have been studied with care so their places in the scheme are reasonably

reliable. The perpendicular cases of Rodriguez and Gurnett[81,82] are included in the figure (circled asterisks). Clearly, the quasi-parallel division is not only undocumented, but largely undefined for warm and hot solar wind conditions.

In dealing with waves, as with particles, energy distributions are important, so the hybrid quantities of Figure 2's vertical axes give an inadequate description of the shock profile. The magnetic power density must be interpreted in terms of its place in the total spectrum, whose form varies from f^{-3} to f^{-4},[73,81,94] and whose integrated power has been found to be about 4×10^{-11} ergs/cm^3 from 1 to 140 Hz[71] and about $x.4 \times 10^{-15}$ ergs/cm^3 from .02 to 4.0 kHz[82]. The vertical scale for P_B in Figure 2 applies at f = 10 Hz. Electrostatic waves are treated differently. The scale of <E> refers to the peak rms field strength for $f \lesssim 3$ kHz, because in the shock there is a noise maximum in this frequency range, formed by many discrete frequencies[35,36,81]. The noise seems to arise largely from ion-acoustic wave generation, some of which is undoubtedly responsible for electron heating in the shock ramp[66].

The role of wave-spectral form in shock characterization can be appreciated by viewing Figure 3. In the center, a magnetic profile along shock normal $\underset{\sim}{n}$ is drawn for a typical supercritical, essentially perpendicular, bow shock in a warm solar wind, with electron and proton velocity space distributions superimposed. At left, sequences of magnetic, and, at right, sequences of electric, spectral "snapshots" in the appropriate scale along $\underset{\sim}{n}$ (guided by Figures 13 and 14 of Rodriguez and Gurnett[82]) have been reproduced

from Figures 3 and 4 of Rodriquez and Gurnett[81]. The magnetic spectra at left show the sharp enhancement of electromagnetic noise corresponding to the main field gradient, but no isolated peak in the measured range of frequencies.

The electric wave noise at right is more complicated. Ahead of the main gradient, where there is a "foot" in the solar wind with associated small amplitude precursors, there is also a clear wave resonance at about 10^4 Hz marking the electron plasma wave frequency f_{pe} corresponding to the upstream density. The resonance is very likely associated with electrons streaming through the solar wind along $\underset{\sim}{B}_{SW}$ at the outer edge of the shock transition layer. Strictly speaking, this feature should be most apparent when $\underset{\sim}{B}_{SW}$ is not exactly perpendicular to $\underset{\sim}{n}$, allowing electrons to stream outward ahead of the shock proper.

Once the main (average) gradient in B is encountered, the f_{pe} noise at 10^4 Hz vanishes and the lower frequency end of the spectrum becomes enhanced. Unlike the magnetic noise, however, the electric noise has a shoulder, or even a local maximum, between .1 and 1 kHz, indicating the presence of ion acoustic wave noise at the local proton frequency f_{pi}. This characteristic of the electric spectra persists, at lower levels, into the downstream plasma. The differences between B and E spectral signatures motivated the different presentations of their scales in Figure 2.

The particle velocity distributions, symbolized by the negative and positive signs in Figure 3, are altered by the shock in several ways. The electrons, hotter in the solar wind than the protons, are partially scattered in the foot of the ramp, but toward the rear of the field gradient, and beyond, they become rather uniformly scattered, giving the flat-topped cross section observed by Montgomery et al.[61]. The protons begin in the solar wind with $T_{\|}/T_{\perp} \approx 2$, and are slowed slightly in the foot with some scattering of particles to higher energies. They become partially heated in the ramp and exhibit a secondary, high energy peak, and by the rear of the main gradient have been scattered to higher "temperature" than the electrons, but with a nonmaxwellian distribution featuring a high energy tail. The directional properties of the proton velocity distributions, i.e., their degree of anisotropy, are known only in the solar wind, so the sketches are illustrative rather than representational. In and behind the shock ramp, distributions represented by double, or even multiple bubbles, might be appropriate.

Returning to Figure 2, the most important message of the figure is that microscale magnetic and electric wave activity are somewhat independent of each other, but are both low for the bow shock in cold, low M solar wind flow and both high for the bow shock in hot, high M solar wind flow. The variation of P_B with β has not revealed any inflections, but the Mach number dependence of $\langle E \rangle$ appears to undergo a significant change at $M \approx 3$. This is a reasonable phenomenon to expect since in the absence of particle collisions it is plasma waves that presumably supply the electric fields needed to scatter and heat both electrons and protons in the shock transition, and the critical Mach number at which resistive dissipation by drift current instabilities becomes inadequate is at $M \approx 3$. Thus when $M > 3$ either a different,

more potent wave mechanism or two separate mechanisms operate to produce higher dissipation than is required when $M \approx 3$. The investigations of Rodriguez and Gurnett[81,82] revealed that for perpendicular and quasi-perpendicular shocks of supercritical magnetosonic Mach number, electric wave noise in the .02-4 kHz band tends to rise with T_e/T_p or, alternatively, since T_e doesn't change very much in the solar wind, tends to fall with increasing T_p. Thus we know T_e/T_p plays its part in determining microscopic shock structure, as expected, but it must be noted that the scatter in the cited correlations was large and significant work is still to be done. Certainly the record of influence of this parameter must be extended to include the entire ranges of the other macroscopic parameters. Similar diagrams based on T_e/T_p, $T_{p\parallel}/T_{p\perp}$ etc. await the outcome of future investigations.

Foreshock

One of the most interesting discoveries of space plasma physics has been the shock-generated region of particles, waves, and wave-particle interactions that inhabit a vast volume of space outside the bow shock. The existence of such a region where $\underset{\sim}{B}_{SW}$ is parallel to $\underset{\sim}{n}$ was broadly predicted by Kellogg[55], who modeled the regions in terms of whistler waves under stationary conditions. Experimental data, almost always describing transient conditions, have disclosed a somewhat different picture that may include the developmental stages of Kellogg's model, but the latter has yet to be documented. We describe the observational results.

Charged particles exist in the foot of the classical shock for various reasons: some have failed to negotiate the shock electric potential; some

have been reversed in direction by the enlarged field in the ramp gradient; some have been deflected by plasma waves in the shock transition. Many of these have guiding center velocity components opposite to, and greater in magnitude than, V_{SW}, and can therefore leave the shock "upwind." These spiral around B_{SW} into the solar wind with net motion outward away from the shock. As they do, they move, in the shock frame, according to the expression

$$\frac{d\underset{\sim}{V}}{dt} = (\Omega_c/B_{SW})\ (-\underset{\sim}{V}_{SW} \times \underset{\sim}{B}_{SW} + \underset{\sim}{V} \times B_{SW}),$$

where Ω_c is the appropriate cyclotron frequency[10,97]. The result is a modified spiral trajectory displaced from $\underset{\sim}{B}_{SW}$ in the shock frame, with net guiding center velocity $\underset{\sim}{V}_r = \underset{\sim}{V}_\parallel + \underset{\sim}{V}_d$, where $\underset{\sim}{V}_d \cdot \underset{\sim}{V}_\parallel = 0$, as shown in Figure 4(a). We designate the upstream region occupied at any instant by some species of particles and/or the waves they excite the "foreshock."

In practice, space physicists have formed the habit of representing the particles as leaving the shock with guiding center velocity pV_{SW} along $\underset{\sim}{B}_{SW}$, while the field itself is carried downstream with the particles, at speed V_{SW}. The result is the same, as Figure 4(a) indicates, but the artificial $p\underset{\sim}{V}_{SW} + \underset{\sim}{V}_{SW}$ resolution of $\underset{\sim}{V}_r$, which arose from observations of waves rather than particles, has lent itself more readily to experimental treatment. We expect this practice to decline as improved particle detection mandates a more direct analysis of the foreshock's material constituents in terms of their actual pitch angle velocities $\underset{\sim}{V}_\parallel$ and $\underset{\sim}{V}_\perp$ determined by plasma-physical processes in the shock. For the present, however, we shall continue to deal with $\underset{\sim}{V}_r$ and θ_{XF} using experimental results on the value of p.

Since the vectors governing the foreshock are $\underset{\sim}{B}_{SW}$ and $\underset{\sim}{V}_{SW}$, and the former is variable in time, the action at a given instant takes place in the infinite set of parallel planes containing the solar wind's flow velocity and its magnetic field. One of these planes is illustrated in Figure 4(b) for a sunward IMF oriented about 30° above the ecliptic. Electrons are hot, around 1.5×10^{5}°K[24], and are the first to be affected at the outermost fringe of the shock[61]. Their $\underset{\sim}{V}_{\|}$ is high ($p_{-} \approx 10$) and they are not displaced very much from $\underset{\sim}{B}_{SW}$ as they progress upstream[24]. The negatives (dashes) in Figure 4(b) signify the region occupied by return electrons. Most return protons are appreciably slower than electrons ($p_{+} \approx 2$) and $\underset{\sim}{V}_{d}$ is a large part of $\underset{\sim}{V}_{r}$. Return protons are therefore displaced considerably from $\underset{\sim}{B}_{SW}$. Their region is denoted by the pluses in Figure 4(b).

Figure 4(b) displays the complexity of the foreshock to first order: there is an electron foreshock and a proton foreshock. But there is also much more. There is of course the three-dimensional foreshock consisting of the infinite set of planes intersecting the shock, parallel to the one depicted. There are also electrons and protons upstream with energies, i.e., velocities, differing from those described above, because there is a spectrum of each return species at any given point of origin in the shock. Further, these spectra may vary from point to point in the shock. In addition, return electrons and protons each generate waves in the oncoming solar wind by their interaction with it, and these waves are carried downstream with the plasma at doppler-shifted frequencies. The most prominent of these waves appear to be caused by the protons with $p \approx 2$. They are detected by magnetometers as quasi-sinusoidal oscillations of amplitude

about $B_{SW}/4$ superposed on the IMF[19,47] and are the principal phenomenon by which the foreshock has been studied and characterized. These waves constitute what may be called the ULF foreshock.

The varieties of particle energies and wave frequencies that have been detected outside the bow shock in the cislunar region have been enumerated in a recent summary review[44]. Here, we shall review newer results and describe the current outstanding issues involving the foreshock.

Two properties of the foreshock have been the subject of five recent studies. These properties are the variability of p and the interaction of foreshock constituents with the solar wind.

A very curious result of early investigation of upstream ULF waves and quasi-parallel structure was the reproducibility of the q-perpendicular to q-parallel transformation at the tangent point of the ULF foreshock boundary when the value p = 1.6 was used to compute the angle of the boundary $\theta_{XF} = \arctan [p \sin \theta_{XB}/(p \cos \theta_{XB} - 1)]$[41,47]. However, most early observations were confined to the subsolar region of the appropriate B-X cross section. A dependence of p on position away from the subsolar point has emerged in a statistical study by Diodato et al.[15] that covered most of the upstream region. Diodato et al. found that the overall best value was p = 2 on the sunward (of the earth) side of the shock and that p increased with distance from the subsolar point.

In application the Diodato et al. result means that when the passing orientation of $\underset{\sim}{B}_{SW}$ places the changeover point from local q-perpendicular to q-parallel structure toward either flank of the shock, the ULF foreshock

boundary, which is tangent to the shock at the changeover point, is a little more forward (i.e., θ_{XF} is a little less) than it would have been were p constant. Equivalently, the backstreaming protons assumed responsible for the forwardmost ULF waves are not displaced as much from $\underset{\sim}{B}_{SW}$ when they originate near the flank as they are when they originate near the subsolar point. The most straightforward physical inferences are that either a larger proprotion of return particle energy appears as $\underset{\sim}{V}_{\|}$, or return particles are more energetic along the flanks, or both. Figure 5(a) shows the crude dependence of θ_{XF} on θ_{XB} implied by the Diodato et al. study. The broken and overlapping nature of the graph results from their study's division of data according to discrete ranges of θ_{XT} (defined in Figure 5(b)).

The second property of the foreshock that has received experimental attention is its influence on the solar wind. Unfortunately the newest results are in conflict with each other. If return particles interact with the plasma to produce waves, as theory and observation suggest[8,19,33,35,36,88] then there ought to be some discernible effect of the interaction on the particles themselves. In fact, we know that return electrons change the heat flux in the electron foreshock, reversing its direction and elevating $T_{e\|}/T_{e\perp}$ slightly by about nine percent[24]. But what of the protons?

A study by Feldman et al.[23] set out to examine the thermal properties of the protons in the proton foreshock and discovered no temperature effect but a slightly lower thermal anisotropy. The data used for the study, however, did not include a magnetometer or return proton detector, so the direction of the electron heat flux was used as a guide to the ambient regime applicable to each data point. Since this criterion identifies only

the electron foreshock, at most an unknown fraction of the observations came from the proton foreshock as well. What is significant is that an anisotropy distinction was found even though the proton foreshock was misidentified in some cases. Presumably the effect would be more pronounced in an undiluted sample.

Auer et al.[6] and Formisano and Amata[28] undertook statistical analyses of the solar wind proton properties associated with the ULF foreshock and reported what appeared to be distinguishable differences between N, V, and $T_{\|}B^2/N^2$ in the foreshock and in the unperturbed solar wind. However, a very extensive statistical study by Diodato and Moreno[16] involving almost 11,000 hourly averages of data from four satellites, using both single and paired spacecraft observations, found no significant differences between the foreshock and the undisturbed solar wind in speed, proton density, or proton temperature. Moreover, Diodato and Moreno demonstrated that an automatic biasing effect of variations in N on shock location could explain the apparent differences reported earlier.

For the time being the composition, behavior, and influence of the foreshock's various components remain open subjects of inquiry. In addition to the obvious need for further experimental results along the same lines as those already described, requirements for numerous other measurements and analyses can easily be generated with a few moment's thought. Most important is the necessity for direct measurements of return particle characteristics throughout the foreshock and for their correlation with local wave observations.

Besides detailing the upstream plasma interaction processes themselves, future investigation must discover the sources of return particles, their distributions, and the forces governing their creation. We have glibly referred to return particles in the foreshock after touching briefly at the beginning of this section on the role played by the "foot" of the shock in supplying them. But the foot accompanies quasi-perpendicular structure, whereas the foreshock, or foreshocks, exist upstream only from the quasi-parallel structure. Where is the foot, or what takes its place, in the q-parallel structure with its multiple gradients, irregular profile, and quasi-periodic oscillations? Neither theory nor experiment, laboratory or extraterrestrial, has yet supplied answers to this question.

Finally, the "steady state" question has yet to be explored even to the extent such exploration is possible in the constantly varying solar wind. The character of quasi-parallel shock pulsations appears superficially to be consistent with that of large-amplitude whistler-like waves, and we know a large portion of the whistler spectrum can propagate outward from the shock along $\underset{\sim}{B}_{SW}$. There is no reason at present to doubt that under stationary solar wind conditions a very extensive whistler foreshock should be evident where θ_{nB} is small, as predicted by Kellogg[55]. Further measurement and analysis should succeed in testing this picture at least to a satisfactory steady state approximation.

Magnetosheath

The entire region between a detached bow shock and the obstacle it shields from a surrounding supersonic flow is not ordinarily thought of when referring to the post shock fluid. In the case of a magnetospheric

bow shock, however, such an extension of terms is unavoidable. Only in that section of a magnetospheric bow shock where quasi-perpendicular structure prevails can the traditional concept of average fluid properties just behind a shock "front" be maintained as a basis for describing the post shock "jump conditions." Elsewhere, the downstream "turbulence" appears to occur on an amplitude scale comparable to or greater than the nominal jump itself and on a spatial scale that may extend the entire distance to the magnetopause. Speculations on the extent to which the magneto-sheath, and even the magnetosphere, participates in quasi-parallel shock structure have been published before[42] and are being actively investigated with available data resources[48], so we avoid repetition here and concentrate on issues related to shock mechanics.

From a plasma-physical standpoint, the downstream regime is of interest for several reasons. If we think first in terms of isolating q-perpendicular and q-parallel cases locally, there are two reasons immediately:

1. The magnetosheath supplies the post-jump conditions needed to characterize the q-perpendicular shock. Moreover, the conditions must be catalogued for a wide range of each upstream parameter, and there are numerous parameters.

2. The definitions of "preshock" and "post shock" are parameter-dependent in individual cases. For example, the magnetic ramp, or gradient, i.e., the "jump," is not strictly superposed on the density ramp. Especially in the cases of extreme parameter values such as very high betas or Mach numbers, a delayed or expanded particle redistribution mechanism may require appreciable post-shock data collection to deduce the physical processes

making up the shock profile. The electrons may be principally responsible for one shock subprofile and the protons for another. One model of proton heating at supercritical Mach numbers, for example[64], does not call for any prominent "viscous" dissipation instability in the ramp, so that plasma wavemode identification would be fruitless and only substantial particle observation well behind the principal magnetic ramp would allow verification of this model. In the case of q-parallel structures, it appears that the magnetosheath is the shock to some extent. We simply do not know at present just where the downstream, or post-shock regime begins.

If we expand our inquiry beyond isolation of separate cases, we find two more penetrating reasons why magnetosheaths are interesting to plasma-shock investigators:

3. Collisionless shocks in the solar wind are three-dimensional, with magnetic field lines connecting one region of the nominal surface, through the post shock standoff region, to the other "side." If one side is q-perpendicular in a uniform field, the other side is q-parallel. Hence there is the possibility of communication by field-aligned phenomena propagating from one part of the nonuniform shock to another through the downstream region. The three-dimensional shock favored by nature is an entity by itself with a continuity of parameter changes differentiating it from the two-dimensional cases that might be found locally, reproduced in the laboratory, or approximated in theory.

4. Finally, the most exciting reason of all, is the opportunity to observe transient plasma processes at high resolution. The magnetosphere is the source of phenomena that communicate the presence of an obstacle

to the solar wind. The magnetosheath therefore is the location in which the processes originate that continuously form and reform the various shock structures in response to changing solar wind conditions. The dimensions of a typical magnetosheath, being on the order of many planetary radii, transform ephemeral phenomena that might take microseconds or nanoseconds in the laboratory into finite events that can be recorded over intervals of seconds and minutes, with almost arbitrary precision.

The reader will have perceived that the straightforward summary of physical results in the preceding sections has been replaced here by an undocumented discussion of justifications for obtaining physical results. This is because spacecraft data have not yet yielded generalizable results of plasma physical application except in a few instances. Results close behind the shock front have been described already in Figures 1 and 3 and in an earlier review[44]. Deeper in the earth's magnetosheath, the most pertinent results are, first, that ion acoustic wave noise continues to modify the solar wind electric wave spectrum well behind the quasi-perpendicular magnetic ramp[81], and, second, that suprathermal protons of energy > 100 keV appear in sharp correlation with large amplitude magnetic field oscillations similar to those of q-parallel pulsations[102].

Improvements in instrumentation and expanded data analysis will be necessary to catalogue magnetosheath properties and relate them to the bow shock.

PLANETARY SHOCKS

Four planets other than the earth have been found by direct measurements with spacecraft instruments to have identifiable bow shocks. These are Mercury, Venus, Mars, and Jupiter. In addition, there is evidence for a limb shock associated with the moon's wake.

Four results regarding planetary shocks are of plasma-physical significance.

1. The overall classification of shock structures into quasi-perpendicular and quasi-parallel categories is generalizable to shocks at least as widely spaced in the solar system as Mercury[71] and Mars[40] and apparently as far as Jupiter[93] too. The local characteristics of shocks of both categories, with regard to magnetic field and electron behavior, are familiar from the earth's bow shock.

2. The relatively swift passage of the respective mission spacecraft through their target magnetospheres has provided bilateral views of nonuniform bow shocks difficult to obtain at earth even with multiple spacecraft observations. Thus the coexistence of q-perpendicular and q-parallel structures on opposite faces of the nominal shock has been demonstrated at Mercury[67,68], Venus[40], and Mars[40]. The data from Mercury are particularly striking because the shock profile during the third pass of Mariner 10 was virtually a mirror image of the pass two years earlier, with the macrostructures encountered in reverse order because the interplanetary field orientation during the third pass was complementary to the orientation during the first pass[67,68,71,72].

3. Each pair of bilateral shock views just cited has, of course, been part of a one-dimensional cross section of a corresponding magnetosphere-solar wind interaction region. These planetary cross sections have therefore provided pictures of the extensive downstream regions of their respective shocks. A prominent, repeated feature has been the penetration of the q-parallel magnetic pulsations deep into their post shock magnetosheaths. Thus these planetary passes have served as principal sources for the inference that the q-parallel shock and its downstream sheath are essentially inseparable, and that pulsation macrostructure stretches all the way from the magnetopause to some, as yet undefined, outer pulsation boundary. Magnetosheath involvement in q-parallel oscillatory structure has yet to be systematically studied in the earth's interaction region.

4. The fourth result from probes to other bodies is that the solar system offers some unusual cases of special interest not to be found in the earth's vicinity. One is the moon, which having no intrinsic field of dimension comparable to itself, but possessing small local conductive or magnetic anomalies, offers the opportunity to examine marginal collisionless shock formation and maintenance near the scale limits required by theory[91]. A second case is Venus, where the obstacle in the solar wind's path is neither a strong magnetosphere nor the absorptive surface of the planet itself (or its atmosphere), but an induced magnetic field arising from currents in the ionosphere. Here, there is a possibility that the solar wind is partially absorbed in the upper atmosphere and that the shock may be so close to the ionosphere that the latter may be part of the downstream region, and shock effects could play a role in the planet's meteorological energy budget. Certainly Venus presents the prospect of displaying some unique plasma-shock phenomenology[84]

A third case is Jupiter, whose magnetosphere may be so rotationally asymmetric, so tilted, and so rapidly revolving as to create wholly unfamiliar nonuniformities and asymmetries covering new areas of parameter space not seen elsewhere[26]. Little data of significance to plasma shock study has yet been released, however.

There is only one example of detailed study of an alien bow shock system from a plasma physical viewpoint. Fairfield and Behannon[21] analyzed magnetic field data from Mariner 10 and found whistler waves propagating upstream from the shock as observed previously at the earth[20]. They went further, however, and identified some of the waves in the Hermean magnetosheath as ion cyclotron waves, a result yet to be achieved at the earth.

ASTROGENIC SHOCKS

All shock phenomenology described so far pertains to shocks formed by supermagnetoacoustic flow past large, effectively stationary obstacles. The other great class of shocks on space pertains to those formed in front of objects moving at supermagnetoacoustic velocity through relatively slow, although nonstationary, plasma. The "objects" contemplated in this class are not solid bodies or their magnetospheres, but masses of plasma of sufficient density and conductivity to appear impenetrable to the background hydromagnetic gas through which they proceed. We shall call such an object a "driver plasma," or "driver gas," and its shock a "driven shock," although we intend to include in this category the "blast shock" which is very much detached from its driving source.

Driven shocks differ from planetary shocks in two major ways; scale and speed. The sources of driven shocks are activity centers on the sun and, by implication, on other stars, and stars themselves. Should these pulse or explode or, in the case of binaries, revolve at high enough orbital velocities, shocks would presumably be generated in the interstellar medium. The scale size of such shocks is therefore of heliospheric or stellar, rather than magnetospheric, proportions. This means that wave-particle effects that may depend on tangential motion along a shock front have an immensely enlarged arena in which to operate. Thus, multiple charged particle acceleration associated with a shock front could result in particles developing cosmic ray energies before being released by local alterations in field or shock geometry to escape into the general interstellar environment.

The speed of a driven shock will depend on the speed of the plasma piston that drives it. The latter varies depending on strength of the source and distance from the source, whereas the flow speed of the solar wind is essentially constant with distance from the sun after an initial acceleration in the outer corona. Although the absolute value of solar wind velocity changes as the source region changes, the nominal speed is a few hundred km/sec and usually varies by no more than a factor of about two to one. Piston plasmas and their shocks, in contrast, may race through the solar atmosphere or outward from the sun as fast as a few thousand km/sec, slowing appreciably as they progress outward through the solar system. Since shocks driven from solar flares or by high-speed streams have a wide range of

velocities, they may include both slow shocks and shocks of extreme Mach number which makes their observation worthwhile, if difficult, at customary data rates. Once again, however, particle acceleration is involved; to the extent that energy exchange between a shock and an individual ion contributes to the ion's gain in energy, driven shocks have the potential of accelerating ions to much higher energies than do planetary shocks.

The property of accelerating ions gives the subject of driven shocks a value for those concerned with high energy cosmic particles, because there is ample reason to believe that collisionless MHD shocks are commonplace in the astrophysical environment. The most primitive example would be the shock expected to exist in front of the heliosphere as a result of its motion in the galactic plane[52]. Presumably such shocks can be attributed to the myriad of stars later than F5 where the existence of stellar winds is inferred[79]. A more sophisticated example has been envisioned by Siscoe and Heinemann[92], who have proposed that binary stellar winds might include opposing shocks bracketing the contact discontinuity between appropriately spaced star-pairs. Finally, the velocities of supernova ejecta and remnants may be as high as several thousands of km/sec and are quite likely to drive high Mach number shocks ahead of them into the interstellar region[11]. If such examples are multiplied, there are numberless opportunities for particles to be accelerated throughout the universe by shock waves of many origins. Study of the energization mechanism taking place in the solar system should have direct application to such astrophysical sources.

Locally, we know that accelerated particles are associated with the bow shock, but neither their origin nor the mechanism of their production has been established. The subject has been discussed separately for electrons[2,3] and for protons[59]. More to the point, there has been both analysis and observation of high energy ions in connection with interplanetary shocks, and this subject is currently under investigation[4,86,87]. It has yet to be learned whether microscopic plasma physical processes play a direct role in the energization of particles by collisionless shocks or whether they are involved only indirectly in formation of the shocks themselves. If microinstabilities are directly responsible, then the prospect opens that not only shocks, but any plasma interaction region where these phenomena can occur will become more accessible to analysis based on spacecraft investigations.

DISCUSSION

Recent Status of Theory. The experimentally-observed presence of electrostatic turbulence which clearly correlates with the more macroscopic MHD structures in the bow shock has resulted in significant theoretical activity, especially in searches for electrostatic microinstabilities that could be driven by currents of one kind or another.

The outstanding unanswered theoretical question concerning dissipative mechanisms in collisionless shocks is simply: what randomizes the protons across the shock?

At first, it was thought that diamagnetic electron drift currents

$$\underset{\sim}{J}_D = ne\underset{\sim}{V}_D = \frac{c}{4\pi} \nabla \times B \sim \frac{\Delta B}{\Delta S}$$

setting up a jump in field ΔB over a shock thickness ΔS would produce Langmuir, Buneman or ion-acoustic turbulence, and hence anomalous resistivity, and that ion-acoustic turbulence would cause the observed ion heating. However, quasilinear theory predicted only electron heating in such turbulence, and even though the formula has been successful at estimating magnetic ramp, i.e., drift region, thicknesses [66,51], doubt has been cast upon this as the main dissipative mechanism (ion randomizer).

Furthermore, observations of OGO-5 data on bow shock crossings revealed shock structures in which electrostatic turbulence appeared in many magnetic gradients, without randomization of the proton flow.

A summary of all the known attempts to resolve this theoretical question by postulation of electrostatic or electromagnetic instabilities and consequent turbulence up to 1973 was given by Greenstadt and Fredricks[46]. Their

conclusion was that electron heating in bow shock precursors by electrostatic microinstabilities is qualitatively rather well understood, while the ion heating mechanism is not. The most reasonable conjecture they implied was that ion-ion sound instabilities are responsible for the observed rapid ion heating across most bow shocks.

This latter conjecture has been put on a stronger basis in work by Galeev[38], who reviews again the possible candidates. Galeev proposes that counterstreaming ions produce turbulence due to instabilities with frequencies near the lower hybrid resonance. However, he points out that whether this mechanism is the correct one depends on verification by direct measurement on future spacecraft missions. Furthermore, as amply documented in earlier sections of this review, there are obvious differences in the microscopic particle dynamics associated with the several morphological categories (perpendicular, quasi-perpendicular, quasi-etc.) of bow shock structures, so that exact identification of the microinstability responsible for the dissipative process in each case is not an easy task.

It should be pointed out that measurements of lower hybrid resonance turbulence in the earth's bow shock structures have not been made in the past due to lack of instrumentation to cover the extremely low frequency electrostatic spectrum (f_{LHR} ~ 10-20 Hz). It is hoped that this deficiency will be removed by proper choice of instrumentation on such missions as the ISEE spacecraft in the future.

Morse[64] has suggested an alternative model for proton heating in the earth's bow shock. He starts from the premise that plasma fluctuations

within the bow shock structure of sufficient amplitude have not yet been observed. Morse then puts forth the hypothesis that the motion of ions in the macroscopic field fluctuations in the shock and its downstream region is sufficient to explain the observed broadening (and decrease of the average velocity) of the proton velocity distribution behind bow shocks, without introducing subshocks or wave instabilities as ion heating mechanisms. This idea is reminiscent of the ion "orbit-crossing" mechanism for ion randomization discussed by Auer et al.[5] and Morawetz[62,63].

Shock Parameters in the Solar System. There need be little doubt that solar system plasma offers an arena in which shock investigations can proceed and theoretical estimates can be tested experimentally. The values of the principal upstream plasma parameters in which shocks may be found in the solar system between Mercury and Jupiter is displayed in Figure 6. In 6(a), parameters β, C_{MS}, and M_{MS} and the nominal quasi-perpendicular unit of thickness c/ω_{pi} are shown vs solar distance in astronomical units (AU) for a typical solar wind speed of 400 km/sec, taken as constant at all distances in the solar equator. The parameters have been computed using the relations $N = 7/r^2$, $T_p = 7 \times 10^4/r$, $T_e = 1.5 \times 10^5/r^{1.5}$, and the Parker model for the spherical field components: $B_r = 5\sqrt{2}/2r^2$, $B_\phi = 5\sqrt{2}/2r$, $B_\theta = 0$. The numbers typify the solar wind at 1 AU. Although some question arose for a time about the radial dependence of B_ϕ[12,83], all the dependences used here have now been documented[9,54,70,75].

The long, solid and dotted curves at top and bottom of 6(a) represent parameters of the solar wind independent of whether any shocks exist there

or not. The Mach numbers, in contrast, depend on some assumption regarding shock and solar wind velocities. The long dashed curve shows the magnetosonic Mach number that would apply to a planetary, or cometary, shock stationary in a 400 Km/sec solar wind at any r, but the circled points mark only the M-values at the permanent planets, identified by their initials just above the horizontal coordinate line. The Q-turbulent/turbulent designation under the line denotes the category of planetary shock structure defined by the M_{MS} and β combinations in the corresponding regions. The two shorter curves illustrate the Mach numbers of two solar flare shocks propagating into the assumed 400 Km/sec wind. They terminate at the lower end, of course, where $M_{MS} = 1$. The declining velocity profiles from which the curves were computed were taken from two examples (Figures 15,16) of Dryer[18]. Such shocks would be turbulent above $M_{MS} \approx 3$, quasi-turbulent below.

The dotted curve in 6(a) accounts for the third source of shocks in the solar wind, namely, stream-stream interaction. Fast streams overtaking slower ones in the solar wind can evolve into compressed structures containing sufficient density to serve as "bodies" coursing through the wind, capable of generating shock pairs, one forward and one reverse. Such fast-stream shocks have been predicted[74], inferred[39,53,96], and demonstrated[95]. The velocity differential between the fast plasma behind a stream's steepening front and the slower plasma ahead of it falls in the range 40 to 120 Km/sec and averages about 60 Km/sec[53], so the behavior of the magnetosonic velocity C_{MS} implies that formation of a forward shock where $\Delta V > C_{MS}$ (i.e., $M_{MS} > 1$) is marginal for a 400 Km/sec background wind, but generally

most probable where C_{MS} drops below 60 Km/sec, i.e., where r > 1.2 AU. Clearly, the Mach numbers of such shocks are unlikely to be supercritical, so the classification of almost all of them would be laminar out to large r-distances, where $\beta << 1$ would make them quasi-laminar, or quasi-turbulent nearer 1 AU where $\beta \approx 1$. They would be expected to be well defined and of relatively simple structure when observed.

Figure 6(b) plots two more parameters determined by the IMF: its longitude angle $\phi_B = \arctan (B_\phi/B_r)$, for the sense away from the sun, and foreshock boundary angle θ_{XF} $\arctan p \sin \theta_{XB}/(p \cos \theta_{XB} - 1)$[43], as defined in Figure 5(b). The θ_{XF} curve in 6(b) is dashed, since it really applies only at the planets, indicated by the circled points. The plotted values were obtained by using p = 2 for the entire range of r[15]. Note that for purposes of illustration, the sense of B has been reversed in 6(c), so $\theta_{XB} = \phi_{XB} + 180°$. The importance of θ_{XB} in 6(b) lies in its implications regarding the extent and location of the average foreshock at each planet. For the first three planets, the foreshock literally reaches out in front of its bowshock, since its forward boundary makes an acute angle with the X-axis. At Mars and Jupiter, however, and by inference, beyond them, the foreshocks do not reach sunward of the subsolar points of their bow shocks, since $\theta_{XF} < 90°$, but they may occupy a considerable region to the sides of the shock flanks. The trend in θ_{XF} also implies that, given the usual fluctuations of the IMF direction, there is a fair probability that the foreshocks of the inner planets will occupy their entire upstream regions a substantial fraction of the time, i.e., when θ_{XB} slews toward 0°. Equivalently the sunward face of Mercury's shock should be expected to exhibit quasi-parallel structure frequently.

One paramount feature of Figure 6 has been saved for final emphasis. At the right of 6(a) and (b), three vertical flags represent the ranges of β, M_{MS}, and ϕ_B that have been recorded by spacecraft at 1 AU. According to these, the bow shock system of the earth alone can occur in any combination of parameters, thus providing a rich table from which to select samples for the study of shock structure, both in transient and steady-state conditions. In reality, not all combinations of the structural parameters are equally probable or equally observed. High β, low M (quasi-turbulent) conditions are unusual for the earth's shock at 1 AU and would presumably be much more probable at Mercury. Laminar cases, on the other hand, are considerably more common than the curves in 6(a) would suggest. On balance, detailed study of the earth's shock system should, with patience, produce a full display of the structural panoply found among the planets.

A Contemporary view of the multilocal approach to observations by spacecraft.

The constellation of sensors needed for diligent examination of plasma shock systems in space is suggested by the juxtaposed data plots of Figures 7 and 8. The figures show an overall view, at low resolution, of a sequence of encounters with interplanetary and bow shocks during two days of repeated crossings by both IMP 7 and 8 on opposite sides of the magnetosphere. We call attention to certain highlights of these plots that illustrate major features of the most significant events.

The formats of both figures are identical. The four upper panels are from IMP-8, the five lower from IMP-7. The IMP-8 panels contain, beginning

at the top, one channel of the Iowa electric wave data, the ambient magnetic field magnitude B, and the ambient field's longitude ϕ_B and latitude λ_B in solar ecliptic coordinates. Field measurements were made by the Goddard instrument. The inserts in the B-panel show the relative positions in the ecliptic of the satellites and the field's projected orientation, together with nominal magnetopause and shock cross sections, during the time intervals where the inserts are placed. Details will be described below.

The IMP-7 panels, continuing toward the bottom, contain views in four sectors, in the ecliptic, of low energy protons detected by Iowa's LEPEDEA experiment. The apex of each caret at left indicates the direction in which protons must be flowing to be detected in the corresponding sector. These directions coincide, counterclockwise, with the axes of the inserts in the B panel. The last panel at the bottom represents the average energy density from a wide channel of the TRW electric wave experiment on IMP-7. Note that the narrower Iowa channel on IMP-8, top panel, falls at about the center of the TRW channel. Both are chosen as monitors of local ion acoustic noise, which predominates in and around the bow shock. Descriptions of the various instruments can be found in references[32,81,89].

Our principal interest in Figure 7 is the foreshock, but we note the interplanetary shock of 1530, whose jump in B is accompanied by a sudden increase in average V_{SW}, a spread (thermalization) of its distribution (uppermost LEPEDEA panel), and enhanced plasma wave noise at both spacecraft. The triangle at bottom denotes the sudden commencement of a magnetic storm at the earth's surface.

The apparent discontinuity in B at left, around 0115, is of unidentified character and will not be discussed.

What Figure 7 shows very clearly is the control of the foreshock by the IMF orientation in three distinct intervals. The key factor was ϕ_B, since λ_B remained for the most part within 30° of the ecliptic.

First interval. Early in the day, until 0800, the IMF projection was directed from sunward to slightly west of the sun-earth line, i.e., $\phi_B \approx 330\text{-}360°$, as depicted by the small arrows in the first insert. The wavy line F in the same insert is the forwardmost proton-foreshock boundary that applied during this early interval. IMP-8 was obviously in the foreshock then, and we see that electric wave noise was enhanced (top panel) and that B oscillated as if upstream waves were present.

Second interval. At about 0800 the IMF rotated into the first quadrant, where it remained for several hours with ϕ_B between 0 and 45° (second insert) and then gradually advanced past 90°. The result at 0800 was to shut IMP-8 out of, and place IMP-7 within, the proton foreshock. We see that at IMP-8 the enhanced electric noise subsided and B became very quiet, while at IMP-7, return protons made their appearance. These are represented by the dark green traces between 10^3 and 10^4 eV in the second proton panel, indicating particles of those energies coming to IMP-7 from the shock to the right of it in the second insert. Shortly before 1600 UT, and slightly after the interplanetary shock of 1530, when ϕ_B had increased to about 70° (and there was also a swing away from the ecliptic in λ_B), as pictured by

the dotted arrow in the second insert, the foreshock boundary (dotted wavy line) passed behind IMP-7 and the return protons disappeared. They reappeared briefly at about 1730 coincident with a slight shift for a few minutes in IMF direction.

Third interval. After about 1745 electric wave noise became enhanced again at IMP-8 and remained so until nearly the end of the day. The field-shock geometry for this interval is shown in the third insert, where it is clearly seen that IMP-8 was outside the expected foreshock boundary. In this case the field projection was actually opposite the small arrows, which have been drawn for visual clarity and to emphasize the direction of possible particle velocities toward the satellite. It cannot now be stated with certainty whether the enlarged E-field amplitudes at IMP-8 were related to the bow shock or to changed interplanetary conditions, but we observe that the wavy foreshock boundary refers only to the region containing slow protons with $u_{\parallel} \approx 2\ V_{SW}$ and their associated ULF upstream waves. Faster protons, and electrons, could have been streaming more closely along B to IMP-8 from the bow shock, causing the local VLF wave noise. This possibility is favored by what we know of return electron behavior and by noting that proton velocities corresponding to the high energy end (10^4 eV) of the green return proton traces earlier at IMP-7 were adequate to convey ions to IMP-8 outside the depicted foreshock boundary in the third insert. This possibility is also compatible with the electric wave record at IMP-7, in the bottom panel. The noise level never seems to have dropped to a sustained background, or quiet, level during the 4th, even though we know IMP-7 was sometimes outside the (slow proton) foreshock boundary, as drawn in the

first insert. However, even early in the day, electrons and high-speed protons could have been reaching IMP-7 from the shock. This is the meaning of the dashed lines in the first insert, and IMP-7's location may have kept it on IMF lines intersecting the shock the entire day.

It remains only to note the short entry and exit of the bow shock itself by IMP-7 just before 2300, as determined from the thermalization and deflection of protons and the sharp peak in plasma wave noise.

We proceed to July 5th, in Figure 8. For this day, we cannot discuss the foreshock within the scope of the present report because the frequent, wide excursions of λ_B from 0° rendered ecliptic representations wholly inadequate. Of greater interest is the sequence of shock encounters which illustrate the wealth of diverse events and conditions that can occur in a single day in space.

At the outset, IMP-7 encountered the bow shock and was enclosed by the magnetosheath for about 30 minutes between 0100 and 0130. The high B that prevailed at IMP-8 at the start of the day suggests that there was a low Mach number in the solar wind that brought the bow shock out to IMP-7, for when the field dropped later, IMP-7 remained in the solar wind several hours until it saw the shock closer to its nominal position at 1315.

The insert in the IMP-8 B-panel shows the wide range of ϕ_B during the hour and a half or so before the 1315 crossing. Of special interest is the swing of the IMF to the fourth quadrant at about 1245, the result of which was to make B essentially tangent to the bow shock in time to determine a nearly perpendicular crossing geometry at 1315. After that, λ_B was close

enough to -90° to keep the shock geometry nearly perpendicular while IMP-7 was in the sheath behind it until about 1615 when the solar wind underwent a sudden change, best indicated in the figure by the discontinuity in B and λ_B. IMP-7 emerged gradually from what appears to have been a quasi-parallel shock and then, at 1715, began a two-hour series of irregular and noisy shock measurements the nature of which our information is insufficient to define. It is clear, however, that the local magnetic geometry was quasi-parallel, and we see that (a) the average velocity of the solar wind was little changed, (b) thermalization was taking place but less intensely than behind the 1315 crossings, and (c) higher energy particles, up to about 10^4 eV were being deflected into all sectors; this did not take place behind the 1315 crossing.

At 1930 a sudden-commencement interplanetary shock reached the vicinity of the earth and drove the bow shock inward, leaving IMP-7 again in the solar wind. The satellite reentered the bow shock at 2210 under conditions varying between quasi-perpendicular and quasi-parallel and remained in the magnetosheath through the end of the 5th.

The IMP-8 data of 5 July featured electric wave activity which was low, with some sporadic enhancement through about 1630, and then increased for the remainder of the day. The cause of the noise amplification cannot be unambiguously identified here. It is always necessary to be cautious about differentiating between interplanetary and foreshock effects. In this case, a very careful geometric analysis of the possible connecting field geometry to the bow shock would be necessary to determine the plausibility of return

electrons or high-speed protons, as factors in producing the amplified wave noise, so we cannot state now whether the plasma wave pattern on the 5th is a continuation of the same processes that appeared to be operating on the 4th.

The content and description of Figures 7 and 8 illustrate that two satellites and four instruments are far from superfluous in reconstructing a sequence of physical processes surrounding the bow shock. Indeed, the absences of IMF data at IMP-7 and LEPEDEA data at IMP-8 from the figures are keenly felt. The former were unavailable because the instrument had previously ceased operating; the latter was available but had not yet been obtained for this report. Either would have served to clarify some of the event identifications. From another point of view, however, the figures demonstrate how comprehensive instrumentation permits analysis to proceed, albeit cautiously, even though nominally essential measurements are missing.

The examples of Figures 7 and 8 also show the way in which space plasma behavior can be examined on a vast scale with devices of comparatively infinitesimal dimensions and negligible influence on the environment they sample. Finally, the figures expose the data techniques of the present and command the methodology of the future. The four plasma panels, for example, were selected from a still wider display of the solar wind's properties developed at Iowa to show the properties of the multidimensional solar wind in metric and velocity space in a manner rapidly understandable to the data analyst. The verbal description of the events of 4-5 July, how-

ever, underscores the upcoming need for equally routine display of the relationship among spacecraft locus, IMF orientation, and bow shock geometry in order to define the delicate control of the foreshock's components and the bow shock's structure by the IMF.

RECOMMENDATIONS

Future study of collisionless shock phenomena will be greatly aided by spacecraft programs designed for both geocentric and interplanetary orbits, if appropriate instrumentation is carried. In addition to conventional field and particle devices, high resolution, omnidirectional plasma detectors will be of great importance.

The earth's bow shock will be the principal source of new, detailed measurement covering almost the whole range of solar wind parameter combinations. Details will be needed in both spatial and temporal dimensions. Table 2 displays the major categories of shock phenomena, the requirements for their investigation, the sources for meeting these requirements, and the status of the sources at present. Multiple listings under a single letter designation mean that coordinated, grouped requirements must be satisfied as a unit. The table indicates that, for the most part, suitable spacecraft instrumentation has been or is scheduled to be available, but that the status of software support and non-mission activities is in general undersupported.

The most important new recommendation outside the table is that a cluster of four closely-spaced vehicles be planned, giving three-dimensional spatial resolution with gyroradial-order separation, omnidirectional plasma

particle detection capability, high sampling rates, and an accompanying solar wind monitor in the sunward plasma, far upstream.

Much of future progress, however, is rooted in the past. Figures 7 and 8 of the foregoing section illustrate, scandalously, the first published example of double measurements in the foreshock with the instrumentation shown. A commitment is urgently needed to recapture the investment in plasma physics lying idly in data warehouses throughout the nation. We recommend such a commitment.

Further, it is of great importance that theoretical modeling of shock phenomena, most of which is years old, be encouraged to catch up with the rich store of observational information already published, let alone awaiting disclosure in data freightyards and tape libraries. One of the most promising avenues is numerical simulation, and we recommend dedication of at least part of the effort of one or more large computer facilities to the digital-modeling of shocks in space.

Finally, we note that comprehension of physical phenomena are best completed when we can reproduce and manipulate them in the laboratory, and we recommend wholeheartedly that promising experimental work with tenuous, if not collisionless, plasma shocks in laboratory apparatus be extended as far as possible into the parameter domains found in space.

ACKNOWLEDGMENTS

We express with pleasure the help of K. L. Ackerson, K. W. Behannon, R. A. Chevallier, L. A. Frank, A. A. Galeev, J. R. Jokipii, R. P. Lepping, C. T. Russell, and F. L. Scarf in supplying helpful discussion and/or advance copies of unpublished data and manuscripts used to prepare this report.

Table 1. Macroscopic Features of the Bow Shock by Parameter Class

Parameter Values	Plasma Conditions	Name of Structure	QUASI-PERPENDICULAR Features	Ref's	QUASI-PARALLEL Features	Ref's
$\beta \ll 1$, $M \lesssim 3$	Cold plasma, low Mach number	LAMINAR	Clean field jump, sometimes with damped periodic waves, no turbulence.	50	Multigradient field transition with embedded, nearly-periodic wavetrains.	(43)
			Relatively little proton temperature jump ($T_{p2}/T_{p1} \approx 2$), maxwellian downstream distribution	30	Upstream waves with strong periodic component, tens of seconds period, unknown plasma distributions.	
$\beta < 1$, $M \gtrsim 3$	Cool plasma, high Mach number	QUASI-LAMINAR	Clean field jump, downstream, nearly-periodic waves, little turbulence	(43)	Multigradient field transition at least 2 R_E thick, with large amplitude pulses; extensive foreshock.	49
			Appreciable proton temperature jump, bimodal distribution, nonmaxwellian high energy tail downstream.		Little change in solar wind streaming velocity; some heating with resulting non-maxwellian proton distributions differing from either solar wind or magnetosheath forms.	
$\beta \approx 1$, $M \lesssim 3$	Warm plasma, low Mach number	QUASI-TURBULENT	Clean field jump, small-scale turbulence.	(43)	No known example.	
			Little proton temperature rise, maxwellian downstream distribution.			
$\beta \approx 1$ $M \gtrsim 3$	Warm plasma, high Mach number	TURBULENT	Irregular field fluctuations obscuring definite average field jump.	29	Multigradient, irregular field transition with large magnitude excursions, upstream waves.	(43)
			Bimodal or multimodal proton distributions, nonmaxwellian downstream		Unknown plasma distributions, but maxwellian deep in magnetosheath.	
$\beta \gg 1$, $M > 3$	Hot, high velocity plasma	HIGH-BETA	Irregular field fluctuations of extremely high peak magnitude, possibly lowering β locally.	31		
			Extensive precursor region with appreciable effect on approaching flow; proton distributions unknown.			

Table 2. Overview of Shock Study Expectations

Phenomenon	Requirements	Sources	Status*
1. Ion thermalization processes in q-perpendicular structure.	A. High resolution observations (< 10-sec cycles) a. Multidirectional plasma electron and ion spectra b. Full-band plasma & EM wave spectra c. Tenth-second magnetometry d. Multidirectional suprathermal particle spectra	Suitably-instrumented new spacecraft	S (ISEE-A,B)
	e. Shock velocity monitor	Subsatellite at 1 R_e separation	S (ISEE-A,B)
	f. Upstream parameter monitor	Separate subsatellite	S (ISEE-A,B)
	B. Quantitative theoretical modeling	a. Theoretical analysis b. Numerical simulation	M N
	C. Experimental modeling	Laboratory simulation	N
	D. Coordinated reduction & analysis	Intimate experimental & theoretical cooperation	P (presently voluntary & sporadic)
2. Foreshock composition & structure	A. Low resolution observations (10-60 sec cycles) a. Multidirectional plasma electron & ion spectra b. Full-band plasma & EM wave spectra c. 1-sec magnetometry d. Multidirectional superthermal ion spectra	Suitably instrumented old & new spacecraft	S (IMP's H,J; ISEE-A,B; ALSEP, et al., data-bank opportunities largely unexplored)
	e. Multilocal observations	Widely-separated spacecraft	M (opportunities largely unexplored)

. . . . . .

*Status Code: S Satisfactory equipment or procedure exists, is scheduled for use, and is supported.
P Partial equipment or procedure has been supported in past or is available, but partially supported.
M Minimal or inadequate equipment, procedure, or support is available.
N No equipment, procedure, or support is available.

Phenomenon	Requirements	Sources	Status*
	f. Upstream parameter monitor	Widely-separated solar wind spacecraft	M (opportunities largely unexplored)
	g. Coordinated reduction & analysis	Intimate experimental & theoretical cooperation	M
	B. Implicit analytic support	Automatic conversion of solar wind parameters to shock & foreshock location	N
	C. Theoretical modeling	Theoretical analysis	P
	D. General data analysis	Existing data banks, fortuitous conjunctions of existing spacecraft	M
3. Comprehensive plasma processes in q-parallel structure	A. Medium resolution observations (1-100 sec cycles) a. Multidirectional plasma electron & ion spectra b. Full-band plasma & EM wave spectra c. .1-sec magnetometery d. Suprathermal particle spectra	Suitably instrumented old and new spacecraft	S
	e. Multilocal wave propagation observations	Satellite cluster	N
	f. Upstream parameter monitor	Widely-separated solar wind spacecraft	S (ISEE-C; IMP's et al.
	g. Coordinated reduction & analysis	Intimate experimental & theoretical cooperation	M
	B. Implicit analytic support	Automatic conversion of solar wind parameters to nominal shock location	N

Phenomenon	Requirements	Sources	Status
	C. Theoretical modeling	a. Theoretical analysis b. Numerical simulation	N
	D. General data analysis	Existing data banks; fortuitous conjunctions of existing spacecraft	M
	E. Experimental modeling	Laboratory simulation	N
4. Magnetosheath composition & behavior	A. Low resolution observations (10-200 sec cycles) a. Plasma electron & ion spectra b. Plasma & EM wave spectra c. 1-sec magnetometry d. Multidirectional suprathermal ion spectra	Suitably instrumented old or new spacecraft	S (several suitable spacecraft)
	e. Multilocal observations	Widely-separated spacecraft	M (unexplored)
	f. Upstream parameter monitor	Widely-separated spacecraft	M (unexplored)
	g. Coordinated reduction & analysis	Intimate experiment cooperation	M
	B. Implicit analytic support	Automatic conversion of solar wind parameters to shock & magnetopause locations	N
	C. Theoretical modeling	a. Theoretical analysis b. Numerical simulation	N
	D. General data analysis	Existing data banks	M
	E. Experimental modeling	Laboratory simulation	N

Phenomenon	Requirements	Sources	Status
5. Interplanetary shock processes	A. High resolution observations (1-10 sec cycles) a. Solar wind ion & Electron spectra b. Plasma & EM wave spectra c. .1 second magnetometry d. Multidirectional suprathermal ion spectra	Suitably instrumented old & new spacecraft	S (IMP-7; ISEE-A,B)
	e. Shock velocity monitor	Subsatellite	S (ISEE-A,B)
	f. Shock uniformity monitor	Widely-separated space probes	P (Pioneer; Helios; unexplored)
	g. Coordinated reduction & analysis	Intimate experimental cooperation	P
	B. Analytic support	Computational shock speed & shape analysis	S
	C. Theoretical modeling	Theoretical analysis of ion acceleration	S

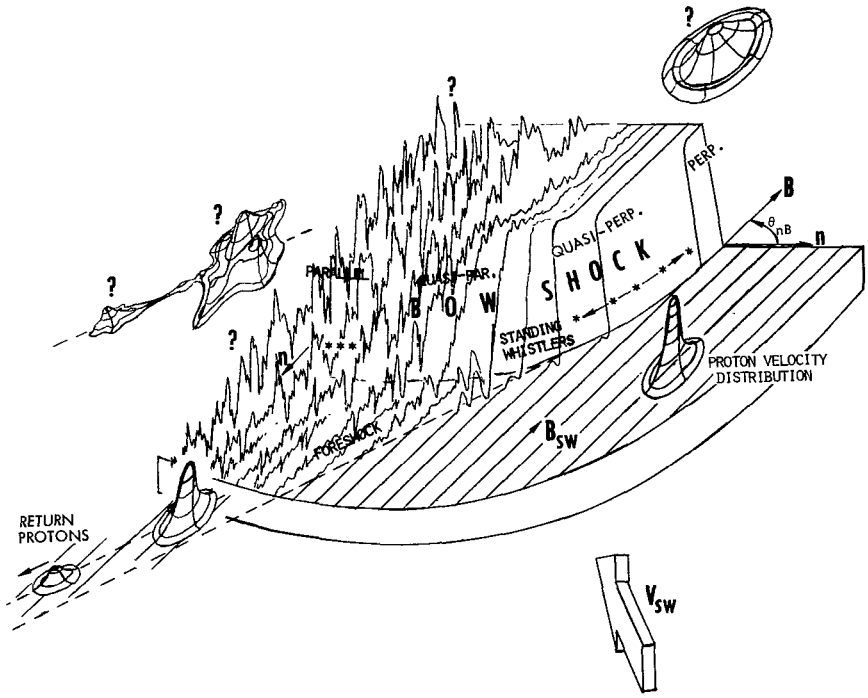


FIGURE 1 Conceptualization of collisionless shock macrophenomenology as manifested in the earth's curved bow shock. Unshocked interplanetary field direction $\underset{\sim}{B}_{SW}$ is indicated on the foreground field "platform." Field *magnitude* is plotted vertically; field *direction* would be deflected somewhat in the $\underset{\sim}{B}_{SW}$ plane by the quasi-perpendicular shock, but would share the agitation of the magnitude in the quasi-parallel shock in all components. The superposed three-dimensional sketches represent solar wind proton thermal properties as number distributions in velocity space.

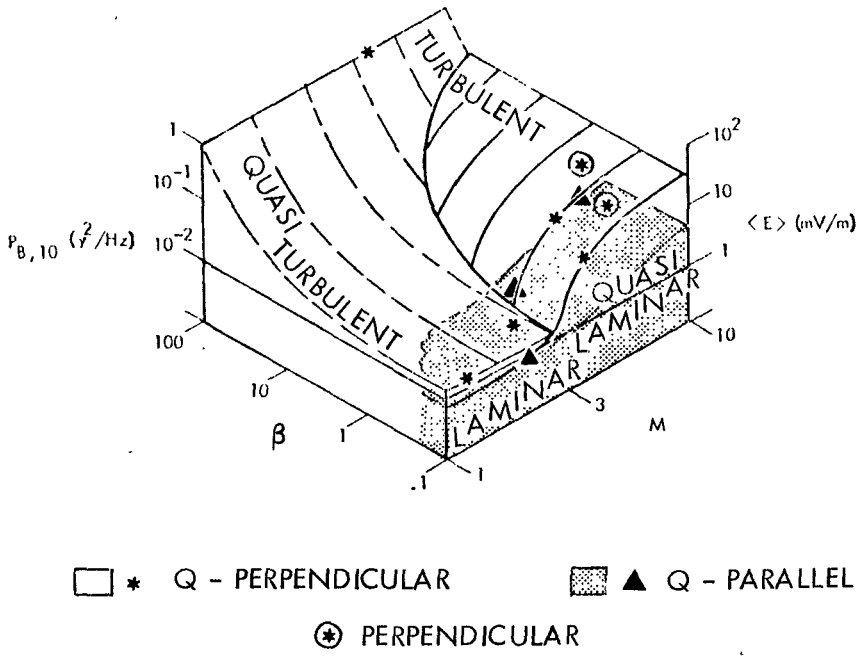


FIGURE 2 Conceptualization of bow shock microphenomenology as represented by electromagnetic noise power density and plasma electric wave noise amplitude in the various shock structures determined by upstream plasma parameters β and M. Individual symbols show the parameter combinations in which details have been documented in specific cases. The clear contours apply to quasi-perpendicular, the shaded (and truncated) contours to quasi-parallel, geometry.

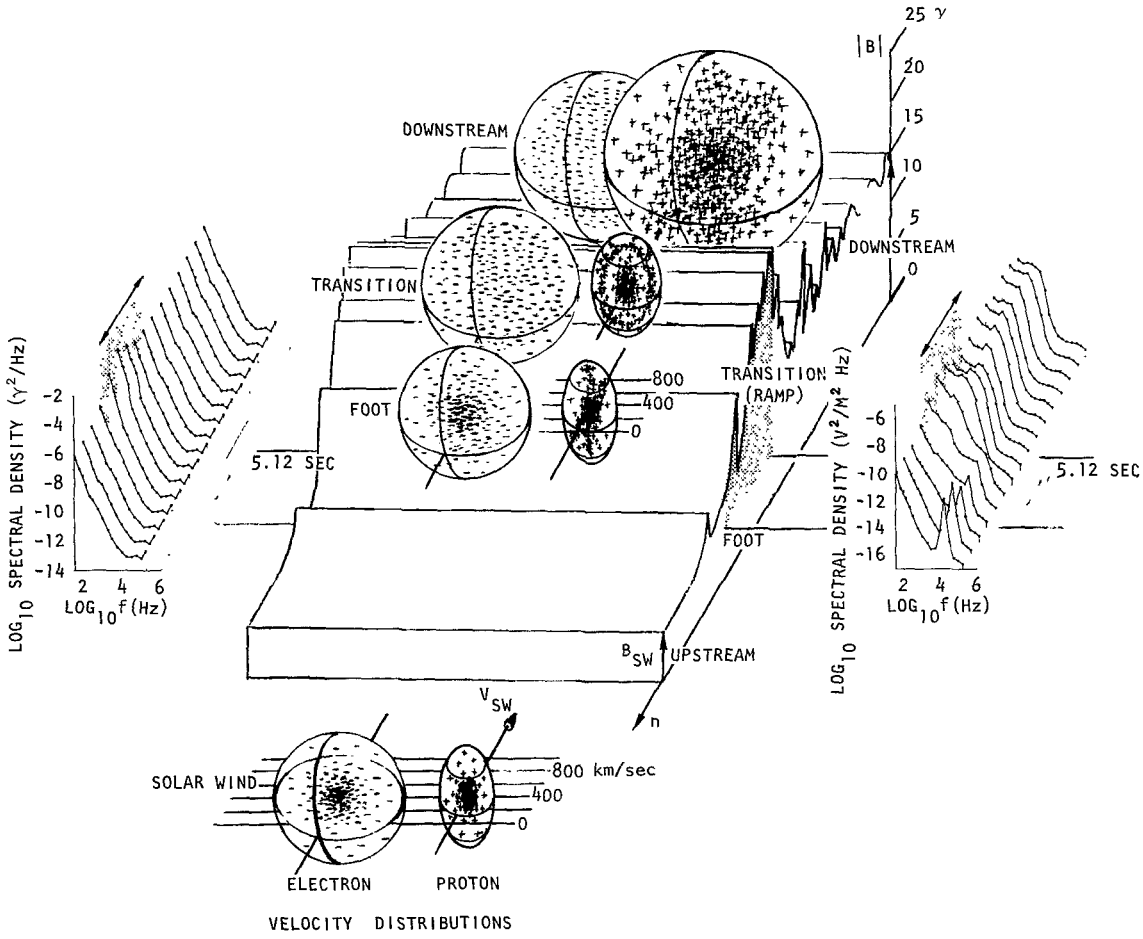


FIGURE 3 Sequences of (electro) magnetic wave, left, electric plasma wave, right, and particle events, center, relative to typical quasi-perpendicular, supercritical magnetic profile plotted in the central block. Electron and proton distributions in velocity space are shown in the cold, fast solar wind, foreground, in the foot of the magnetic structure where the electrons are initially heated and the protons retarded and partially scattered, in the center of the principal magnetic gradient, or ramp (shaded) where the electrons are fully scattered and the protons partially heated and scattered to form a bimodal distribution, and finally behind the magnetic front, where electrons and protons are both found heated and scattered into nonmaxwellian distributions.

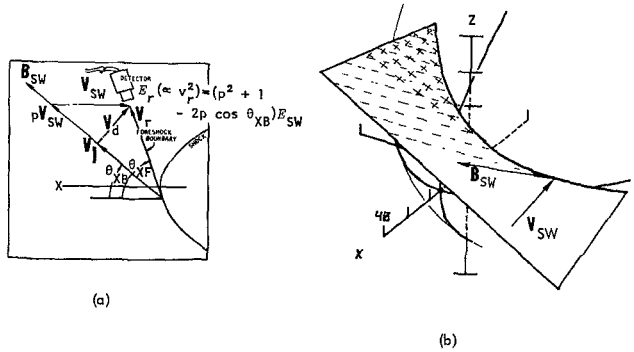


FIGURE 4 Geometry of return proton detection. (a) Particle guiding-centers in the plane of $\underset{\sim}{V}_{SW}$ (ie., X) and $\underset{\sim}{B}_{SW}$ advance along $\underset{\sim}{B}_{SW}$ at speed $V_{\parallel}$ while drifting perpendicular to $\underset{\sim}{B}_{SW}$ at speed V_d, but resulting velocity $\underset{\sim}{V}_r$ is treated as if $\underset{\sim}{V}_r = pV_{SW}\underset{\sim}{B}_{SW}/B_{SW} + \underset{\sim}{V}_{SW}$. (b) A cross section of the foreshock is formed in each $\underset{\sim}{B}_{SW} - \underset{\sim}{V}_{SW}$ plane by electrons and protons streaming away from the shock according to the diagram in (a).

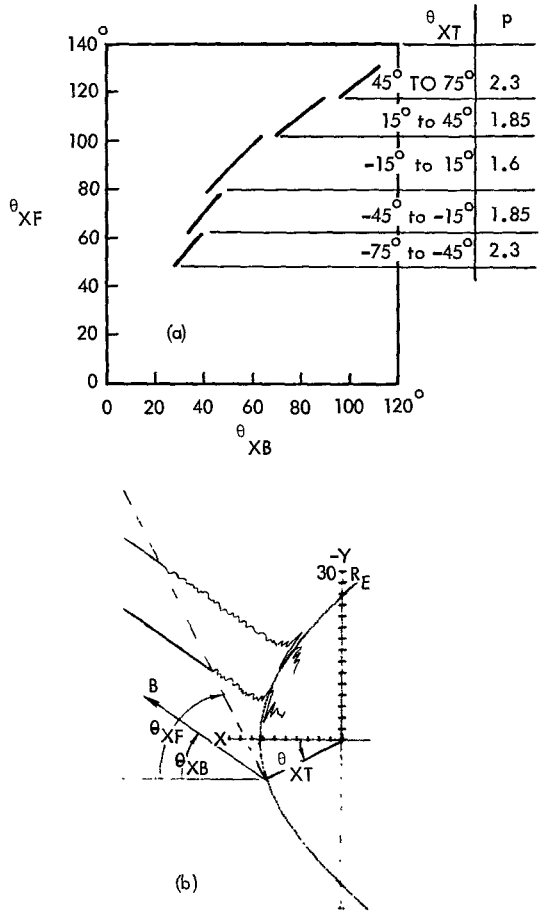


FIGURE 5 (a) Dependence of the foreshock's proton wave boundary angle θ_{XF} on field angle θ_{XB} for different positions along the shock, indicated by θ_{XT}. The quantities are defined in relation to each other in (b).

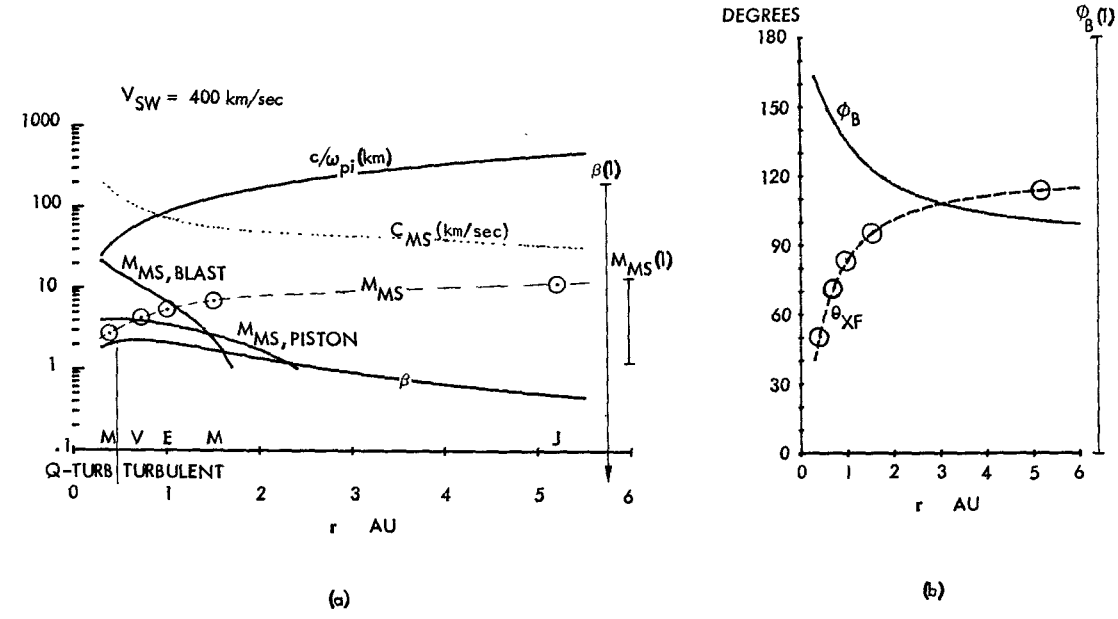


FIGURE 6 (a) Typical parameters of the solar wind *vs* heliocentric distance r from the sun to Jupiter; (b) Average field longitude θ_B and foreshock boundary angle θ_{XF} *vs* heliocentric distance.

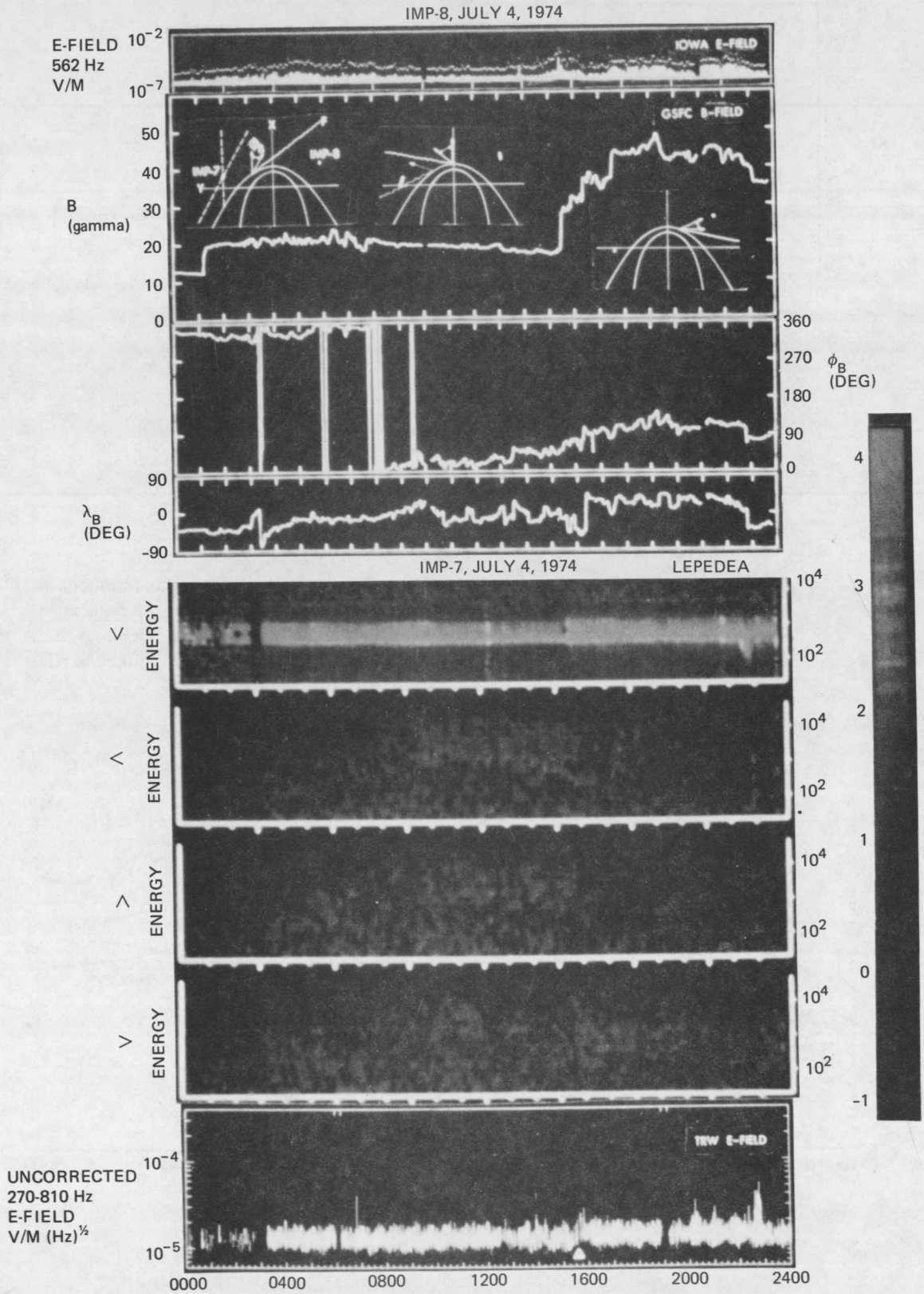


FIGURE 7 Simultaneous measurements by IMP's 7 and 8 on 4 July 1974. Positions of the satellites relative to the bow shock system are shown in the diagrams superposed on the second panel from the top.

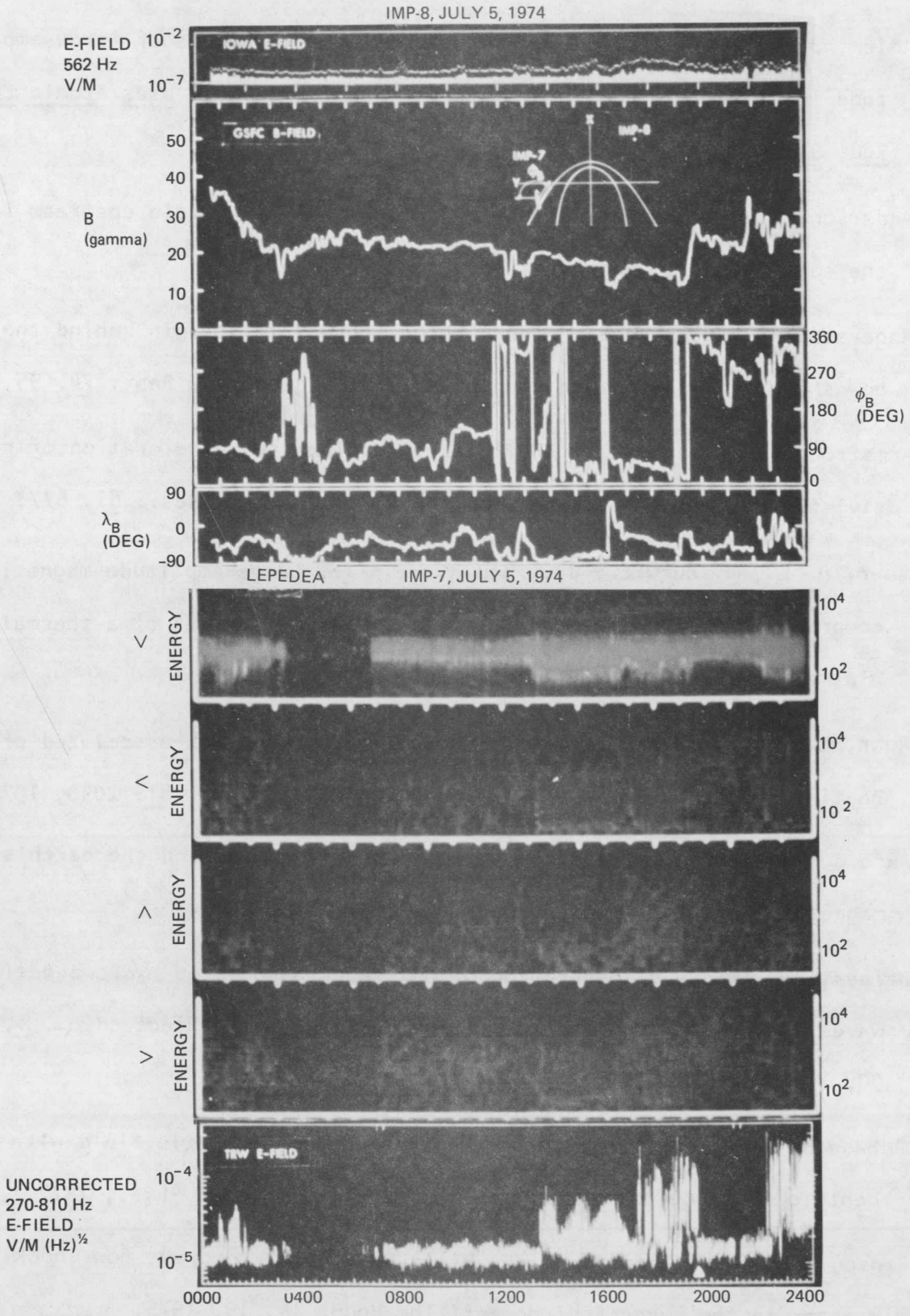


FIGURE 8 Same as Figure 7 for 5 July 1974.

REFERENCES

1. Adlam, J. H. and J. E. Allen, Hydromagnetic disturbances of large amplitude in a plasma, Proc. U.N. Interntl. Conf. Peaceful Uses Atomic Energy, 2nd, 31, 221, 1958.

2. Anderson, K. A., Energetic electrons of terrestrial origin upstream in the solar wind, J. Geophys. Res., 73, 2387, 1968.

3. Anderson, K. A., Energetic electrons of terrestrial origin behind the bow shock and upstream in the solar wind, J. Geophys. Res., 74, 95, 1969.

4. Armstrong, T. P. and S. M. Krimigis, Interplanetary acceleration of relativistic electrons observed with IMP 7, J. Geophys. Res., 81, 677, 1976.

5. Auer, P. L., H. Hurwitz, Jr., and R. W. Kilb, Large-amplitude magnetic compression of a collision-free plasma, 2, Development of a thermalized plasma, Phys. Fluids, 5, 298, 1962.

6. Auer, R.-D., H. Grünwaldt, and H. Rosenbauer, Bow-shock-associated proton heating in the upstream solar wind, J. Geophys. Res., 81, 2030, 1976.

7. Axford, W. I., The interaction between the solar wind and the earth's magnetosphere, J. Geophys. Res., 67, 3791, 1962.

8. Barnes, A., Theory of generation of bow-shock-associated hydromagnetic waves in the upstream interplanetary medium, Cosmic Electrodyn., 1, 90, 1970.

9. Behannon, K. W., Variation of the interplanetary magnetic field with heliocentric distance, submitted to Rev. Geophys. & Space Phys., 1977.

10. Benson, J., J. W. Freeman, H. K. Hills, and R. R. Vondrak, Bow Shock Protons in the Lunar Environment, The Moon, 14, 19, 1975.

11. Chevallier, R. A., The interaction of supernovae with the interstellar medium, Ann. Rev. of Astron. & Astrophys., 15, in press, 1977.

12. Coleman, P. J., Jr., E. J. Smith, L. Davis, Jr., and D. E. Jones, The radial dependence of the interplanetary magnetic field: 1.0-1.5 AU, J. Geophys. Res., 74, 2826, 1969.

13. Daughney, C. C., L. S. Holmes, and J. W. M. Paul, Measurement of spectrum of turbulence within a collisionless shock by collective scattering of light, Phys. Rev. Lett's., 25, 497, 1970.

14. De Hoffmann, F., and E. Teller, Magnetohydrodynamic shocks, Phys. Rev., 80, 692, 1950.

15. Diodato, L., E. W. Greenstadt, G. Moreno, and V. Formisano, A statistical study of the upstream wave boundary outside the earth's bow shock, J. Geophys. Res., 81, 199, 1976.

16. Diodato, L. and G. Moreno, A two-spacecraft study of the foreshock perturbations of the solar wind protons, CNR, Frascati, submitted to J. Geophys. Res

17. Dobrowolny, M. and V. Formisano, The structure of the earth's bow shock, Rev. Nuovo Cim., 3, 419, 1973.

18. Dryer, M., Interplanetary shock waves generated by solar flares, Space Sci. Rev., 15, 403, 1974.

19. Fairfield, D. H., Bow shock associated waves observed in the far upstream interplanetary medium, J. Geophys. Res., 74, 3541, 1969.

20. Fairfield, D. H., Whistler waves observed upstream from collisionless shocks, J. Geophys. Res., 79, 1368, 1974.

21. Fairfield, D. H. and K. W. Behannon, Bow shock and magnetosheath waves at Mercury, J. Geophys. Res., 81, 3897, 1976.

22. Fairfield, D. H. and W. C. Feldman, Standing waves at low Mach number laminar bow shocks, J. Geophys. Res., 80, 515, 1975.

23. Feldman, W. C., J. R. Asbridge, and S. J. Bame, Bow shock perturbation of the upstream solar wind proton component, J. Geophys. Res., 79, 2773, 1974.

24. Feldman, W. C., J. R. Asbridge, S. J. Bame, and M. D. Montgomery, Solar wind heat transport in the vicinity of the earth's bow shock, J. Geophys. Res., 78, 3697, 1973.

25. Fishman, F. J., A. R. Kantrowitz, and H. E. Petschek, Magnetohydrodynamic shock wave in a collision-free plasma, Rev's. Mod. Phys., 32, 959, 1960.

26. Formisano, V., Jupiter's bow shock structure, Lett. Nuovo Cim., 10, 182, 1974.

27. Formisano, V., The physics of the earth's collisionless shock wave, presented to the Third International Congress on Waves and Instabilities in Plasmas, Paris, June 1977.

28. Formisano, V., and E. Amata, Solar wind interaction with the earth's magnetic field. 4. Preshock perturbation of the solar wind, J. Geophys. Res., 81, 3907, 1976.

29. Formisano, V. and P. C. Hedgecock, On the structure of the turbulent bow shock, J. Geophys. Res., 78, 6522, 1973.

30. Formisano, V., G. Moreno, F. Palmiotto, and P. C. Hedgecock, Solar wind interaction with the earth's magnetic field. 1. Magnetosheath, J. Geophys. Res., 78, 3714, 1973.

31. Formisano, V., C. T. Russell, J. D. Means, E. W. Greenstadt, F. L. Scarf, and M. Neugebauer, Collisionless shock waves in space: A very high-beta structure, J. Geophys. Res., 80, 2030, 1975.

32. Frank, L. A., K. L. Ackerson, and R. P. Lepping, On hot tenuous plasmas, fireballs, and boundary layers in the earth's magnetotail, J. Geophys. Res., 81, 5859, 1976.

33. Fredricks, R. W., A model for generation of bow-shock-associated upstream waves, J. Geophys. Res., 80, 7, 1975.

34. Fredricks, R. W., G. M. Crook, C. F. Kennel, I. M. Green, F. L. Scarf, P. J. Coleman, Jr., and C. T. Russell, OGO 5 observations of electrostatic turbulence in bow shock magnetic structures, J. Geophys. Res., 75, 3751, 1970.

35. Fredricks, R. W., C. F. Kennel, F. L. Scarf, G. M. Crook, and I. M. Green, Detection of electric field turbulence in the earth's bow shock, Phys. Rev. Lett's., 21, 1761, 1968.

36. Fredricks, R. W., F. L. Scarf, and L. A. Frank, Nonthermal electrons and high-frequency waves in the upstream solar wind. 2. Analysis and interpretation, J. Geophys. Res., 76, 6691, 1971.

37. Fredricks, R. W., F. L. Scarf, and I. M. Green, Distributions of electron plasma oscillations upstream from the earth's bow shock, J. Geophys. Res., 77, 1300, 1972.

38. Galeev, A. A., Collisionless shocks, Physics of Solar Planetary Environment, ed. D. J. Williams, American Geophysical Union, 464, 1976.

39. Gosling, J. T., A. J. Hundhausen, and S. J. Bame, Solar wind stream evolution at large heliocentric distances: Experimental demonstration and the test of a model, J. Geophys. Res., 81, 2111, 1976.

40. Greenstadt, E. W., Dependence of shock structure at Venus and Mars on orientation of the interplanetary magnetic field, Cosmic Electrodyn., 1, 380, 1970.

41. Greenstadt, E. W., Binary index for assessing local bow shock obliquity, J. Geophys. Res., 77, 5467, 1972.

42. Greenstadt, E. W., Field-determined oscillations in the magnetosheath as possible source of medium-period, daytime micropulsations, Proc. of Conf. on Solar Terrestrial Relations, Univ. of Calgary, 515, 1972.

43. Greenstadt, E. W., Structure of the terrestrial bow shock, Solar Wind Three, ed. C. T. Russell, Institute of Geophysics & Planetary Physics, UCLA, 440, 1974.

44. Greenstadt, E. W., Phenomenology of the earth's bow shock system. A summary description, Magnetospheric Particles & Fields, ed. B. M. McCormac, Reidel Publishing Co., Dordrecht, Holland, 13, 1976.

45. Greenstadt, E. W., Energies of backstreaming protons in the foreshock, Geophys. Res. Lett's., 3, 553, 1976.

46. Greenstadt, E. W. and R. W. Fredricks, Plasma instability modes related to the earth's bow shock, Magnetospheric Physics, ed. B. M. McCormac, D. Reidel, Dordrecht, Holland, 355, 1974.

47. Greenstadt, E. W., I. M. Green, G. T. Inouye, D. S. Colburn, J. H. Binsack, and E. F. Lyon, Dual satellite observation of the earth's bow shock. 2. Field-aligned upstream waves, Cosmic Electrodyn., 1, 279, 1970.

48. Greenstadt, E. W. and J. V. Olson, Pc 3,4 activity and interplanetary field orientation, J. Geophys. Res., 81, 5911, 1976.

49. Greenstadt, E. W., C. T. Russell, V. Formisano, P. C. Hedgecock, F. L. Scarf, M. Neugebauer, and R. E. Holzer, Structure of a quasi-parallel, quasi-laminar bow shock, J. Geophys. Res., in press, 1976.

50. Greenstadt, E. W., C. T. Russell, F. L. Scarf, V. Formisano, and M. Neugebauer, Structure of the quasi-perpendicular, laminar bow shock, J. Geophys. Res., 80, 502, 1975.

51. Greenstadt, E. W., V. Formisano, C. T. Russell, M. Neugebauer, and F. L. Scarf, Ion acoustic stability analysis of the earth's bow shock, submitted to Geophys. Res. Lett's, 1977.

52. Holzer, T. E., Interaction between the solar wind and the interstellar medium, Exploration of the Outer Solar System, ed. E. W. Greenstadt, M. Dryer, and D. S. Intriligator, AIAA, New York, 21, 1976.

53. Hundhausen, A. J. and J. T. Gosling, Solar wind structure at large heliodentric distances: An interpretation of Pioneer 10 observations, J. Geophys. Res., 81, 1436, 1976.

54. Jokipii, J. R., Radial variation of solar wind parameters, Geophys. Res. Lett's., 3, 141, 1976.

55. Kellogg, P. J., Flow of plasma around the earth, J. Geophys. Res., 67, 3805, 1962

56. Kellogg, P. J., Solitary waves in cold, collisionless plasma, Phys. Fluids, 7, 1555, 1964.

57. Kennel, C. F., and F. V. Coroniti, Is Jupiter's magnetosphere like a pulsar's or earth's?, The Magnetosphere of Earth & Jupiter, ed. V. Formisano, D. Reidel, Dordrecht, Holland, 451, 1975.

58. Kennel, C. F. and R. Z. Sagdeev, Collisionless shock waves in high β plasmas, 1, J. Geophys. Res., 72, 3303, 1967.

59. Lin, R. P., C.-I. Meng, and K. A. Anderson, 30- to 100-keV protons upstream from the earth's bow shock, J. Geophys. Res., 79, 489, 1974.

60. Montgomery, D., Development of hydromagnetic shocks from large amplitude Alfvén waves, Phys. Rev. Lett's., 2, 36, 1959.

61. Montgomery, M. D., J. R. Asbridge, and S. J. Bame, Vela 4 plasma observations near the earth's bow shock, J. Geophys. Res., 75, 1217, 1970.

62. Morawetz, C. S., Magnetohydrodynamic shock structure without collisions, Phys. Fluids, 4, 988, 1961.

63. Morawetz, C. S., Modification for magnetohydrodynamic shock structure without collisions, Phys. Fluids, 5, 1447

64. Morse, D. L., A model for ion thermalization in the earth's bow shock, J. Geophys. Res., 81, 6126, 1976.

65. Morse, D. L., W. W. Destler, and P. L. Auer, Non-stationary behavior of collisionless shocks, Phys. Rev. Lett's., 28, 13, 1972.

66. Morse, D. L. and E. W. Greenstadt, Thickness of magnetic structures associated with the earth's bow shock, J. Geophys. Res., 81, 1791, 1976.

67. Ness, N. F., K. W. Behannon, and R. P. Lepping, The magnetic field of Mercury, 1., J. Geophys. Res., 80, 2708, 1975.

68. Ness, N. F., K. W. Behannon, and R. P. Lepping, Observations of Mercury's Magnetic Field, Icarus, 28, 479, 1976.

69. Ness, N. F., K. W. Behannon, R. P. Lepping, Y. C. Whang, and K. H. Schatten Magnetic field observations near Venus: Preliminary results from Mariner 10, Science, 183, 1301, 1974.

70. Neugebauer, M., Large-scale and solar-cycle variations of the solar wind, Space Sci. Rev., 17, 221, 1975.

71. Ogilvie, K. W., J. D. Scudder, R. E. Hartle, G. L. Siscoe, H. S. Bridge, A. J. Lazarus, J. R. Asbridge, S. J. Bame, and C. M. Yeates, Observations of Mercury encounter by the plasma science experiment on Mariner 10, Science, 185, 145, 1974.

72. Ogilvie, K. W., J. D. Scudder, V. M. Vasyliunas, R. E. Hartle, and G. L. Siscoe, Observations at the planet Mercury by the plasma electron experiment -- Mariner 10, Goddard Space Flight Center Report X-692-76-223, 1976.

73. Olson, J. V., R. E. Holzer, and E. J. Smith, High-frequency magnetic fluctuations associated with the earth's bow shock, J. Geophys. Res., 74, 4601, 1969.

74. Parker, E. N., Interplanetary dynamical processes, Interscience, New York, 1963.

75. Parker, G. D. and J. R. Jokipii, The spiral structure of the interplanetary magnetic field, Geophys. Res. Lett's., 3, 561, 1976.

76. Paul, J. W. M., Review of experimental studies of collisionless shocks propagating perpendicular to a magnetic field, ESRO Spec. Publ. 51, Frascati, Italy, 97, 1969.

77. Paul, J. W. M., G. C. Goldenbaum, A. Ioyoshi, L. S. Holmes, and R. A. Hardcastle, Measurement of electron temperatures produced by collisionless shock waves in magnetized plasma, Nature, 216, 363, 1967.

78. Paul, J. W. M., L. S. Holmes, M. J. Parkinson, and J. Sheffield, Experimental observations on the structure of collisionless shock waves in a magnetized plasma, Nature, 208, 133, 1965.

79. Roberts, P. H., Stellar Winds, Solar Wind Three, ed C. T. Russell, Univ. of Calif., Los Angeles, 231, 1974.

80. Robson, A. E., Experiments on oblique shock waves, ESRO Spec. Publ. 51, Frascati, Italy, 159, 1969.

81. Rodriguez, P. and D. A. Gurnett, Electrostatic and electromagnetic turbulence associated with the earth's bow shock, J. Geophys. Res., 80, 19, 1975.

82. Rodriguez, P. and D. A. Gurnett, Correlation of bow shock plasma wave turbulence with solar wind parameters, J. Geophys. Res., 81, 2871, 1976.

83. Rosenberg, R. L. and P. J. Coleman, Jr., The radial dependence of the interplanetary magnetic field: 1.0-.7 AU, Inst. Geophys. & Planet. Phys. Publ. #1196-26, Univ. of California, Los Angeles, 1973.

84. Russell, C. T., The Venus bow shock: Detached or attached?, J. Geophys. Res., in press, 1977.

85. Sagdeev, R. Z., Cooperative phenomena and shock waves, collisionless plasmas, Reviews of Plasma Physics, 4, ed M. A. Leontovich, Consultants Bureau, New York, 1966.

86. Sarris, E. T., S. M. Krimigis, and T. P. Armstrong, Observations of a high-energy ion shock spike in interplanetary space, Geophys. Res. Lett's., 3, 133, 1976.

87. Sarris, E. T., and J. A. Van Allen, Effects of interplanetary shock waves on energetic charged particles, J. Geophys. Res., 79, 4157, 1974.

88. Scarf, F. L., R. W. Fredricks, L. A. Frank, and M. Neugebauer, Non-thermal electrons and high-frequency waves in the upstream solar wind, I, Observations, J. Geophys. Res., 76, 5162, 1971.

89. Scarf, F. L., R. W. Fredricks, I. M. Green, and G. M. Crook, Observations of interplanetary plasma waves, spacecraft noise, and sheath phenomena on IMP 7, J. Geophys. Res., 79, 73, 1974.

90. Scudder, J. D., D. L. Lind, and K. W. Ogilvie, Electron observations in the solar wind and magnetosheath, J. Geophys. Res., 78, 6535, 1973.

91. Siscoe, G. L. and B. Goldstein, Solar wind interaction with lunar magnetic fields, J. Geophys. Res., 78, 6741, 1973.

92. Siscoe, G. L. and M. A. Heinemann, Binary stellar winds, Astrophys. & Space Sci., 31, 363, 1974.

93. Smith, E. J., L. Davis, Jr., D. E. Jones, P. J. Coleman, Jr., D. S. Colburn, P. Dyal, and C. P. Sonett, Jupiter's magnetic field, magnetosphere, and interaction with the solar wind: Pioneer 11, Science, 188, 451, 1975.

94. Smith, E. J., R. H. Holzer, M. G. McLeod, and C. T. Russell, Magnetic noise in the magnetosheath in the frequency range 3-300 Hz, J. Geophys. Res., 72, 4803, 1967.

95. Smith, E. J. and J. H. Wolfe, Observations of interactive regions and co-rotating shocks between one and five AU: Pioneer 10 and 11, Geophys. Res. Lett's., 3, 137, 1976.

96. Sonett, C. P. and D. S. Colburn, The SI^{+}-SI^{-} pair and interplanetary forward-reverse shock ensembles, Planet. Space Sci., 13, 675, 1975.

97. Spitzer, L., Jr., Physics of fully ionized gases, Interscience, Wiley, New York, 1962.

98. Spreiter, J. R., and A. Y. Alksne, Plasma flow around the magnetosphere, Magnetospheric Physics, ed D. J. Williams and G. D. Mead, William Byrd Press, Inc., Richmond, Virginia, 11, 1969.

99. Tidman, D. A., Turbulent shock waves in plasmas, Phys. Fluids, 10, 547, 1967.

100. Tidman, D. A., The earth's bow shock wave, J. Geophys. Res., 72, 1799, 1967.

101. Tidman, D. A. and N. A. Krall, Shock Waves in Collisionless Plasmas, John Wiley-Interscience, New York, 1971.

102. West, H. I. and R. M. Buck, Observations of > 100 keV protons in the earth's magnetosheath, J. Geophys. Res., 81, 569, 1976.

103. Zhigulev, V. N., On the phenomenon of magnetic detachment of the flow of a conducting medium, Soviet Phys., Doklady, 4, 514, 1959.

104. Zhigulev, V. N. and E. A. Romishevskii, Concerning the interaction of currents flowing in a conduction medium with the earth's magnetic field, Soviet Phys., Doklady, 4, 859, 1960.

Magnetic Field Reconnection

Bengt U. Ö. Sonnerup
Radiophysics Laboratory
Dartmouth College
Hanover, New Hampshire 03755

1. Introduction

Magnetic-field reconnection has been proposed as a basic energy-conversion process which may occur in many parts of the universe. Its primary function in the cosmic scheme is to prevent the build up of excessive amounts of magnetic energy in association with intense electric current sheets formed in highly conducting plasmas. The reconnection process is thought to cause a relaxation of such configurations, either partially or completely, and either continuously or sporadically, toward their lowest energy (current-free) state. The magnetic energy released during reconnection is converted into kinetic and internal energy of the plasma. The process causes the transfer of magnetic flux and plasma from topological cells with excessive flux to cells deficient in flux. This fact provides the basis for a precise definition of reconnection to be given in Section 3.4. Reconnection is also often referred to as magnetic field merging or magnetic field annihilation but, as will be seen, the three terms should not be used synonymously.

Figures 1-5 show examples of cosmic current sheets where reconnection may occur. Figure 1 represents the field produced by two photospheric dipoles which gradually move toward each other[94]. In the absence of reconnection, a current sheet of increasing length forms between the dipoles in the highly conducting solar atmosphere above them. If reconnection suddenly sets in, the magnetic field may relax toward a potential one, as indicated in the last picture of the sequence. This represents a possible, perhaps even plausible, mechanism for a solar flare[57,92]. Figure 2 illustrates current-sheet formation caused by the stretching of magnetic loops on the sun during rapid plasma ejection[14]. Figure 3 shows current sheets* separating interplanetary magnetic sectors with different polarity[123]. Figure 4 shows the magnetopause current

*stretching of magnetic loops on the sun during rapid plasma ejection

layer, formed as the solar wind presses the interplanetary magnetic field against the terrestrial field, as well as the tail current sheet, resulting directly from tangential stresses exerted by the solar wind on the magnetic field in the two tail lobes. The topology shown in the figure was first proposed by Dungey[32]. Figure 5 shows the magnetic field configuration expected for a rapidly spinning planetary magnetosphere such as that of Jupiter[48,67].

All of the above examples, and many possible other ones, such as supernova remnants[70], accretion disks[76], and galactic dynamos[87], illustrate cosmic situations in which magnetic field reconnection may occur. However, we do not know with certainty that the process does in fact take place in any or all of these geometries. And if it does take place, we still do not know much in detail about its dynamics. Are both continuous and sporadic reconnection possible, and if so, what are the plasma parameters and geometries in which these two modes are to be expected? What are the conditions for onset of reconnection? What is the energy conversion rate? In spite of twenty years of theoretical effort, recently summarized in a brilliant manner by Vasyliunas[118], as well as several laboratory experiments[10,31,84,85,114] and computer experiments[4,42,43], no universal agreement exists concerning the answers to most of these basic questions. Even in the most recent literature, opinions about the cosmic occurrence of the process range from full acceptance[118] to outright rejection[2,3]. On the other hand, there is conclusive evidence that reconnection occurs in tokamaks and other fusion devices as an end product of the resistive tearing-mode instability[122,120,121].

One of the difficulties with the cosmic reconnection research effort to date is that to a large extent it has lacked the detailed integration of theoretical and experimental work essential to the effective advancement of our knowledge concerning the process. On the one hand, an extensive but

rather abstract body of theoretical work exists[118], concerned primarily with the steady-state process and utilizing the fluid description. The latter is likely to be inadequate for the analysis of certain critical aspects of the process. On the other hand, laboratory experiments[10,11] indicate the importance of sporadic reconnection. However, the plasma parameters in these experiments are sufficiently different from those prevailing in most cosmic applications so as to pose serious difficulties in the application of the laboratory results in cosmos. A wealth of indirect observational evidence in the terrestrial magnetosphere, both at the magnetopause and in the tail, suggests that if the process occurs, it is likely to do so sporadically rather than continuously. In current observational magnetospheric work, the reconnection process is often invoked to account for a great variety of observations but with little effort to check theoretical predictions in detail or to consider alternate interpretations. The result is that the observational case for the occurrence of the process in the magnetosphere is not as solid as it might be. For other astrophysical applications, the situation is even worse.

On balance, our best opportunity for learning about reconnection as a viable cosmic energy conversion process is likely to be in the earth's magnetosphere. It is difficult to account for the overall dynamic behavior of the magnetosphere without invoking time-dependent transfer of magnetic flux from closed to open field lines and vice versa. And such transfer is one of the principal features of the reconnection process. The magnetosphere offers the unique advantage of permitting *in situ* plasma and field observations with probes that are much smaller than relevant plasma length scales. Thus an intense magnetospheric observational program with a focus on reconnection, coupled with a theoretical effort aimed at the geometries and plasma parameters

prevailing at the magnetopause and in the magnetotail would seem to have high potential for success. What is learned about reconnection in the magnetosphere may then be applied to other cosmic systems which do not permit *in situ* observations. It is seen that a research effort focused on magnetospheric reconnection is likely to lead to significant advances in our understanding of many other astrophysical and cosmic problems.

It is the purpose of this paper to provide a concise qualitative summary of the present state of reconnection theory and observations, with special reference to the earth's magnetosphere, and to bring into focus a number of specific problems and questions concerning the reconnection process in its magnetospheric application which should be studied both theoretically and observationally. The organization of the paper is as follows. First, a number of basic concepts are introduced via a qualitative discussion of steady two-dimensional reconnection in Section 2, and of possible nonsteady and/or three-dimensional configurations in Section 3. With this background, the more detailed technical discussion in subsequent sections can be presented in a compact fashion. Specifically, Section 4 deals with the external flow region, which is usually described in terms of the fluid approximation. Section 5 discusses one-fluid and two-fluid approaches to the plasma dynamics in the diffusion region, which is the site of the field reconnection process itself, and in which plasma microinstabilities are likely to be important. Section 6 discusses possible mechanisms for the generation of finite resistivity in the diffusion region and for the onset of reconnection. Section 7 contains a brief summary of present observational evidence for or against magnetospheric reconnection. Finally, Section 8 provides a summary of outstanding problems along with certain recommendations concerning the organization of future reconnection studies.

Three comments should be made about the scope of the paper. First, it does not attempt to provide a historical perspective. Rather it is organized to elucidate basic physical principles and recent significant approaches to the development of adequate theories of cosmic reconnection. Second, the paper does not attempt to cover all direct and indirect evidence for or against reconnection in the magnetosphere, on the sun, or elsewhere in cosmos. Third, the paper does not deal with applications in tokamaks and other laboratory devices where the physical boundary conditions are such that spatially periodic behavior results. It should be stressed, however, that vigorous interaction between fusion plasma physicists and cosmic physicists on the problem of reconnection is likely to be of substantial benefit to both groups.

Note that the reconnection process may be discussed entirely without reference to moving field lines and that indeed the latter concept might become invalid if substantial electric fields parallel to the magnetic field should develop. However, in the present simple model no such parallel fields occur except in the region very near the magnetic null. The use of the concept of moving field lines is then just another way of referring to the electric field E_o. In this model, the use of the term "reconnection" to describe the process is best understood in terms of moving field lines. As the lines C_1C_2 and D_1D_2 move with $\underline{v}_E$ toward each other through positions $C_1'C_2'$ and $D_1'D_2'$ they ultimately reach locations $C_1"C_2"$ and $D_1"D_2"$ where the lines meet at the origin. The surfaces through these lines and perpendicular to the plane of the figure are called separatrices, because they separate families of field lines of different topological origin. When the lines have reached this critical position, they appear to be cut and reconnected so that at still later times they are connected as $C_1'''D_1'''$ and $C_2'''D_2'''$, as shown in the figure. It is evident that the reconnection may be thought of as leading to a transport of magnetic flux from flux cells (1) and (2) across the separatrices into cells (3) and (4).

2.2 External Plasma Dynamics

Up to this point the description of the reconnection process has contained no reference to plasma dynamics. Indeed, the process may well have been imagined to occur in a vacuum. In such an instance, or if the field configuration is imbedded in a weakly conducting plasma, few restrictions exist on the magnitude of E_o, i.e., on the magnitude of $\underline{v}_E$. And the magnetic field will be equal to, or nearly equal to, a vacuum configuration with an angle α of nearly $\pi/2$ between the intersecting separatrices at the origin. The coupling between the electromagnetic field and the plasma is weak or absent. But in

virtually all cosmic applications of interest, the field configuration in Figure 6 would be imbedded in a plasma of high electrical conductivity. Indeed, in many cases Coulomb collisions may be considered entirely absent and the conductivity, if such a term is to be used, is associated with plasma turbulence and/or inertia and gyro effects, occurring near the magnetic null. Away from that point, the coupling between the $\underline{B}$ field and the plasma is strong and the plasma dynamics of the process will have dramatic effects in determining the detailed magnetic field configuration and perhaps in limiting the magnitude of the electric field E_o. We now outline some basic features of the plasma dynamics of the reconnection process.

First, it is observed that in a collision-free plasma the guiding centers of charged particles move with some velocity $\underline{v}_D$ under the influence of the electromagnetic field in Figure 6. In the drift approximation, which is expected to be valid, except in the immediate vicinity of the origin, the component of $\underline{v}_D$ parallel to the xy plane and perpendicular to $\underline{B}$ is identical with the flux transport velocity $\underline{v}_E$. Thus, in that plane, and as long as $\underline{E} \cdot \underline{B} = 0$, the magnetic field lines may be thought of as moving with the plasma or vice versa. We note that the simplified magnetohydrodynamic description also yields this result in the limit of an infinite electrical conductivity. The region away from the magnetic null in which plasma and fields move together is referred to as the *convection region*. Qualitatively the plasma motion is the one shown by the velocity arrows in Figure 6. Plasma approaches the origin along the positive and negative x axes and leaves along the positive and negative y axes. The motion may be the result of an external electric field E_o applied between capacitor plates at $z = \pm h$. Alternatively, E_o may be a polarization field created by an impressed plasma flow, specified in terms of a prescribed inflow rate at large $|x|$ values, say. The details

of the overall flow and field configuration will depend on these and other boundary conditions in a manner discussed in Section 4.1. However, all MHD models are expected to have in common the occurrence of large-amplitude standing waves in which the plasma is accelerated into the exit flow along the $\pm y$ direction, as shown in Figure 7. In incompressible analyses, these waves are Alfvén waves; in compressible flow they are slow shocks approaching the switch-off limit. The occurrence of these standing wave patterns is related to the fact that the propagation speed of these modes is very small in directions nearly perpendicular to the magnetic field. Thus, by arranging the angle between the wave normal and the upstream magnetic field to be sufficiently near 90°, the wave front can remain stationary even for very small inflow speeds along the $\pm x$ direction. The set of waves divides the flow field into two inflow regions and two outflow regions. These regions do not coincide exactly with the four flux cells in Figure 6. Because the separatrices are located upstream of the standing waves, parts of cells (3) and (4) overlap the inflow regions.

The standing waves contain concentrated electric currents, directed along the z axis as shown in Figure 7. The $\underline{j} \times \underline{B}$ force associated with these currents serves two purposes: it balances the difference in perpendicular momentum and in pressure of the plasma across the shock, and it accelerates the plasma in a direction tangential to the shock. It should be emphasized that currents are by no means confined to flowing only in the wave fronts. Distributed currents j_z may occur throughout the flow field. In particular, as will be shown in Section 4.1, the current distribution in the inflow region may influence the reconnection process in a crucial way.

An approximate balance of the magnetic shear stress at the shock and the exit momentum flow* yields

*In this calculation it is assumed that the plasma has a negligible velocity component along the y direction as it enters the shock. This assumption is not always valid. See Sections 4.1 and 4.2.

$$\rho_1 v_1 v_2 \simeq \frac{B_1 B_2}{\mu_0} \qquad 2\text{-}1$$

where ρ_1 is the plasma density in the inflow. Further, v_1, B_1, and v_2, B_2, are inflow and outflow speeds and magnetic fields, respectively. They are related via

$$E_o = v_1 B_1 = v_2 B_2 \qquad 2\text{-}2$$

If v_1 is eliminated between equations (1) and (2), we find

$$v_2 \simeq \frac{B_1}{\sqrt{\mu_0 \rho_1}} \equiv v_{A_1} \qquad 2\text{-}3$$

and

$$M_{A_1} \equiv \frac{v_1}{v_{A_1}} \simeq \frac{B_2}{B_1} \qquad 2\text{-}4$$

Thus it appears that, regardless of the inflow speed, the exit speed v_2 is always of the order of the Alfvén speed v_{A_1}, based on inflow conditions. Also, for fixed B_1, the magnitude of the magnetic field B_2 in the exit flow increases with increasing Alfvén number M_{A_1} in the inflow. When $M_{A_1} = 0$, $B_2 = 0$ and the configuration reduces to a current sheet. When $M_{A_1} \simeq 1$ the two fields are approximately equal, i.e., $B_2 \simeq B_1$.

In steady-state reconnection models, the inflow Alfvén number M_{A_1} is commonly used as a measure of the *reconnection rate*.

For very small values of M_{A_1}, and in a collision-less plasma, the plasma ejection along the $\pm y$ axis, postulated in the model in Figure 6, may become gradually replaced by an ejection at $z = -h$ and $z = +h$, respectively, of electrons and positive ions meandering in the current layer, as suggested by Alfvén[1] and discussed further by Cowley[23]. The charge separation effects in that case lead to an electric field E_z which is a function of the coordinate z. This limit will not be dealt with in the present paper.

2.3 Region Near the Magnetic Null

The preceding discussion has dealt with plasma motion away from the magnetic neutral point at the origin in Figure 6. Let us now briefly consider the region immediately adjacent to that point. As the origin is approached, the flux transport velocity $\underline{v}_E$ tends to infinity. Thus it is evident that the plasma can no longer move with $\underline{v}_E$ in the xy plane. In fact, as the plasma approaches the origin from both sides it must be brought to rest for symmetry reasons. In hydrodynamic terms, the magnetic neutral point is also a double stagnation point. The region in which the plasma velocity deviates significantly from $\underline{v}_E$ is referred to as the *diffusion region*; its dimensions are denoted by $2x^*$ and $2y^*$ as indicated in Figure 6. In this region finite conductivity effects of some type must come into play, allowing the current density to remain finite at the null point for $E_o \neq 0$. Three main possibilities exist.

(i) In a collisional plasma with large but finite electrical conductivity σ, the half width x^* of the diffusion region is expected to adjust itself in such a way that a balance is established between the rate of magnetic flux convected into the diffusion region and the rate of diffusion of that flux through the semistagnant plasma in the diffusion region. The ratio of these two transport rates is measured by the magnetic Reynolds number $R_m \equiv \mu_0 \sigma v_1 x^*$. Thus we expect $R_m \simeq 1$, i.e., x^* is of the order of the resistive length:

$$x^* \simeq (\mu_0 \sigma v_1)^{-1} \qquad 2\text{-}5$$

We note that x^* decreases with increasing conductivity and increasing v_1. Since $v_1 = v_E = E_0/B_1$, B_1 being the magnetic field at $(x = \pm x^*, y = 0)$, increasing v_1 corresponds to increasing E_0, assuming B_1 to remain fixed.

(ii) In a collision-free plasma one might expect the value of x^* to be determined instead by the scale of the particle orbits near the null point. Four such scales may be of relevance: the electron and ion gyroradii and the electron and ion inertial lengths. Further discussion of these scales is presented in Section 5. An equivalent electrical conductivity may be imagined in this case, such that the effective residence time of a particle (an electron or an ion) replaces the usual collision time τ in the expression $\sigma = ne^2\tau/m$ (m = particle mass). This residence time is found to be inversely proportional to v_1 so that $x^* \sim (\mu_0\sigma v_1)^{-1}$ becomes independent of v_1 and hence of E_0 for fixed B_1. For further discussion, see section 6.1.

(iii) In each of the above two cases, the current density or the gradients in the diffusion region may become sufficiently large to cause plasma micro-instabilities. The resulting plasma turbulence will lead to a reduction in the effective conductivity, as discussed in section 6.2.

Whether the plasma dynamics in the diffusion region is described in a continuum fashion, i.e., by use of an effective conductivity, or in terms of individual particle orbits near the magnetic null point, it is easy to see that the net current $\underline{I}$ in the diffusion region will be along the positive z axis so that $\underline{E} \cdot \underline{I} > 0$. Thus the diffusion region, along with the entire shock system, acts as a dissipator of electromagnetic energy.

We note that the overall conservation of mass in the diffusion region yields

$$\rho_1 v_1 y^* = \rho_2 v_2 x^* \qquad 2\text{-}6$$

which may be combined with 2-2 and 2-4 to yield

$$M_{A_1} \equiv \frac{v_1}{v_{A_1}} \simeq \frac{B_2}{B_1} \simeq \frac{\rho_2}{\rho_1}\frac{x^*}{y^*} \qquad 2\text{-}7$$

Assuming the density ratio ρ_2/ρ_1 to vary only moderately with M_{A_1}, we see that the diffusion region is very elongated along the y axis for small M_{A_1} values. Additionally, in a collisional plasma the thickness x^* increases with decreasing M_{A_1}, as shown by equation 2-5:

$$x^* \simeq (\mu \sigma v_{A_1})^{-1} (M_{A_1})^{-1} \qquad 2\text{-}8$$

Combining equations 2-7 and 2-8 it appears that $x^* \sim M_{A_1}^{-1}$, $y^* \sim M_{A_1}^{-2}$ in a collisional plasma (case i) while $x^* \sim$ const., $y^* \sim M_{A_1}^{-1}$ in a collision-free case dominated by inertial resistivity (case ii).

Finally, we estimate the separatrix angle α in the outflow (see Figure 6). Near the magnetic null point we may write

$$\left.\begin{aligned} B_x &= ay \\ B_y &= bx \end{aligned}\right\} \qquad 2\text{-}9$$

where a and b are positive constants, and the angle $\alpha = 2 \tan^{-1}\sqrt{a/b}$. Estimating $ay^* \simeq B_2$ and $bx^* \simeq B_1$ we find by use of equation (7)

$$\alpha = 2 \tan^{-1}\sqrt{\frac{B_2 x^*}{B_1 y^*}} \simeq 2 \tan^{-1}\left(\frac{x^*}{y^*}\sqrt{\frac{\rho_2}{\rho_1}}\right) \simeq 2 \tan^{-1}\left(\sqrt{\frac{\rho_1}{\rho_2}}\, M_{A_1}\right) \qquad 2\text{-}10$$

indicating that the range of Alfvén numbers M_{A_1} from zero to $\sqrt{\rho_2/\rho_1}$ corresponds to an α range of zero to $\pi/2$. The latter value corresponds to $b = a$, i.e., to a current-free state because $j_z = (b - a)/\mu_0$.

2.4 Energy Conversion

The reconnection model described in this section serves as a steady-state converter of electromagnetic energy into plasma kinetic and internal energy. For example, the rate of electromagnetic energy flow into and out of the diffusion region may be estimated as follows:

$$\left.\begin{aligned} \text{Inflow} &= 8y^*h\, E_0\, B_1/\mu_0 \\ \text{Outflow} &= 8x^*h\, E_0\, B_2/\mu_0 \end{aligned}\right\}$$

where the diffusion region has been taken to be a rectangular box with sides $2y^*$, $2x^*$ and $2h$. Thus the net rate of inflow of electromagnetic energy is

$$W_{EM} = 8y^*h\, E_0\, B_1/\mu_0 \left(1 - \frac{x^*}{y^*}\frac{B_2}{B_1}\right)$$

which upon use of equations 2-2 and 2-7 may be written

$$W_{EM} \simeq 16y^*h\left(\frac{B_1^2}{2\mu_0}\right) v_{A_1} M_{A_1}\left(1 - M_{A_1}^2 \frac{\rho_1}{\rho_2}\right) \qquad 2\text{-}11$$

It is evident from this approximate expression that the energy conversion rate has a maximum at some value of the reconnection rate M_{A_1} intermediate between 0 and a maximum value, which in the present approximate set of relations appears to be $M_{A_1} = \sqrt{\rho_2/\rho_1}$. Note that $W_{EM} = 0$ both for $M_{A_1} = 0$ and for $M_{A_1} = \sqrt{\rho_2/\rho_1}$. For the latter value of M_{A_1}, the configuration near the null is current-free and symmetric ($b = a$; $\alpha = \pi/2$). In such circumstances one may expect $\sqrt{\rho_2/\rho_1} \simeq 1$. Thus $M_{A_1} \simeq 1$ appears as a theoretical upper limit for the reconnection rate (based on plasma conditions at $x = x^*$ $y = 0$). It is, however, by no means assured that boundary conditions at large distances or plasma processes in the diffusion region will always permit this upper limit to be reached.

The net rate of increase of kinetic energy of the plasma may be expressed as follows

$$W_{KE} \simeq \rho_1 v_1\, 8y^*h\left(\frac{v_2^2}{2} - \frac{v_1^2}{2}\right) \simeq 8y^*h\left(\frac{B_1^2}{2\mu_0}\right) v_{A_1} M_{A_1}(1 - M_{A_1}^2) \qquad 2\text{-}12$$

and conservation of energy requires the difference $W_{EM} - W_{KE}$ to be the rate of increase of the internal energy of the plasma, W_I. This latter rate may include thermal as well as nonthermal parts, for example in the form of run-away electrons.

The analysis given above applies to the diffusion region. But usually only a minute part of the total energy conversion occurs there, the main part taking place in the shocks. In approximate terms, the formulas 2-11 and 2-12 may be modified to be valid for the entire reconnection geometry by replacing y^* by L, where $2L$ is the height of the total configuration, as shown in Figure 6. Also, all quantities bearing the subscript 1 (which are evaluated at $x = x^*$ $y = 0$) should be replaced by quantities bearing the subscript ∞, i.e., they should be evaluated at $x >> x^*$, $y = 0$. Depending on the nature of the boundary conditions, the inflow may be such that M_{A_∞} differs significantly from M_{A_1} (see Sections 4.1 and 4.2).

The phrase magnetic field annihilation has been used to describe the reconnection process. In the light of the preceding discussion, this term appears appropriate only in the limit of small M_{A_1} values where the magnetic field B_2 in the exit flow is small (or absent as in Alfvén's model, mentioned earlier[1,23]). Henceforth, annihilation will refer to situations where M_{A_1} is sufficiently small so that the diffusion region occupies the entire length of the current sheet, i.e., $y^* \geq L$. By combination of equations 2-7 and 2-8 this is seen to occur for $0 < M_{A_1} \leq \sqrt{(\rho_2/\rho_1)/(\mu_0 \sigma v_{A_1} L)}$.

In reconnection, energy conversion occurs on a time scale comparable to the Alfvén wave time $\tau_A = L/v_{A_1}$ (assuming the inflow regions to extend to $|x| \simeq L$), while in annihilation the scale is $\sqrt{\tau_A \tau_D}$, τ_D being the time for purely resistive decay of a current sheet i.e., $\tau_D = \mu_0 \sigma L^2$. τ_D is enormous in most cosmic applications, so that reconnection rather than annihilation is required to account for the rapid energy release in solar flares, geomagnetic substorms, etc.

3. Flux Transfer in Time-Dependent and Three Dimensional Configurations

The two-dimensional steady reconnection model outlined in Section 2 is useful as a vehicle for introducing certain basic aspects of reconnection. But it appears likely that in any real cosmic applications of the process, three-dimensional and temporal effects are important, perhaps even dominant. For this reason it is useful to consider briefly a few reconnection configurations which incorporate these effects. To date, the plasma dynamics associated with such geometries has not been dealt with in a substantial way, so that the discussion is confined mainly to the electromagnetic field topology and flux transfer aspects of the process. In the following subsections we describe the two-dimensional but time-dependent double inverse pinch configuration, a simplified steady-state three-dimensional magnetopause topology and a possible three-dimensional time-dependent magnetotail configuration. Finally, in Section 3.4, a general definition of reconnection is given.

3.1 Plane Time-Dependent Geometry

A plane vacuum magnetic field geometry associated with the double inverse pinch laboratory experiments[10] is shown in Figure 8. The X type magnetic null point is again located at the origin. The magnetic field is maintained by the currents I in the two metal rods at the center of flux cells (1) and (2), and a return current $2I$, flowing in the plasma along an outer envelope, which coincides with the outermost field lines in flux cell (3). In the experiments, the current I increases with time so that magnetic flux is generated continually at the two rods, i.e., in cells (1) and (2). If we assume for a moment that no plasma is present, the flux in cell (3) increases proportionately so that magnetic flux may be thought of as being transported from the rods into cells (1) and (2) and from there across the

separatrix into cell (3). It is of interest to calculate the electric field responsible for this flux transport. The vector potential for the vacuum configuration is given by

$$\underline{A} = -\hat{\underline{z}} \frac{\mu_0 I}{2\pi} \ln \frac{c^2+d^2}{r_1 r_2} = \hat{\underline{z}} A_z(x,y,t) \qquad 3\text{-}1$$

where the rod separation is $2c$, the minor diameter of the return-current envelope is $2d$, and the radii r_1 and r_2 are measured from the two rods as shown in Figure 8. Note that $\underline{A} = 0$ at the envelope. In the experiment, the current I and the envelope diameter both increase with time; in a more general case, the rod separation might be imagined to depend upon time also. But for our purposes it suffices to consider the time variation of the current I and the diameter d. Then, the electric field is

$$\underline{E} = -\frac{\partial \underline{A}}{\partial t} = \hat{\underline{z}} \frac{\mu_0 \dot{I}}{2\pi} \ln \frac{c^2+d^2}{r_1 r_2} + \hat{\underline{z}} \frac{\mu_0 I}{2\pi} \frac{2d\dot{d}}{c^2+d^2} =$$

$$-\underline{z} \frac{\dot{I}}{I} A_z(x,y,t) + \hat{\underline{z}} \frac{\mu_0 I}{2\pi} \frac{2d\dot{d}}{c^2+d^2}$$

Since at each instant A_z remains constant on a magnetic field line, the instantaneous electric field has the same value on a given field line but its value changes from one line to another. In particular, on the separatrix it has the value

$$E_z = \frac{\mu_0 \dot{I}}{2\pi} \ln\left(1 + \frac{d^2}{c^2}\right) + \frac{\mu_0 I}{2\pi} \frac{2d\dot{d}}{c^2+d^2} \qquad 3\text{-}2$$

Thus for increasing current I and diameter d, E_z is positive as required for flux transport into cell (3).

In the presence of a plasma, the field configuration is modified as follows The electric field now drives plasma currents in the vicinity of the magnetic null line, causing a field deformation of the type shown by the dashed lines in Figure 8. The separatrix intersection angle α falls below its vacuum value

of $\pi/2$. These effects imply an excess of magnetic flux in cells (1) and (2), a deficiency in cell (3), compared to the vacuum configuration, which is the lowest energy state. Thus, a certain amount of free magnetic energy is stored in the system. However, at the same time a considerable amount of flux transport into cell (3) takes place. That is, reconnection occurs continuously*. The principal difference between the present case and the steady-state model in Section 2 is the spatial nonuniformity of the instantaneous electric field. This effect occurs because in the nonsteady case some of the flux transported in the xy plane is being deposited locally, causing a field magnitude increase at each point. Associated with this flux accumulation, a plasma compression must also occur. But this would appear to be a relatively minor effect so that the steady model in Section 2 may provide an adequate instantaneous description of the flow away from the rods and the return envelope. Thus the essential qualitative features of the reconnection flow may be obtained by examination of a sequence of steady-state configurations.

Impulsive flux transfer events are observed in the double inverse pinch experiments. It appears that, as the magnetic field and associated plasma currents near the null point grow, anomalous resistivity associated with ion sound turbulence sets in abruptly with an associated rapid increase of electric field and decrease of currents at the null point. The net result is a much more rapid flux transfer into cell (3) and an associated relaxation of the entire magnetic field configuration toward its potential form with the separatrix intersection angle α increasing toward $\pi/2$. Evidently the stored free magnetic energy described in the previous paragraph is being rapidly converted into plasma energy. These events occur on a time scale much shorter than that associated with $\dot{I}$. Hence it is unlikely that they may be described,

*By contrast, Ref. *15* analyzes a hyperbolic-field collapse, where α decreases from $\pi/2$ to 0, without any reconnection.

even approximately, by a sequence of steady-state configurations. But the conditions for onset of such an event may perhaps be identified by examination of such a sequence.

3.2 Steady Three-Dimensional Geometry

A three-dimensional magnetic-field configuration of interest for steady-state magnetopause reconnection is obtained by the superposition of a dipole and a uniform field of arbitrary direction. This topology, shown in one cross section in Figure 9, has been discussed extensively in the literature[24,32,127]. Two hyperbolic magnetic null points X_1 and X_2 are formed in the plane containing the dipole moment vector and the uniform field vector. A basic topological property of such a null point is that many field lines enter it forming a separatrix surface and two single field lines leave it along directions out of that surface, or vice versa. The separatrix surfaces associated with X_1 and X_2 intersect along a circular ring located in a plane through the two points and perpendicular to the plane of Figure 9. This ring is referred to alternatively as a singular line, a reconnection or merging line, a critical line, an X line, or a separator line. At a chosen point on the ring the magnetic field does not vanish in general, but it is directed along the ring. Only at X_1 and X_2 is the field intensity zero. If the uniform field is exactly antiparallel to the dipole field a degenerate situation arises in which the magnetic field vanishes at each point on the ring.

A schematic picture of the two separatrix surfaces is shown in Figure 10, in a configuration that may be appropriate for magnetopause reconnection. The upper part of the figure shows a view in the antisolar direction of field lines on the separatrix surface associated with the null point X_2; the lower part shows the same view of the X_1 separatrix. The total picture is an overlay

of the two diagrams with the reconnection line connecting X_1 and X_2. Part of the solar-wind electric field E_{sw} is impressed across the configuration and must be sustained along the reconnection line. Thus, in the vicinity of that line a strong electric field component is present along the magnetic field. Unless special circumstances exist, such parallel electric fields do not arise in highly conducting plasmas. However, it is believed that the field lines on the separatrix and its immediate vicinity bend to become nearly parallel to the reconnection line extremely close to that line, as shown in Figure 10. Thus parallel electric fields occur only within the diffusion region which surrounds the reconnection line and in which finite resistivity effects permit their presence. Figure 10 suggests that it may be possible to study reconnection in this geometry by use of a locally two-dimensional model which is then applied to each short segment of the reconnection line. Such a model will be similar to that discussed in Section 2, but with an added magnetic field component $B_z(x,y)$. Thus the reconnection of fields that are not antiparallel is obtained. Further discussion of such geometries is given in Section 4.4. The dynamics of the motion near the points X_1 and X_2 has not been studied to date. It may well be that these points mark the end points of a reconnection line segment on the front lobe of the magnetopause surface.

Referring to Figure 5, which represents a cut through the earth's magnetosphere in the noon midnight meridional plane, it is seen that reconnection at the magnetopause, as described above, serves to transport magnetic flux from the interplanetary cell (1) and from the front-lobe magnetospheric cell (2) into the polar cap cells (3) and (4).

3.3 Time-Dependent Three-Dimensional Geometry

As a final example of reconnection geometries of cosmic interest,

consider the magnetic-field topology associated with the formation of a reconnection bubble in the geomagnetic tail. The evolution of the field geometry in the noon-midnight meridional plane is shown in Figure 11. Note the formation of an *X* type and an *O* type neutral point. The bubble originally has a very small longitudinal dimension. As it grows in size in the noon-midnight plane, it also occupies an increasing longitude sector. The actual three-dimensional magnetic field topology of such a bubble is not known, but it may be represented schematically by an *X* type and an *O* type null line as in Figure 12. The points *A*, *X*, *B* and *O* in that figure all emerge at the same place at the time of onset of reconnection. Subsequently they move apart as the reconnection process continues and the bubble grows. An electric field exists along the reconnection line *AXB* but none, or almost none, along the *O* line *AOB*. This field presumably has an inductive and an electrostatic part which tend to cancel along *AOB* while adding along *AXB*.

3.4 Definitions

On the basis of the preceding discussion we now formalize the definition of several terms, used in the magnetic-field reconnection literature:

(i) A *separatrix* is a surface in space which separates magnetic field lines belonging to different topological families. By necessity the separatrix is everywhere tangential to the magnetic field. The field lines constituting the surface originate at a hyperbolic neutral point in the field.

(ii) A *separator* is the line of intersection between two separatrices or the line of intersection of one separatrix with itself. The separator is also called reconnection line, merging line, or *X* line. The terms neutral line, singular line, or critical line should be avoided, since they may refer to the *O*-type topology as well.

(iii) The *diffusion region* is a plasma channel, surrounding the separator, in which resistive diffusion, caused by collisional processes, turbulence, or inertial effects, is important. In a highly conducting plasma, the diffusion region is imbedded in a much larger *convection region*, in which magnetized plasma moves toward and away from the separator, in the *inflow* and *outflow* regions, respectively, and in which dissipative effects are confined to shocks.

(iv) *Magnetic-field reconnection* is said to occur when an electric-field component E_o (induced or electrostatic) is present along a separator or a macroscopic portion thereof. It is proposed that the term *magnetic-field annihilation* be reserved for the case where the separator has degenerated (for dynamic purposes*) to a surface (e.g., the surface separating two half spaces containing antiparallel uni-directional fields). The term *magnetic-field merging* may be taken to encompass both reconnection and annihilation.

(v) The local instantaneous *reconnection rate* at a chosen point on a separator is measured by the instantaneous magnitude of the electric-field component E_o along that line. It is desirable to express this rate in a nondimensional form by dividing the electric field by the product of a characteristic velocity and a characteristic magnetic field. The latter two quantities may be taken to be the Alfvén speed v_{A_r} and magnetic field B_r at a chosen reference point, denoted by the subscript r, in the inflow, such as $(x = x^*, y = 0)$ or $(x = L, y^* = 0)$. Since E_o/B_r represents a characteristic flow speed, the dimensionless reconnection rate takes the form of an Alfvén number:

$$M_o = \frac{E_o/B_r}{v_{A_r}}$$

In steady, two-dimensional ($B_z = 0$, $\partial/\partial z = 0$), models the electric field E_o is constant throughout the xy plane so that $E_r = E_o$. With the reference point in

*see comments in sections 2.4 and 6.1.

the convection region, and on the X axis where $B_x = 0$, E_o/B_r is then the plasma flow speed toward the separator at the reference point and M_o is the local Alfvén number, $M_o = M_{A_r}$. In nonsteady flow, the electric field at the reference point, E_r, in general differs from E_o, and $M_o \neq M_{A_r}$.

Vasyliunas[118] has defined magnetic merging as "the process whereby plasma flows across a surface that separates regions containing topologically different magnetic field lines"; he takes the magnitude of that flow as a measure of the merging rate. For reconnection in a highly conducting plasma, such that $R_m = \mu_0 \sigma v L >> 1$, the two definitions are essentially equivalent. However, the one adopted here, in terms of an electric field component along the separator works also for flows at arbitrary R_m. It corresponds to the occurrence of flux rather than plasma transport across the separatrix, because flux transport is but an alternate way of referring to the electric field* Note also that for the degenerate case of magnetic field annihilation there is no plasma flow across a separatrix. There is, however, an electric field and a corresponding magnetic flux transport.

*This equivalence is seen most clearly[119] by casting Faraday's law into the form of a conservation equation, viz., in subscript notation, $\partial B_i/\partial t + \partial/\partial x_j (\varepsilon_{ijk} E_k) = 0$, where ε_{ijk} is the antisymmetric (Levi-Civita) unit tensor.

4. The Convection Region

The plasma dynamics in the regions away from the immediate neighborhood of the reconnection line usually is described by use of continuum equations. Nonsteady solutions have not been found to date, which describe rapid configuration changes such as might be associated with impulsive flux transfer events in the double inverse pinch experiment (for a circuit model, see Bratenahl and Baum,[11]). Three-dimensional solutions also have not been obtained. Hence the discussion in the present section is confined to steady-state plane reconnection.

The incompressible assumption corresponds to the limit $\beta \to \infty$, where β is the ratio of plasma pressure to magnetic pressure. While this limit is invalid in most cosmic applications, it has the advantage of yielding simple analysis. Thus it provides an opportunity to study certain basic features of the reconnection flow without undue mathematical complications. We first describe two incompressible reconnection flows with fundamentally different behavior. Certain compressibility effects are considered in the second subsection. The third subsection discusses asymmetric reconnection configurations, perhaps applicable to the magnetopause. The fourth subsection deals with the reconnection of magnetic fields that are not antiparallel, a common situation at the magnetopause. Finally, a partial single-particle model is discussed briefly.

4.1 Two Incompressible Symmetric Flow Models

Figure 13, reproduced from Vasyliunas[118], shows a field and flow map for a reconnection model initially analyzed by Petschek[90] and subsequently considerably refined and improved by Vasyliunas. The model contains a set of four Alfvén discontinuities which in compressible flow may be identified

as slow-mode shocks and across which the plasma is accelerated into the exit flow regions. Note that the plasma flow converges toward the x axis in the inflow and that the magnetic field intensity decreases on that axis for decreasing $|x|$ values. As pointed out by Vasyliunas, this behavior is characteristic of fast-mode expansion of the plasma as it approaches the reconnection line. Because the fast-mode propagation speed is infinite in the incompressible limit, such expansion is by necessity an elliptic effect, that is, no standing expansion wavelets are possible. The maximum reconnection rate in this model corresponds to an Alfvén number M_{A_1} of about one in the inflow just adjacent to the diffusion region. But because of the increase in flow speed and decrease in magnetic field associated with the fast-mode expansion, the Alfvén number, M_{A_∞}, at large distances upstream is considerably less than unity. Values in the range $.05 < M_{A_\infty} < .2$ for the maximum rate are obtained (see Ref. *118* Fig. 12). Recently, Soward and Priest[110] have reexamined Petschek's reconnection geometry by use of an asymptotic approach, valid away from the reconnection line. Their analysis in all essential respects supports the conclusions summarized above.

Figure 14 also taken from Ref. *118*, shows a flow and field map for a different model[103], which is the sole nonsingular member of the similarity solutions derived by Yeh and Axford[131]. This model contains a second set of Alfvén discontinuities located upstream of the slow shocks and originating at external corners in the flow, as shown in the figure. These discontinuities represent the incompressible limit of slow-mode expansion fans centered at the external corners. They cause a large deflection of the plasma flow away from the x axis and a substantial increase in field magnitude. It is now generally agreed that these discontinuities will not occur in any real situation. Rather they represent a suitable mathematical lumping of slow-mode expansion effects

in the inflow. The maximum reconnection rate in this model* is $M_{A_1} = (1 + \sqrt{2})$. On the x axis this value remains constant, independent of $|x|$. However, this is a result of the lumping of the slow-mode effects. In a model where these effects are spread over the inflow region the value of M_A on the x axis will decrease with decreasing $|x|$ in association with a decrease in plasma velocity and an increase in magnetic field. Thus, in reality it is unlikely that the inflow into the diffusion region can occur at M_{A_1} as high as $(1 + \sqrt{2})$; more likely that value corresponds to the maximum M_{A_∞} at large $|x|$ values. Further discussion of this point is given in Sections 4.2 and 5.1.

The two models discussed above represent two extreme sets of conditions in the inflow: pure fast-mode and pure slow-mode expansion. In any real application both effects may be present. Vasyliunas[10] has pointed out that from a mathematical viewpoint the difference between the two models is related to the boundary conditions at large distances from the reconnection line. Far upstream, the fast-mode model is essentially current free and has a nearly uniform flow and magnetic field, while the slow-mode model contains substantial currents which bend the magnetic field lines and cause a deflection of the flow away from the x axis. Vasyliunas has further suggested that the former state of affairs may obtain when a demand for magnetic flux originates at the current sheet itself (the xy plane) or in the exit flow, as may be the case in the geomagnetic tail, while the latter set of conditions may correspond to externally forced inflow such as at the magnetopause. In this context, it is worth noting that slow-mode expansion effects have been argued[132] to be present outside

*The estimates given in Section 2, viz., $v_2 \simeq v_{A_1}$ and $(v_1)_{max} \simeq v_{A_1}$ assumed a negligible flow component along the y axis as the plasma enters the shock. Such a component is present in this model, the result being that the exit flow speed v_2 and the maximum inflow speed $(v_1)_{max}$ both exceed v_{A_1} by a factor $(1 + \sqrt{2})$.

the subsolar magnetopause regardless of whether or not reconnection occurs there.

4.2 Compressible Symmetric Models

A detailed compressible analysis of the external region of Petschek's reconnection geometry is not available at present. On the other hand, the slow-mode expansion model has been extended to include compressibility effects. An isothermal analysis was given by Yeh and Dryer[130]. But the isothermal assumption leads to unacceptable entropy variations with decreasing entropy across the shocks and increasing entropy across the expansion waves. More recently, an analysis has been performed by Yang and Sonnerup[125], which assumes isentropic flow in the inflow and uses the ordinary jump relations for slow shocks. It is found that the expansion-wave discontinuities in the incompressible solution do indeed dissolve into slow expansion fans centered at the external corners in the flow (see Figure 15). It might be thought that the reflection of these fans in the x axis, and the subsequent interaction of the reflected waves with the shocks, shown schematically in Figure 16, may be treated exactly by the method of characteristics. However, it is found that the flow from region ① in the figure, across the last expansion wavelet and the innermost portion of the shock, cannot be dealt with without the inclusion of elliptic (fast-mode) effects. This is extremely difficult to do. Thus, in the main part of their work, Yang and Sonnerup, after calculating the isentropic plasma and field changes across the fans, considered them to be lumped into a single discontinuity, i.e., they ignored the reflection altogether. While such a procedure is perhaps justified in a first attempt to study compressibility effects in the external flow, it nevertheless seriously limits the usefulness of the resulting solutions. The width of the slow expansion fans in the inflow increases dramatically with increasing compressibility,

i.e., with decreasing values of $\beta_1 \equiv 2\mu_0 p_1/B_1^2$, so that for $\beta_1 \simeq 1$ the lumping of the fan into a single discontinuity is difficult to defend. Furthermore, except perhaps for very large β values, conditions immediately outside the diffusion region are not adequately represented so that the solution may not be used to provide external boundary conditions for compressible matched diffusion-region analyses. However, the analysis is valid at large distances from the origin, and it is of interest to examine its predictions concerning flow and plasma conditions in the exit regions. When conditions typical of geomagnetic tail reconnection are substituted, flow speeds in the range of 1000 km/s are calculated, in rough agreement* with observed proton speeds in the tail during energy-release events[39,63]. The analysis also predicts exit flow speeds considerably greater than the fast-mode propagation speed so that standing transverse fast shocks may be present in the two exit flow regions, causing a decrease in flow speed and an associated increase in plasma density, temperature, and in the exit magnetic field.

Yang and Sonnerup[125] also calculated the change in plasma and flow properties along the x axis in Figure 16, caused by the reflection of the slow expansion fan, but ignoring the elliptic effects mentioned earlier. The solid curve in Figure 17 shows the resulting relationship between the Alfvén numbers M_{A_1} and M_{A_∞}, in regions (1) and (∞) of Figure 16, respectively. For comparison, the corresponding relationship for the fast-mode model, developed by Soward and Priest[110], is shown by the dashed curves. It is evident that the different distant boundary conditions for the fast-mode and the slow-mode models may lead to profoundly different inflow conditions into the diffusion region for the two models.

*The agreement is however not sufficiently detailed to support this particular reconnection configuration over others.

4.3 Asymmetric Models

A qualitative reconnection model for the asymmetric flow and field conditions existing at the magnetopause was first described by Levy *et al.*[72,91]. In this model, shown in Figure 18, the magnetosheath plasma is assumed to carry with it a compressed interplanetary magnetic field which is due south so that a field reversal results at the magnetopause (see Figure 4). In impinging on the earth's field, the plasma encounters a wave system consisting of an intermediate wave (rotational discontinuity; large amplitude Alfvén wave) followed by a narrow slow expansion fan. The intermediate wave, which marks the magnetopause, accomplishes the field direction reversal and an associated plasma acceleration parallel to the magnetopause and away from the reconnection line. The magnetic-field magnitude remains constant across this wave but it then increases to its higher magnetospheric value in the slow expansion fan across which the plasma pressure also is reduced to match the pressure in the magnetosphere, taken to be equal to zero in the model. The quantitative details of this model have not been worked out for fast-mode expansion in the inflow. But for the incompressible slow-mode expansion model, various types of asymmetries in the flow and field have been analyzed [26, 82, 106]. Certain compressible counterparts of these geometries have been studied by Yang[124] by use of the procedure of lumping slow-mode effects in the inflow, discussed in Section 4.2. In particular, the case of vacuum conditions in the magnetosphere has been reported in detail[126]. A typical resulting field and flow configuration is as shown in Figure 18. The model predicts the following features of the magnetic field: (i) a small magnetic field component normal to the magnetopause; (ii) rotational behavior of the magnetic-field component tangential to the magnetopause; (iii) a gradual increase in magnetic field intensity inside the magnetopause. In terms of plasma

flow, the model predicts: (i) flow of magnetosheath plasma normal to and across the magnetopause with speed equal to the Alfvén speed based on the normal magnetic field component; (ii) tangential acceleration of the plasma to speeds of the order of 500-750 km/s as it crosses the magnetopause; (iii) no change in plasma density or temperature as it crosses the magnetopause; (iv) an isentropic decrease in density and temperature and an associated velocity increase as the plasma expands in the slow expansion fan inside the magnetopause. To date, the predicted plasma behavior has not been observed. At various times some but usually not all of the predicted magnetic signatures have been seen[108]. An example is shown in Figure 19.

4.4 Reconnecting Fields Forming an Arbitrary Angle

In his original paper on reconnection, Petschek[90] observed that in incompressible two-dimensional reconnection flow, a constant magnetic field component B_z could be added to any solution without changing the flow or magnetic field configuration in the xy plane (refer to Figure 6). Thus, it is possible to generate solutions that describe the reconnection of fields that form an arbitrary angle. This provides a link between the traditional cosmic reconnection models and the problems of reconnection in nearly force-free field configurations, such as the tokamak[120,121,122]. This procedure has provided the basis of a number of attempts to describe the dependence of the cross-magnetospheric electric potential difference on magnetosheath field direction, assuming the former to be caused by magnetopause reconnection[49,58,59,107]. The result of these geometrical analyses is that the potential should have a principal angle dependence given by the function*

$$f(\theta) = \frac{(B_o/B_i - \cos\theta)^2}{1 + (B_o/B_i)^2 - 2(B_o/B_i)\cos\theta} \qquad 4\text{-}1$$

*In Ref. 49 the functional dependence is $(f(\theta))^{\frac{1}{2}}$

where θ denotes the angle between the magnetic field B_i in the magnetosphere and B_o in the magnetosheath. For $cos\theta \geq B_o/B_i$ no reconnection occurs* and $f(\theta) \equiv 0$. Observations[13] indicate a low energy injection rate from the solar wind into the magnetospheric ring current system when $\theta < \pi/2$, a result which appears compatible with Equation 4-1. Observations[12] of the magnetospheric cusp location as a function of θ also are in qualitative agreement with this equation.

Recently, Cowley[25,28] has pointed out that incompressible solutions exist in which B_z is a function of x and y. Thus the assumption underlying Equation 4-1, of one and the same value of B_z on the magnetospheric and the magnetosheath side of a typical magnetopause reconnection model, may be invalid. It is noted that this assumption corresponds to a situation where the net current in the magnetopause (and in the diffusion region) is parallel to the separator.

In the incompressible MHD approximation the equations describing the flow and field in the xy plane are completely uncoupled from the differential equations for the velocity component v_z and for B_z. However, as pointed out by Cowley[25], an indirect coupling exists via the requirement that the line integral $\oint \underline{E} \cdot \underline{dl} = 0$ for any closed loop which encircles the diffusion region. This requirement poses an additional constraint on the shape of the diffusion region, a constraint that does not arise when $B_z = 0$, or $B_z = const.$, and that does not appear to be automatically satisfied by the diffusion region of plane reconnection geometries. Thus it is not clear at the present time whether Cowley's criticism of Equation 4-1 has a firm foundation in incompressible MHD theory. But even if it doesn't, the use of a constant B_z in the real compressible magnetopause flow situation to construct reconnection geometries for arbitrary θ values remains hypothetical. Further theoretical study of this problem requires detailed analysis of

*If the magnetosheath field magnitude exceeds the magnetosphere field, the subscripts o and i are to be interchanged.

compressible external and diffusion-region flows and appropriate matching of these flows at the edge of the diffusion region.

4.5 Collisionless Model

No complete or nearly complete collisionless models for the external reconnection region have been developed to date. But certain partial results have been obtained by Hill[59]. He suggests that for small β values the principal field reversal is accomplished by a current sheet located on the y axis, as shown in Figure 20. with only a small amount of field-line bending at the slow shocks. Hill does not treat the flow and field configurations in the inflow or in these weak shocks. Rather he assumes that, away from the magnetic null point at the origin, the field lines form an angle χ with the current sheet. He then proceeds to discuss the properties of the sheet. One-dimensional self-consistent Vlasov equilibria of such sheets have been obtained numerically by Eastwood[36,37]; an approximate analytic theory using the adiabatic invariant[34,65,104] associated with the meandering particle orbits in the sheet has also been given[38]. However, the result primarily used by Hill is obtained directly from the overall stress balance at the sheet, without reference to the sheet structure: in a frame of reference sliding along the y axis (see Figure 20) with a speed such that the reconnection electric field E_o vanishes, the usual firehose limit must apply. Assuming the magnetic moment of individual particles to be preserved, Hill shows that for inflow along the x axis* this condition may be expressed as a local reconnection rate

$$\frac{E_o}{v_A B} = (1 - \alpha)^{\frac{1}{2}} \sin\chi\cos\chi \qquad 4\text{-}2$$

*Hill's analysis also includes an unspecified velocity component v_y of the incoming particles, which we have set equal to zero.

where $v_A \equiv B/\sqrt{\mu_0 \rho}$ and B are the Alfvén speed and magnetic field, respectively, in the region adjacent to the current sheet. Further, the anisotropy factor α of the incident particles is defined by

$$\alpha = 2\mu(p_{\parallel} - p_{\perp})_{in}/B^2 \qquad 4\text{-}3$$

Note that $\alpha = 0$ corresponds to isotropic pressure, $\alpha = 1$ to firehose conditions (taking account of the fact that the total plasma density ρ is twice the density, ρ_{in}, of the incident particles).

The preceding result is incomplete in that the angle χ must be a function of the reconnection rate. Also, the rate should refer to conditions on the x axis in the inflow region. Equation 4-2 is nevertheless interesting because it suggests that pressure anisotropy in the incoming plasma may be an important factor. In particular, it appears that reconnection may cease altogether for $\alpha = 1$.

Using the same approach, Hill has also considered the case of reconnection of fields of unequal magnitude and with a constant B_z component present. He obtains the formula 4-1 for the angular dependence of the reconnection potential.

The particle energization in a model of Hill's type is seen to be the direct result of inertia and gradient drifts in the current sheet, moving positive ions in the direction of the reconnection electric field, electrons in the opposite direction. It is also important to note that the energized plasma will be streaming out nearly parallel to the y axis, i.e., for small angles χ, nearly parallel to the magnetic field on both sides of the current sheet. By contrast, the symmetric fluid dynamical models yield an exit plasma flow that is perpendicular to the weak magnetic field in the two exit flow regions.

5. Fluid Description of the Diffusion Region

A complete theoretical treatment of the reconnection problem requires not only a self-consistent solution for the external flow, but also an internal, or diffusion-region, solution which describes the essential dissipative physical processes operating in that region, and which joins smoothly to the external solution. In the present section we review attempts to describe the diffusion region in terms of continuum equations which incorporate effects of plasma microinstabilities, if any, by means of an effective conductivity σ.

A brief discussion of one-fluid theories is given in section 5.1. In magnetospheric applications of reconnection, the collisional resistive length $(\mu_0 \sigma v_1)^{-1}$ is much smaller than relevant inner plasma scales, such as the electron inertial length. Thus one-fluid theory is directly applicable only if turbulent processes generate an effective resistive length which exceeds these inner scales. But even if that condition is not met, one-fluid theory serves the important purpose of providing an opportunity for a careful mathematical treatment in one region of plasma-parameter space.

The two-fluid description of the diffusion region is dealt with in sections 5.2 and 5.3. The former discusses the importance of the electron inertial length in determining the width $2x^*$ of the diffusion region when no collisional or turbulent resistivity is present. In the latter section, the importance of Hall currents and of the ion-inertial length and gyroradius are discussed.

5.1 One-Fluid Models

Detailed studies of the flow and field configuration in the diffusion region, using one-fluid magnetohydrodynamics, have utilized two

approaches: series expansion around the magnetic null point, and development of integral or "lumped" equations for the entire diffusion region. Additionally, certain exact solutions exist.

Series expansions have been given by Priest and Cowley[93] and by Cowley[27]. They found that in incompressible flow, and assuming analytic behavior at the null point, a plasma velocity of the form $v_x = -k_1 x$, $v_y = k_1 y$ to lowest order yields a magnetic field behavior of the form $B_x = k_2 y^3$ $B_y = k_3 x$, i.e., the field-line configuration at the neutral point is chi-like (χ) rather than ex-like (X). The latter type of configuration may however be obtained by assuming a third-order, rather than a first-order zero in $v_x(x)$ at $x = 0$. Furthermore, Yeh[128] has shown that the flow and field behavior near the null point need not be analytic. Logarithmic terms are possible in the expansion. Whether or not such terms are in fact present can be determined only by matching of the series expansion solution to an appropriate external solution, which has not yet been done. It also appears that the inclusion of compressibility effects will invalidate the above-mentioned result of Priest and Cowley. Finally, Cowley[27] has shown that series expansions yielding a non-constant field $B_z(x,y)$ are possible. But the question of whether such solutions may be matched to corresponding external solutions with non-constant B_z (see Ref. 25) has not been resolved.

The first lumped analysis of the diffusion region was performed by Parker[86] with application to Sweet's resistive current-sheet model[113] of a solar flare. The analysis yielded the following expression for the reconnection rate in incompressible flow

$$M_{A_1} \simeq (\mu_0 \sigma v_{A_1} y^*)^{-1/2} \qquad 5\text{-}1$$

This result is obtained directly from equations 2-5 and 2-7 with $\rho_2 = \rho_1$. Originally, the formula was used with y^* replaced by L, the half-length of the current sheet. It then describes field annihilation (see section 2.4) and yields a value of M_{A_1} that is far too small to account for an energy-release time of the order of 10^3 sec in a solar flare, or for the observed potential difference of 20-100 kV across the terrestrial magnetosphere. Later Petschek[90] used the formula 5-1, now with $2y^*$ representing the height of the diffusion region and with $y^* \ll L$, to describe the diffusion-region flow in his model. The basic point made by Petschek is that equation 5-1 determines, not the reconnection rate, but the height y^* of the diffusion region. In agreement with the discussion in section 2.3, y^* is then seen to be proportional to $M_{A_1}^{-2}$.

A more detailed lumped analysis was performed by Sonnerup[103] in an attempt to develop a diffusion region solution for the slow-mode reconnection geometry in figure 14. The treatment is incomplete because it does not take account of the momentum balance. Additional criticism has been offered by Vasyliunas[118] on the basis that the model implicitly assumes an abrupt switch from finite to infinite electrical conductivity at the outer edge of the diffusion region. Considering the extreme simplification of the external flow in this model, with slow expansion effects lumped into a single discontinuity (see section 4.1 and figure 14), this latter criticism is probably not of serious consequence. The following qualitative results of the analysis are likely to remain valid for the slow-mode reconnection model in figure 14:

(i) A relationship exists between M_{A_1} and the magnetic Reynolds number $R_y \equiv \mu_0 \sigma v_{A_1} y^*$ which is of the form given by Parker, equation 5-1, for small values of M_{A_1}, and which yields $R_y = 0$ for $M_{A_1} = (1+\sqrt{2})$. Thus, the dimensions of the diffusion region shrink toward zero as M_{A_1} approaches its maximum value.

(ii) When the slow-mode expansion in the inflow is concentrated into discontinuities as in figure 14, the increase in magnetic field and decrease in flow speed, described in sections 4.1 and 4.2 must occur in the outer portions of the diffusion region. Thus, as the origin is approached along the $\pm x$ axis, the magnetic field first increases then decreases to zero while $|v_x|$ decreases monotonically to zero. Thus, in the outer portion of the diffusion region the field behaves almost in a frozen-in manner. This may account for the large diffusion region width x^* found in the analysis[103,118].

Two exact solutions exist which describe flow near the magnetic null point. Yeh[129] obtained shock-free similarity solutions by assuming resistivity and viscosity to increase linearly with distance from the origin. It is not clear how such assumptions can be reconciled with an exterior solution in which dissipative effects are confined to shocks. A different type of exact solution has been discussed by Priest and Sonnerup[95,109]. It describes an incompressible two or three-dimensional MHD resistive stagnation-point flow at a current sheet. The field lines are straight and parallel to the current sheet. Thus, purely resistive magnetic field annihilation without reconnection occurs, as illustrated in Figure 21. These solutions represent a generalization of a case studied by Parker[88]. The resulting magnetic field profiles are shown in Figure 22. Three features are of interest. First, the characteristic width of the current layer is of the order of the resistive length as expected. Second, the frozen field condition applies at large $|x|$ values and leads to a gradual increase in the field magnitude $|B_y|$ as $|x|$ and $|v_x|$ decrease. As resistivity effects become increasingly important $|B_y|$ reaches a maximum and then decreases to zero at $x = 0$. This is precisely the behavior described in (ii) above. Third, a nonconstant value of B_z is possible.

5.2 Two-Fluid Effects; Electron Scale Lengths

In the two-fluid description of an electron-proton plasma, the ordinary Ohm's law is replaced by a generalized form (see, e.g., Rossi and Olbert[96]).

$$\underline{E} + \underline{v} \times \underline{B} = \frac{\underline{j}}{\sigma} + \frac{m_e}{ne^2}\left[\frac{\partial \underline{j}}{\partial t} + \underline{\nabla} \cdot (\underline{j}\,\underline{v} + \underline{v}\,\underline{j})\right] - \frac{1}{ne}\underline{\nabla} \cdot \underline{\underline{P}}_e + \frac{1}{ne}\underline{j} \times \underline{B} \qquad 5\text{-}2$$

As pointed out by Vasyliunas, the terms on the right-hand side may introduce a variety of plasma scale length's into the problem. The first term yields the resistive length $\lambda_r = (\mu_0 \sigma v_1)^{-1}$; the second set of terms yields the electron inertial length $\lambda_e = (m_e/\mu_0 ne^2)^{1/2}$; the off-diagonal terms of the electron stress tensor term $\underline{\underline{P}}_e$ yield the electron gyroradius; the last term, describing the Hall effect, yields $(v_A/v_1)\lambda_i$ where $\lambda_i = (m_i/\mu_0 ne^2)^{1/2}$ is the ion inertial length. The importance of the diagonal terms in $\underline{\underline{P}}_e$ has not been studied; with isotropic pressure and isentropic flow, these terms are cancelled identically by an electrostatic field.

To illustrate the effects of the electron inertia terms we now generalize the stagnation-point flows discussed by Sonnerup and Priest[109] to include two-fluid effects. Assuming incompressible flow and diagonal stress tensors for ions and electrons, the flow and fields are of the form

$$\left.\begin{aligned} \underline{v} &= -\hat{\underline{x}}k_1 x + \hat{\underline{y}}k_2 y + \hat{\underline{z}}k_3 z \\ \underline{B} &= \hat{\underline{y}}B_y(x) + \hat{\underline{z}}B_z(x) \end{aligned}\right\} \qquad 5\text{-}3$$

where the quantities k_1, k_2 and k_3 are constants such that $k_1 = k_2 + k_3$. These assumptions lead to a Bernoulli-type pressure integral

$$p = p_0 - \frac{1}{2}\rho\,(k_1^2 x^2 + k_2^2 y^2 + k_3^2 z^2) - (2\mu_0)^{-1}\,(B_y^{\,2} + B_z^{\,2})$$

of the momentum equation, and to the following components of the induction equation (the curl of equation 5-2).

$$\left.\begin{aligned} \lambda_e^2 k_1 \times B_y''' + (k_2\lambda_e^2 - \frac{1}{\mu_0\sigma})B_y'' - k_1 x B_y' - k_2 B_y = 0 \\ \lambda_e^2 k_1 \times B_z''' + (k_3\lambda_e^2 - \frac{1}{\mu_0\sigma})B_z'' - k_1 x B_z' - k_3 B_z = 0 \end{aligned}\right\} \quad 5\text{-}4$$

It is seen that only the electron-inertial length and the resistive length appear in equations 5-4. The Hall current term in 5-2 is curl-free and is cancelled exactly by a Hall electric field $E_x(x)$. Thus the ion inertial length does not appear. The solutions of equations 5-4 in the resistive limit are illustrated in Figure 22. In the inertial limit, the odd solutions are shown in Figure 23 for various values of $\alpha \equiv k_2/k_1$. As expected, the width of the magnetic-field reversal region is now of the order of the electron inertial length regardless of the flow rate. Further, it is observed that for $\alpha = 1$, i.e., for plane flow, the current density is logarithmically infinite at $x = 0$, a conclusion also drawn by Coroniti and Eviatar[22]. Thus some form of plasma microinstability or other effect by necessity must be present to reduce the current density to a finite value.

The off-diagonal terms in the electron pressure tensor will provide a finite electron gyroradius correction to the preceding results. When the electron gyroradius greatly exceeds λ_e, it will replace the electron inertial length as the minimum width of the layer[118]. However, a detailed calculation of these effects is difficult because the appropriate form of the off-diagonal pressure tensor terms is not known in a field reversal region of width comparable to the orbit scale.

Vasyliunas[118] pioneered the study of electron inertial effects in the diffusion region. In an approximate lumped analysis which neglected compressibility (later included by Coroniti and Eviatar[22]), off-diagonal stress

tensor elements, and Hall-current effects, he showed that the diffusion region in the inertial limit (i.e., neglecting resistivity) is hyperbolic in shape with width

$$x^*(y) \simeq (M_{A_1}{}^2\, y^2 + \lambda_e^2)^{1/2} \qquad 5\text{-}5$$

The exact analysis given here in all essential respects confirms Vasyliunas' results for small values of $M_{A_1}{}^2\, y^2/\lambda_e^2$. It also provides magnetic field profiles $B_y(x)$ whereas Vasyliunas neglected B_y within the diffusion region. Note that the formula (22) yields $x^*(0) \simeq \lambda_e$; $y^* \simeq \lambda_e/M_{A_1}$, in agreement with the behavior quoted in section 2.3 (between equations 2-8 and 2-9).

5.3 Two-Fluid Effects; Ion Scale Lengths

It is not clear how a diffusion region of the small physical dimensions implied by equation 5-5 can be joined to an external solution with slow shocks of thickness[20] comparable to, or greater than, the ion inertial length. Thus we are led to ask why the ion inertial length and the ion gyroradius did not appear in the previous analysis. The former is introduced via the Hall current term $\underline{j} \times \underline{B}/ne$ in equation 5-2. It may be cancelled by a Hall electric field only when it is curl free, which was the case for the stagnation-point flow in section 5.2. The ion gyroradius is introduced via the off-diagonal terms in the ion pressure tensor. We now demonstrate that in plane reconnection flow these effects imply the presence of Hall-current components j_x and j_y as well as a macroscopic flow $v_z(x,y)$ and a field component $B_z(x,y)$. Assuming $\partial/\partial z \equiv 0$ and omitting electron inertia and pressure terms for brevity, the z components of the momentum and induction equations are, respectively,

$$\rho v_x \frac{\partial v_z}{\partial x} + \rho v_y \frac{\partial v_z}{\partial y} = \frac{\partial}{\partial x} P_{xz} + \frac{\partial}{\partial y} P_{yz} + \frac{1}{\mu_0}\left(B_x \frac{\partial B_z}{\partial x} + B_y \frac{\partial B_z}{\partial y}\right) \qquad 5\text{-}6$$

$$(B_x \frac{\partial v_z}{\partial x} + B_y \frac{\partial v_z}{\partial y}) - (v_x \frac{\partial B_z}{\partial x} + v_y \frac{\partial B_z}{\partial y}) = - \frac{1}{\mu_0 \sigma} \nabla^2 B_z + \frac{1}{ne} (B_x \frac{\partial j_z}{\partial x} + B_y \frac{\partial j_z}{\partial y}) \quad 5\text{-}7$$

where the particle density n and the conductivity σ have been assumed constant. In the second equation, the Hall term (the last term on the right) is seen to be of the form $(\underline{B} \cdot \underline{\nabla} j_z)/ne$, i.e., it vanishes only when the current j_z remains constant along a field line. Otherwise it becomes a driving term in the second equation forcing values of v_z and B_z different from zero. Similarly, in the first equation it appears that the stress terms will usually force values of v_z and B_z different from zero. These terms are expected to introduce the ion gyroradius as a characteristic length into the problem. However, it is difficult to discuss these effects in detail, because the form of the stress terms in a thin field-reversal region is not known. Thus, we confine attention to the Hall current term, $(\underline{B} \cdot \underline{\nabla} j_z)/n_e$. While this term vanishes in the stagnation-point flow discussed in the previous section (and indeed along the x axis of any configuration), it cannot vanish throughout the diffusion region. We may estimate the magnitude of B_z by approximately equating the first term on the left and the second term on the right in equation 5-7:

$$v_x \frac{\partial B_z}{\partial x} \sim \frac{1}{ne} B_x \frac{\partial j_z}{\partial x}$$

With $v_x \sim v_1$, $B_x \sim B_2 \sim B_1 M_{A_1}$, $j_z \sim B_1/\mu_0 x^*$ and $\partial/\partial x \sim 1/x^*$ we then find

$$B_z/B_1 \sim \lambda_i/x^*$$

where $\lambda_i \equiv (m_i/\mu_0 ne^2)^{1/2}$ is the ion inertial length. Thus it appears that values of x^* much less than λ_i would give rise to unacceptably high values of B_z. A similar comparison between the terms $B_x \partial v_z/\partial x$ and $(ne)^{-1} B_x \partial j_z/\partial x$ yields $v_z/v_{A_1} \sim \lambda_i/x^*$. Again, $x^* \ll \lambda_i$ leads to an unacceptable result. Vasyliunas[119] has pointed out that off-diagonal electron pressure tensor

terms, not shown in equation 5-7, may possibly cancel the Hall term. However, there seems to be no obvious physical reason to expect such a cancellation. And problems with the off-diagonal stress tensor terms in equation 5-6 would still remain.

Detailed analysis of the effects described above is not available at present. However, a nonvanishing field component $B_z(x,y)$ would imply the presence of Hall currents $j_x = \mu_0^{-1} \partial B_z/\partial y$ and $j_y = -\mu_0^{-1} \partial B_z/\partial x$ in the diffusion region. The expected current flow and field pattern is shown schematically in Figure 24. The behavior of the B_y and B_z components indicated in the figure should be easy to identify in magnetic-field vector measurements from a satellite which crosses the diffusion region.

The reason for the appearance of the Hall current component j_x with the direction shown in figure 24 may be understood by noting that for $\sigma \simeq \infty$ the generalized Ohm's law (equation 5-2)implies that apart from electron-inertia and gyroradius effects the magnetic field is frozen into the electron component of the plasma. Thus the electrons flowing toward $x = 0$ are brought to rest over a distance of the order of the electron inertial length or the electron gyroradius. If the ions are similarly brought to rest over a distance comparable to the ion inertial length or the ion gyroradius, a relative motion of electrons and ions results, leading to currents j_x in the direction shown in the figure. Charge conservation then implies the presence of j_y as shown.

The z component of the force $\underline{j} \times \underline{B}$ associated with the Hall currents also leads to an acceleration of the plasma in the $\pm z$ direction. This effect is caused by drift and meandering motion of the ions in the current sheet. It may be interpreted as an ion current in the layer. Indeed, if $x^* \sim \lambda_i$, the principal current component j_z is carried by the ions; if $x^* \sim \lambda_e$ it is carried by the electrons.

It is evident that in the absence of plasma resistivity (classical or turbulent) the electron length scales must play a significant role in the diffusion region structure. But from the preceding discussion it appears possible that the ion length scales may determine the overall width $2x^*$ of the diffusion region while the electron length scales give the size of the detailed structures of j_x, j_y, and j_z near $x = 0$. From the preceding discussion, it is concluded that, even without plasma turbulence, the electromagnetic structure of the diffusion region may be far more complicated than previously assumed.

6. Non-Fluid Effects in the Diffusion Region

In magnetospheric and interplanetary applications of the reconnection process, collisional resistive effects in the diffusion region are negligible. Thus an effective resistivity in that region must derive either from inertia effects or from plasma turbulence. The former effects were dealt with in sections 5.2 and 5.3 from a fluid point of view. However, to develop a physical understanding of inertial phenomena in the diffusion region, it is useful to obtain an approximate expression for the effective inertial conductivity. This is done in section 6.1. Section 6.2 examines plasma instabilities which may generate steady-state turbulence in the diffusion region, but with details provided only for the ion-acoustic instability. Section 6.3 discusses several threshold effects that might be of importance for the onset of reconnection and for the identification of situations in which reconnection may not occur. Particle acceleration in nonsteady reconnection is discussed in section 6.4.

It will become quickly apparent that most of the material in this section is speculative in nature. Different processes may occur in different applications. It appears that no systematic effort has been made to sort out which mechanisms dominate in different parts of plasma parameter space.

In reading the present section, it will be useful to refer to the typical values of several physical parameters given in Table 1.

6.1 Inertial Resistivity

The concept of inertial resistivity was first discussed, in the context of magnetic field reconnection, by Speiser[108]. The basic idea is that a particle spends only a finite amount of time in the diffusion region and thus can pick up only a finite amount of energy from the electric field $E_0\hat{z}$. The inertial conductivity is written as

$$\sigma_{inert} = \frac{\overline{ne^2\tau}}{m} \qquad 6\text{-}1$$

Table 1 Physical parameters for various reconnection sites.

Quantity \ Site	Laboratory (Argon plasma)	Solar flare (pre-flare state) I	Solar flare (pre-flare state) II	Magneto-pause	Magneto-tail	Interpl. sector boundary
Number density $n_1 (m^{-3})$	1.5×10^{20}	10^{21}	10^{15}	2×10^{7}	10^{6}	5×10^{6}
Ion temperature T_{i_1} (°K)	10^{4}	10^{4}	10^{6}	10^{6}	10^{7}	5×10^{5}
Temperature ratio T_{e_1}/T_{i_1}	1	1	1	.5	.1	1
Magnetic field B_1 (Weber m^{-2})	1	5×10^{-2}	10^{-2}	3×10^{-8}	1.5×10^{-8}	5×10^{-9}
$\beta_{i_1} = 2\mu_0 n_1 k T_{i_1}/B_1^2$	5.2×10^{-5}	.14	3.5×10^{-4}	0.8	1.5	3.5
$\lambda_{coll} = (\mu_0 \sigma_{coll} v_1)^{-1} (m)$	7.8×10^{-3}	.40	2.0×10^{-6}	1.3×10^{-4}	4.2×10^{-5}	8.0×10^{-4}
$\lambda_{turb} = (\mu_0 \sigma_{turb} v_1)^{-1} (m)$	5.9×10^{-2}	.20	.94	2.5×10^{5}	5.6×10^{5}	1.9×10^{6}
$\lambda_e = (m_e/\mu_0 n_1 e^2)^{1/2} (m)$	4.3×10^{-4}	1.7×10^{-4}	1.7×10^{-1}	1.2×10^{3}	5.3×10^{3}	2.4×10^{3}
$\lambda_i = (m_i/\mu_0 n_1 e^2)^{1/2} (m)$	1.2×10^{-1}	7.3×10^{-3}	7.2	5.1×10^{4}	2.3×10^{5}	10^{5}

Notes: (1) 1 Weber/m^2 = 10^4 gauss = 10^9 γ

(2) $\sigma^{-1}_{coll} = 6.5 \times 10^3 \ ln\Lambda/T_{e_1}^{3/2}$ (Spitzer[112])

(3) σ_{turb} is given by equation 6-12.

(4) $v_1 = .1 v_{A_1} = .1 \ B_1/\sqrt{\mu_0 n_1 m_i}$

where $\bar{n}$ is the average particle density in the diffusion region and $\bar{\tau}$ is the effective time available for acceleration in the electric field. The formula 6-1 may be used with either the electron mass $m = m_e$ or the ion mass $m = m_i$ depending on whether the diffusion region current is principally an electron current or an ion current. Both cases may be treated the same way so that the particle mass m will be left without a subscript.

The average displacement, $\overline{\Delta z}$, along the electric field E_0, of a particle in the diffusion region may be obtained from a simple mass balance over a box of dimensions $2x^* \times 2y^* \times \overline{\Delta z}$:

$$2n_1 v_1 2y^* \overline{\Delta z} = \bar{n}\bar{v}_z 2x^* 2y^* = 2n_2 v_2 2x^* \overline{\Delta z} \qquad 6\text{-}2$$

where $\bar{v}_z$ is the average particle speed along the electric field $E_0\hat{z}$, and $\bar{v}_z = \overline{\Delta z}/\bar{\tau}$. As before, the subscripts 1 and 2 refer to conditions at the points $(x^*,0)$ and $(0,y^*)$, respectively. From the first equality in equation 6-2 we thus obtain

$$\bar{\tau} \simeq \frac{\bar{n}}{n_1}\frac{x^*}{v_1} \qquad 6\text{-}3$$

When this expression for $\bar{\tau}$ is substituted into equation 6-1 there results

$$\sigma_{\text{inert}} = \frac{\bar{n}^2 e^2 x^*}{n_1 v_1 m} \qquad 6\text{-}4$$

Thus the inertial conductivity is very high for low reconnection speeds, v_1, i.e., when the configuration approaches a current sheet. This is the behavior referred to in section 2.3.

Expressing the basic balance of field convection and diffusion as $\mu_0 \sigma_{\text{inert}} v_1 x^* = 1$, which is valid for small reconnection rates*, we find

$$x^* = \frac{n_1}{\bar{n}} \lambda_1 \qquad 6\text{-}5$$

where λ_1 is the inertial length $\lambda_1 = (m/\mu_0 n_1 e^2)^{1/2}$. For $m = m_e$ this result is in agreement with Vasyliunas' formula (equation 5-5). See also Coroniti

*With inertial resistivity, pure field annihilation is found to occur for $0 < M_{A_1} \leq \lambda_1/L$ (compare section 2.4).

and Eviatar[22]. But the calculation gives no clue as to whether the electron or the ion inertial length is to be used.

It should be realized that the value of x^* given by equation 6-5 represents a lower limit. The calculation assumed the time $\bar{\tau}$ available for free acceleration in the electric field to be equal to the residence time of a particle within the control box. In reality $\bar{\tau}$ must always be less than the residence time because the magnetic field does not vanish within the entire box. Thus the inertial conductivity is less than that given by equation 6-4 and x^* is correspondingly larger than in equation 6-5.. When diamagnetic currents become important, i.e., for $\beta_1 = 2\mu_0 p_1/B_1^2 > 1$, it may be shown that x^* gradually approaches a magnitude of the order of the gyroradius instead.

The previous estimate of x^* applies only for small values of the reconnection rate. To understand this fact, we note that the expression $\mu_0 \sigma v_1 x^* = 1$, which was used in obtaining equation 6-5, derives directly from Ohm's law in the approximate form $j_z = \sigma E_0 = \sigma v_1 B_1$ with $j_z \simeq \mu_0^{-1} \partial B_y/\partial x \simeq B_1/\mu_0 x^*$. For large reconnection rates it becomes important to incorporate the term $\underline{v} \times \underline{B}$ in the electric field, as well as the curvature term $\partial B_x/\partial y$ in the expression for j_z. The latter effect leads to a multiplicative factor $(1 - M_{A_1}^2 \rho_1/\rho_2)$ on the right-hand side of equation 6-5 (see equation 6-7 below) so that the size of the diffusion region decreases toward zero as the reconnection rate approaches its maximum value. Thus, for large reconnection rates, the required width of the diffusion region may be substantially less than the relevant plasma scale (the inertial length or gyroradius, depending on β_1). It is difficult to reconcile such a situation with the nature of the particle orbits in that region. Therefore, it is conceivable that steady-state reconnection with inertial resistivity as the dominant effect in the diffusion region is not possible for large reconnection rates.

The average electrostatic particle energization in the diffusion region may be obtained directly from the first equality in 6-2:

$$\overline{\mathcal{E}} = e\overline{\Delta z}\, E_0 = \overline{ne}\overline{v}_z x^* \frac{E_0}{n_1 v_1} \qquad 6\text{-}6$$

But we also have $\overline{ne}\overline{v}_z = \overline{j}_z = \mu_0^{-1}\,(\partial B_y/\partial x - \partial B_x/\partial y)$. Approximating $\partial B_y/\partial x$ by B_1/x^*, $\partial B_x/\partial y$ by B_2/y^*, noting that $B_2/B_1 = v_1/v_2 \simeq \rho_2 x^*/\rho_1 y^*$ we find

$$j_z = \overline{ne}\overline{v}_z \simeq (B_1/\mu_0 x^*)\left(1 - \frac{\rho_1}{\rho_2} M_{A_1}^2\right) \qquad 6\text{-}7$$

and from 6-6

$$\overline{\mathcal{E}} \simeq \frac{B_1^2}{\mu_0 n_1}\left(1 - \frac{\rho_1}{\rho_2} M_{A_1}^2\right) \qquad 6\text{-}8$$

where the relations $E_0 = v_1 B_1$ and $v_1/v_2 \simeq v_1/v_{A_1} = M_{A_1}$ have been used. This formula agrees with equation 2-11. Maximum acceleration occurs for small reconnection rates and densities: for magnetospheric conditions $\overline{\mathcal{E}} \sim 1 - 10$ *keV*. It is emphasized that equation 6-8 represents the average energy gain. A small number of particles moving nearly along the reconnection line may gain larger amounts of energy in the electric field E_0.

6.2 Plasma Turbulence

A variety of plasma instabilities may serve to generate plasma turbulence in the diffusion region and an associated turbulent conductivity σ_{turb}. We now discuss such effects in an assumed quasi-steady state of reconnection. Onset effects are dealt with in section 6.3.

The tearing instability, either in its collisional resistive version [45,97,117] or in the collision-free electron-inertial version [17,30,44,61,69], has been studied intensely in the context of reconnection. It generates a pattern of alternating *X* type and *O* type magnetic neutral points in a current sheet. But most analyses of this instability pertain to current-sheets with a vanishing magnetic-field component B_x perpendicular to the

sheet. In other words, in the present application either the reconnection rate M_{A_1} is very small or the magnetic field behavior in the diffusion region is chi-like (χ) rather than ex-like (X), as discussed in section 5.1. Schindler[99] has pointed out that for $B_x \neq 0$ the collision-less tearing instability may still proceed as long as the gyro-period $\tau_g = 2\pi m/eB_x$ of a particle in the field B_x exceeds the instability growth time $\tau_0 \simeq (x^*/v_{th})(x^*/R_L)^{3/2}$ where v_{th} is the thermal speed and R_L the gyroradius. This condition may be applied either to electrons (electron tearing) or ions (ion tearing). In rough terms, non gyrotropic behavior of the particles is required for these instabilities to be possible. While the nature or existence of steady-state tearing turbulence does not appear to have been established, one cannot exclude the possibility that such turbulence could be of importance in the diffusion region[19,21,47].

Parker[89] has suggested that interchange instability may serve to enhance the flow rates in Sweet's[113] current sheet model. In the geomagnetic tail, the instability would be driven, not in the diffusion region itself, but rather by the pressure gradient and field curvature in the near-earth section of the tail plasma sheet (see figure 4). A detailed analysis, including the impeding effects of the ionosphere, has been given recently by Kan and Chao[66]. It indicates growth times of the order of a few hours with ionospheric coupling, a few minutes without such coupling. The situation relative to the level of steady-state turbulence is not clear.

Huba *et al*[64] have proposed that the lower-hybrid-drift instability may provide anomalous resistivity in the diffusion region. It appears that the threshold for this instability is sufficiently low to permit the diffusion region width to be of the order of the ion inertial length.

Haerendel[50] has discussed the possibility that the electron-cyclotron

drift instability, which has a current threshold somewhat less than that of the ion-acoustic instability, may generate turbulence in a diffusion region of width $2x^*$ equal to a few electron-inertial lengths. However, its importance has been questioned by Coroniti and Eviatar[22] on the basis that the gyrocoherence required by the instability may not be available in the diffusion region. They also note that when the electron drift speed exceeds the threshold for the ion-acoustic instability the electron-cyclotron drift mode goes over nonlinearly to the ion-acoustic one.

The ion-acoustic instability has been proposed[9,22,41,101] as a likely agent for the generation of turbulence in the diffusion region. It will be dealt with in some detail in the remainder of this subsection. This is done for illustrative purposes and not as an indicator of a universal preference for this particular mechanism. On the other hand, the ion-acoustic instability appears in fact to occur in laboratory reconnection experiments[9,85]. But it is probably not relevant to magnetospheric reconnection.

For a current-driven instability such as the ion-acoustic one to occur, the current density in the diffusion region must exceed a certain minimum value, corresponding to a critical current velocity v_c, i.e., $j > nev_c$. If Hall currents are present, as discussed in section 5.3, the total current must be considered. Here we shall confine attention to the component j_z. According to equation 6-7 we then find

$$j_z = \frac{B_1}{\mu_0 x^*}\left(1 - \frac{\rho_1}{\rho_2} M_{A_1}^2\right) \geq \bar{n}ev_c \qquad 6\text{-}9$$

where, for the ion-acoustic instability

$$v_c = \sqrt{\frac{kT_i}{m_e}}/f(T_e/T_i) \qquad 6\text{-}10$$

Combination of equations 6-9 and 6-10 yields

$$x^*/\lambda_e \leq \sqrt{2/\beta_i}\,\left(1 - \frac{\rho_1}{\rho_2} M_{A_1}^2\right) f(T_e/T_i) \qquad 6\text{-}11$$

where $\beta_i \equiv 2\mu_0 \bar{n} k T_i / B_1^2$. The function $f(T_e/T_i)$ is shown in Figure 25. It is seen that $f(T_e/T_i)$ is of the order unity for $T_e \simeq T_i$ so that for small M_{A_1}, and for β_i of order unity or less, the critical diffusion region width is of the order of the electron inertial length. For large values of β_i, x^* must be considerably less than λ_e suggesting that only a subportion of the diffusion region may contain ion-acoustic turbulence (compare the logarithmic singularity in j_z at $x = 0$, discussed in section 5.2). Coroniti and Eviatar[22] indicate $\beta < 5$ as a condition for their analysis to remain valid. For greater β values, the size of the turbulent region approaches the wave-length of the ion-acoustic turbulence. For M_{A_1} of order unity the diffusion region must also be very small for ion-acoustic turbulence to occur.

For high temperature ratios T_e/T_i, the instability may occur for x^* considerably larger than λ_e but probably not as large as λ_i (see figure 25). A large temperature ratio may perhaps be generated temporarily by electron run-away in a current sheet at the onset of reconnection (see next subsection). For example, in the double inverse pinch experiment the collisional resistive length considerably exceeds λ_e (see Table 1), so that run-away must occur to initiate the ion-acoustic instability. But it appears unlikely that a large temperature ratio T_e/T_i could be sustained on a steady basis in a diffusion region of width much greater than λ_e since most parts of such a region the run-away would have to occur transverse to a strong magnetic field. We conclude that steady-state ion-acoustic turbulence, driven by the current component j_z, is unlikely to be important unless the diffusion region width, $2x^*$, is of the order of the electron inertial length. At the same time it is observed that the Hall current component j_y discussed in section 5.3 (figure 24) may be sufficiently intense to drive the instability in parts of a diffusion

region of total width comparable to the ion inertial length.

Coroniti and Eviatar[22] have examined the question of the turbulent saturation of the ion acoustic instability in detail. They conclude that the current velocity will remain close to the threshold value given by equation 6-10. The resulting weak steady-state turbulence is adjusted to give the value of turbulent conductivity required to satisfy $\mu_0\sigma_{turb}v_1x^* \simeq 1$ with x^* given by the equality in equation 6-11. On the other hand, common estimates of the effective electrical conductivity associated with the ion-acoustic instability, in a state of turbulent saturation, such as (see, e.g., references *40*, and *115*).

$$\sigma_{turb} = \frac{\bar{n}e^2}{m_e}\left(\frac{10^2}{\omega_{pe}}\frac{\sqrt{kT_i/me}}{(j/ne)}\sqrt{\frac{T_i}{T_e}}\right) \qquad 6\text{-}12$$

where $\omega_{pe} = (\bar{n}e^2/\varepsilon_0 m_e)^{1/2}$, give a much too low value of the conductivity, even at the critical current velocity $j/\bar{n}e = v_c$. In other words, with reasonable reconnection speeds and with x^* satisfying 6-11, one finds $\mu_0\sigma_{turb}\, v_1x^* \ll 1$, which is impossible in a steady state. In Table 1, this fact is manifested by the inequality $\lambda_{turb} \gg \lambda_e$, where λ_{turb} is the turbulent resistive length.

6.3 Onset of Rapid Reconnection

There is ample observational evidence relating to solar flares, to the earth's magnetotail, and to the double inverse pinch experiment, to indicate that occasionally rapid reconnection is switched on in an abrupt, almost explosive manner. At the earth's magnetopause, if reconnection actually occurs there, the switch-on appears more gentle and may be a direct consequence of the interplanetary field turning southward so that the angle between the reconnecting fields exceeds some critical value (compare section 4.4). It is natural to assume that the more explosive events might be associated with a plasma instability and/or an abrupt decrease in the effective con-

ductivity in a current sheet or in the diffusion region of a slowly reconnecting configuration. Five such possibilities, all speculative at present, are mentioned below:

(i) Thermal instability. It has been proposed[16,19,56] that the flash phase of a solar flare may be associated with a thermal instability. For example, explosive solutions of the electron energy equation, i.e., solutions which yield an infinite temperature in a finite time, are known to occur when collisional Joule dissipation dominates the equation. This instability is not relevant for magnetospheric applications or for the upper solar atmosphere (case II, in Table 1).

(ii) Beta threshold. It may be hypothesized[106] that, in a collision-free plasma, reconnection is suppressed for high β_1 values but may occur for small β_1. Thus, any current sheet in which β_1 decreases gradually from some initially large value may be converted to a rapidly reconnecting configuration when a critical β_1 value is reached. In the geomagnetic tail, an abrupt decrease in β_1 value occurs if the plasma sheet in which the tail current sheet is imbedded shrinks to a thickness equal to the current sheet thickness. On the other hand, at the subsolar magnetopause, Lees[71] and Zwan and Wolf[132] have described a magnetosheath plasma depletion mechanism (by escape along the magnetic field lines) which would tend to maintain a value of β_1 of order unity or less. The β_1 threshold is not relevant to the double-inverse pinch laboratory experiment, and probably not to solar flares because β_1 is small in these applications (Table 1).

(iii) Current-threshold. Assume that a current sheet with little or no reconnection gradually thins from an initial width of an ion gyroradius or more toward the electron inertial length, in response to an increased externa

total pressure, $p_1 + B_1^2/2\mu_0$. In this process the current density in the sheet increases gradually. When the threshold for onset of current-driven plasma instabilities is reached, e.g., for the ion-acoustic instability, when equation 6-11 is satisfied, a reduction in effective electrical conductivity takes place in the layer. If this reduction occurs sufficiently rapidly, the inductance of the system will allow us to consider the current density initially to remain essentially unchanged. Instead an inductive electric field $E_z(x,t)$ is developed within the sheet to maintain the current density. The magnitude of this electric field is larger, the larger the conductivity reduction.

For a turbulent conductivity of the size used by Coroniti and Eviatar[22], E_z is of the size usually estimated for steady-state reconnection. This electric field, which is initially confined to the current sheet, is subsequently spread by fast-mode expansion waves propagating outward from the sheet as the configuration converts itself to one of steady or quasi-steady reconnection.

If the conductivity is reduced to the level given by equation 6-12, E_z may be one or two orders of magnitude larger than typical steady-state values. Smith[102] has pointed out that the width $2x^*$ of the current sheet must then increase. As pointed out in section 6.2, for reasonable flow rates we find $\mu_0\sigma_{turb}v_1x^* \ll 1$ when $x^* \simeq \lambda_e$ and with σ given by equation 6-12. Thus an increase in x^* occurs in order to bring $\mu_0\sigma v_1x^*$ toward unity as required in a steady state. The rate $-\partial B/\partial t$ associated with the increase in x^* is the principal source of E_z. But the main result of the increase in x^* is that condition 6-11 ultimately is violated so that the ion-acoustic instability is quenched. Smith[102] proposes that the process may then repeat itself. A state of pulsating reconnection is established. See also Bratenahl and Baum[11]. While the above arguments were given in terms of the ion-acoustic

instability, other mechanisms may produce similar effects.

(iv) Tearing threshold. In a collision-free current sheet with a vanishing normal magnetic field component, electron tearing should be normally present, unless it is suppressed by some agent such as pressure anisotropy[18] or velocity shear[60]. With a nonvanishing normal magnetic field component B_x, a threshold for the onset of collision-free tearing does exist, as mentioned in section 6.2. If $|B_x|$ is originally large, no tearing occurs. But as $|B_x|$ gradually decreases it will set in when the gyroperiod in B_x exceeds the growth time. Schindler[99,100] has noted that this threshold may be exceeded for ions (but not electrons) in the geomagnetic tail current sheet during the thinning of that sheet which occurs in the expansive phase of the geomagnetic substorm. Since the tail at this time has free energy available for dissipation the ultimate result of the onset of ion tearing should be a large-scale relaxation (via reconnection) of the tail towards a state of minimum free energy, rather than merely the generation of tearing turbulence in the sheet. Further development of the ion-tearing instability theory has been given by Galeev and Zeleny[46,47].

(v) Interchange instability. An abrupt onset of interchange turbulence in the geomagnetic tail[66,89] may occur if the ionosphere becomes decoupled from the tail plasma sheet by the development of electric fields parallel to the magnetic lines of force.

6.4 Particle Acceleration

One of the most important, and at the same time most poorly understood, aspects of magnetic-field reconnection is its presumed ability to accelerate particles to high energies. Observations in the magnetospheric

tail indicate the occurrence of energetic electron and proton bursts[6,68,116] during times when reconnection may be going on. And it should be remembered that our ability to observe reconnection on the sun and in the far reaches of cosmos depends critically on the generation of energetic particles and on the electromagnetic radiation they subsequently produce.

Particle acceleration may occur either in turbulent small-scale electric fields or in the large-scale reconnection electric field E_z. Both types of acceleration are expected to be operative principally in high-current regions: the diffusion region and the shocks. To discuss turbulent acceleration one must understand the nature of the dominant micro-processes in these regions. Since no such understanding is at hand, the discussion in this section is confined to particle acceleration in the large-scale reconnection electric field.

In many, but not all, cosmic applications, the total potential difference associated with a steady reconnection electric field is sufficiently high to account in principle for observed particle energies. However, it is only in the small diffusion region that particles have the opportunity to move along the electric field for any considerable distance. And even there, most particles have short residence times and undergo a correspondingly small energization, as shown by equation 6-8. Thus, steady-state reconnection does not appear to be an effective mechanism for the acceleration of particles to very high energies[105]. Additionally, in applications such as the geomagnetic tail it is necessary to account for particle energies which exceed the steady-state cross-tail voltage by an order of magnitude or more. One is therefore led to consider the possibility of particle acceleration during nonsteady reconnection[116]. Two possible advantages are gained. First, the inductive electric fields may, in principle at least, become much stronger than the quasi-

static ones during steady reconnection. Second, the nonconservative nature of $\underline{E}$ permits acceleration within more localized regions of space. For example, betatron acceleration to high energy may occur in a small region of space where the particles experience a large increase in magnetic field intensity. By contrast electrostatic acceleration requires particles to move large distances along the separator.

The lack of nonsteady reconnection models prevents a detailed analysis of particle acceleration. But the simple model given below may serve as an illustration of how electron energization might occur in the diffusion region. A resistively decaying one-dimensional current sheet, perhaps generated as described in section 6.3, may be crudely described by

$$\underline{B} = \begin{cases} \hat{y}\, B_1 \dfrac{x}{x^*} & |x| < x^* \\ \hat{y}\, B_1 \dfrac{|x|}{x} & |x| > x^* \end{cases} \qquad 6\text{-}13$$

where the sheet width x^* is an increasing function of time and B_1 is the constant field outside the sheet. Assuming no inflow into the sheet, the associated electric field is

$$\underline{E} = \hat{z}\frac{1}{2} B_1 \left(1 - \frac{x^2}{x_*^2}\right) \frac{dx^*}{dt} \qquad 6\text{-}14$$

The direction of this field is such that it drives the particles toward the center of the current sheet (x=0). A particle accelerating freely at x=0 in this electric field may be shown to gain an amount of energy given by

$$\Delta\mathcal{E} = mc^2 \left\{\sqrt{1+\left(\frac{eB_1\Delta x^*}{2mc}\right)^2} - 1\right\} \qquad 6\text{-}15$$

provided it doesn't leave the system (at $z = \pm h$) during the time it takes the current sheet to widen by Δx^*. In the geomagnetic tail $B_0 \simeq 20nt$, and for $\Delta x^* = 500$ *km* equation 6-15 predicts a possible energy gain of 1.07 *MeV* for electrons (protons would gain a similar amount of energy only if $\Delta x^* \sim 15000$ *km*)

Since an electron in this energy range transverses the entire tail in less than a second, it would appear that an unreasonably large value of dx^*/dt is needed. But this is not necessarily so. If the widening current sheet is located along the separator *AXB* of the reconnection bubble in figure 12, electrons may be accelerated as they move from *A* to *B* along the separator *AXB*. They may then return from *B* to *A* by gradient drift in the vicinity of the *O* type neutral line *BOA* where the electric field vanishes. Subsequently they reenter the acceleration region at *A*. By cycling electrons through this loop many times the energy gains predicted by equation 6-15 may be achieved even for small values of dx^*/dt.

The illustrative example discussed above emphasizes that it may be necessary to consider three-dimensional time-dependent configurations in order to account for particle acceleration in the reconnection process. For further illustrative calculations, see refs. *56, 116*.

7. Magnetospheric Evidence

Much of the observational evidence concerning the possible occurrence of reconnection in the magnetosphere has been summarized by Burch[12]. Relevant references may be found in his paper and are not, for the most part, repeated here.

A large amount of evidence exists indicating a relationship of various magnetospheric activity indices to the southward component of the interplanetary magnetic field. Also, spatial asymmetries in a variety of polar-cap processes appear to be correlated with the orientation of the interplanetary magnetic field. Such evidence is compatible with, but does not prove, the occurrence of reconnection at the magnetopause. This body of observations will not be discussed here. Instead, we focus, in section 7.1, on observations relating directly to the transfer of magnetic flux from closed to open field lines, and vice versa, in the magnetosphere. If such transfer in fact occurs, reconnection of some form must take place. If not, there is no need for it. Section 7.2 contains a brief discussion of direct measurements of magnetic field and plasma in the vicinity of what may have been reconnection sites.

7.1 Flux Transfer Evidence

The case for the occurrence of flux transfer in the magnetosphere from closed to open field lines is based on four sets of observations, discussed below:

(i) Existence of open field lines in the tail.

Anderson and Lin[5] have studied the shadowing effects on solar electrons ($\mathcal{E} > 20$ kev), produced by the moon when it is located in the geomagnetic tail. They provide persuasive evidence that a substantial amount of magnetic

flux in the two tail lobes occurs on open field lines, i.e., on lines that intersect the earth's surface in one place only. But the observations do not establish how large a fraction of the tail flux is on open lines at a given distance from the earth. Thus, it is not known for a fact how large a fraction of the earth's polar-cap field lines, i.e., lines emerging at latitudes above the auroral oval, that are open. A popularly held view is that all are open. But Heikkila[53], questioning the soundness of this view, has drawn attention to observations by McDiarmid et al.[77,78] which indicate the common occurrence of trapped particle pitch-angle distributions in the day-side cusp region as well as poleward of discrete auroral arcs.

(ii) Flux erosion from the front-lobe magnetosphere.

The magnetopause is observed to move closer to the earth when the interplanetary field develops a southward component. At the same time, the day-side polar cusp moves to lower latitudes. These effects cannot be accounted for by simple compression of the magnetosphere. Maezawa[79] has estimated that flux on closed field lines is removed from the magnetosphere front lobe in an amount estimated at about 10^8 weber during a typical event. Either this flux is transferred to open field lines in the polar cap by dayside magnetopause reconnection or it is moved into the tail while remaining on closed field lines. In the latter case, the flux might be added to the lobe of closed field lines in the tail or it might possibly be placed on open field lines by reconnection at the tail magnetopause. The popularly held view is that the flux is transferred to open flux by reconnection somewhere on the dayside magnetopause.

(iii) Flux addition in the open tail lobes.

A substantial body of evidence indicates that the magnetic field intensity in the tail starts to increase shortly after the onset of a southward component of the interplanetary magnetic field while at the same time the asymptotic tail cross section increases[80]. The observed concurrent gradual thinning of the tail plasma sheet (which is believed to contain the closed tail field lines) argues against these effects being caused by an increase of flux on closed field lines in the tail. Rather they indicate an increase of flux on open field lines in the two tail lobes by an amount estimated at $1 - 2.5 \times 10^8$ weber. If the auroral oval (and the dayside cleft) is associated with the separatrix between closed and open field lines, the motion of this oval to lower latitudes following the southward turning of the interplanetary field[62] supports this interpretation. But the evidence, while strong, is not conclusive. If closed field lines occur in the tail outside (i.e., above and below) the plasma sheet, the flux on open field lines could conceivably remain unchanged.

(iv) Polar cap electric fields and convection.

Electric field measurements[55] in the polar ionosphere indicate an average voltage difference across the polar cap of the order of 65 kVolt, corresponding to a magnetic flux transport across the cap from dayside to nightside at a rate of about 2×10^8 weber/hour[80]. Ion flow measurements[52] over the polar cap show flow patterns that carry particles and, unless $\underline{E} \cdot \underline{B} \neq 0$, magnetic flux poleward across the day-side cleft in a narrow longitude sector. Most but not all of the flow in the cap region occurs near the equatorward edge of the cap adjacent to an abrupt flow reversal, below which the return flow to the dayside occurs. While the exact

location of the separatrix between closed and open field is not known, it is difficult to locate it in such a way that these results do not imply a transfer of flux from closed to open field lines.

In spite of the ambiguities in the interpretation of the observations listed above, their mutual consistency in terms of flux transfer rates is impressive and lends credence to the idea that flux transfer from closed to open field lines does occur in the magnetosphere. However, a far greater body of simultaneous observations by satellites at different locations in the magnetosphere needs to be examined in order to establish the validity of the idea in a conclusive manner. It is noted that on the average, any flux transfer from closed to open field lines must be balanced by a reverse transfer from open to closed lines. Tail reconnection, occurring sporadically in connection with the expansive phase of magnetic substorms, is thought to accomplish this latter transfer but conclusive evidence is not available (see section 7.2).

7.2 Measurements Near Reconnection Sites

In a strict sense, direct evidence for reconnection consists of *in-situ* observations of the hyperbolic magnetic field configuration associated with a separator and an electric field along that line. The electric field observation may convincingly be replaced by the observation of plasma energized in the reconnection process (compare sections 2.4 and 4.2).

Hones et al[63] have reported observations of proton fluxes and of magnetic fields in the geomagnetic tail at geocentric distances in the range of 25-32 earth radii. They have found substorm events in which tailward proton flows at speeds up to 1000 km/s and an associated southward component of the magnetic field occurred during the storm expansion phase, followed by earthward flow and a northward field component during recovery. Such observations

are consistent with a separator moving tailward past the satellite. However, evidence concerning the magnetic-field component perpendicular to the tail current sheet is not entirely convincing unless the field is measured near the center of the sheet, which was not the case. And recently Lui et al[74,75] have challenged observations purporting to show the formation of a near earth reconnection line during substorms. Observations of proton jetting in the tail[39], of energetic particle bursts[6,68,98], and of lunar shadow patterns of electron fluxes[73], while generally compatible with tail reconnection, nevertheless cannot be claimed to provide unambiguous proof of the occurrence of the process.

At the dayside magnetopause, magnetic field components perpendicular to the magnetopause have been observed[108], although not as a permanent feature not even when the magnetosheath field opposes the terrestrial one. The narrow jets of energized plasma, predicted by magnetopause reconnection models (e.g., figure 19) and flowing nearly tangential to the magnetopause, have not been seen*, even though satellites such as HEOS 2 have had the right position and attitude to observe them[50,51]. These facts along with recent observations of a plasma boundary layer inside the dayside magnetopause[29,35] suggest that magnetopause reconnection, if it occurs, may be more sporadic and more localized than originally expected. Furthermore, the possibility of reconnection in the cusps and elsewhere on the magnetopause surface, rather than near the sub-solar point, needs to be examined[50].

The absence of observations of plasma energized by dayside reconnection has led Heikkila[54] to suggest that no such reconnection occurs, i.e., that the magnetopause is an electrostatic equipotential. This suggestion is

*However, a layer of energetic electrons of unknown origin has been discovered outside the tail magnetopause[7,81].

difficult to reconcile with the presence of magnetic field components perpendicular to the magnetopause, unless one is willing to accept potential differences of the order of 50 kvolt along field lines extending from the magnetopause into the solar wind; or unless one argues that such perpendicular components are never present over any substantial part of the dayside magnetopause.

8. Summary and Recommendations

In this paper we have given a reasonably detailed description of the present status of our understanding of reconnection. The picture that emerges is of a process, simple in concept but extremely complicated and multifaceted in detail. Nonlinear magnetohydrodynamic processes in the external flow region, governed by distant boundary conditions, are coupled to non-linear microscopic plasma processes in the diffusion region in a manner not clearly understood. And it appears that reconnection may operate in entirely different ways for different plasma parameters and for different external boundary conditions. Steady reconnection may be allowed in some cases, forbidden in others, with intermediate situations involving impulsive or pulsative events.

On the whole, our theoretical and empirical knowledge of reconnection is poor. Yet the process plays a key role in solar-flare theory as well as in our present concept of the dynamic magnetosphere. And it appears as an unwanted feature in tokamaks and other fusion configurations. These facts, along with the potential importance of reconnection in other parts of cosmos, amply justify vigorous research efforts related to reconnection in the following five areas: solar-flare and astrophysical observations, magnetospheric observations, laboratory experiments, computer simulation, and analytical model building. The first area, while extremely important, is too broad to be commented upon here. In the remaining areas the following recommendations are made:

Magnetospheric observations and experiments should include:

(i) A coordinated program to establish (or deny) the occurrence of flux transfer across separatrix surfaces, and to study other global consequences of reconnection.

(ii) Direct observations of magnetic field, plasma, energetic particles, and fluctuating as well as steady electric fields, near magnetospheric reconnection sites. Multi-satellite missions are needed to separate spatial and temporal effects.

(iii) Perhaps active experiments, such as the release of barium clouds near reconnection sites.

Laboratory experiments. The observation of impulsive flux transfer events and of ion-acoustic turbulence in the double inverse pinch experiment illustrates the importance of such experiments in shaping our understanding of reconnection. Yet, (excluding fusion devices) the double inverse pinch appears to be the only operating reconnection experiment in the U.S. today. A substantially expanded laboratory program is needed with four principal goals:

(i) Simulation of solar-flare reconnection.

(ii) Simulation of magnetospheric reconnection.

(iii) Study of basic plasma processes of importance in reconnection, such as slow-shocks and anomalous resistivity.

(iv) Exploration of reconnection in plasma heating devices.

Computer simulation provides a potentially very powerful tool for the study of reconnection. Magnetohydrodynamic codes, and ultimately self-consistent particle-fields codes should be developed. It is particularly important to build into such simulations the effects of inertial and anomalous resistivity in the diffusion region.

Analytical models of reconnection should emphasize the following interrelated problems:

(i) Nonsteady and three-dimensional effects.

(ii) Plasma processes in the diffusion region.

(iii) Particle acceleration.

(iv) Reconnection of fields that are not antiparallel.

It is through vigorous activities in the aforementioned areas, and effective interaction between scientists involved in them, that our understanding of the reconnection process may be most rapidly advanced. To bring about such a state of affairs, two proposals are made:

(A) That a special working group be assembled with the charge of promoting effective research on all aspects of the reconnection problem and with membership drawn from the five research areas discussed above.

(B) That NASA and other funding agencies develop coordinated programs of support for reconnection research.

The importance of the reconnection concept is such that we can ill afford the present somewhat haphazard approach to its study.

Acknowledgment

Research supported by the National Science Foundation under Grant ATM 74-08223 A01 and by the National Aeronautics and Space Administration under Grant NSG 7292 to Dartmouth College.

The author has benefitted from comments on the paper by: P. J. Baum, A. Bratenahl, W. J. Heikkila, F. W. Perkins, E. R. Priest, and V. M. Vasyliunas. Conversations with A. Bratenahl and W. J. Heikkila have enlightened the author on many aspects of reconnection and its possible application in the magnetosphere.

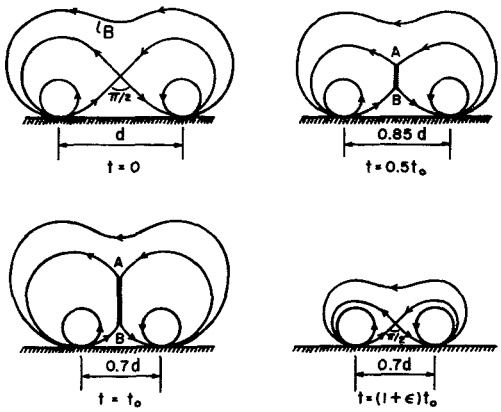


FIGURE 1 Qualitative time sequence for two dipoles moving toward each other on the solar surface. A current sheet *A-B* develops during time $0<t<t_o$. Rapid reconnection sets in at $t = t_o$ and relaxes the configuration toward a potential field in the short time ϵt_o.

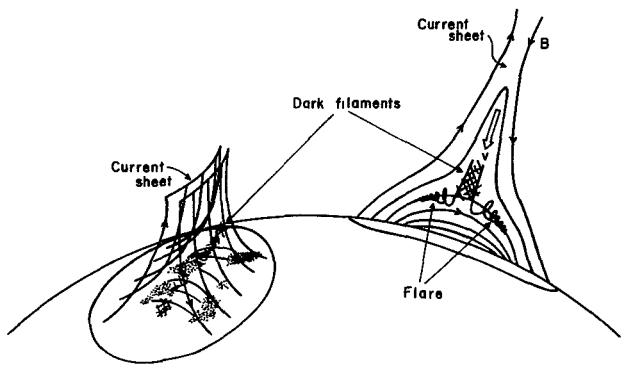


FIGURE 2 Current-sheet formation caused by the stretching of magnetic loop on the sun (after Carmichael[14]).

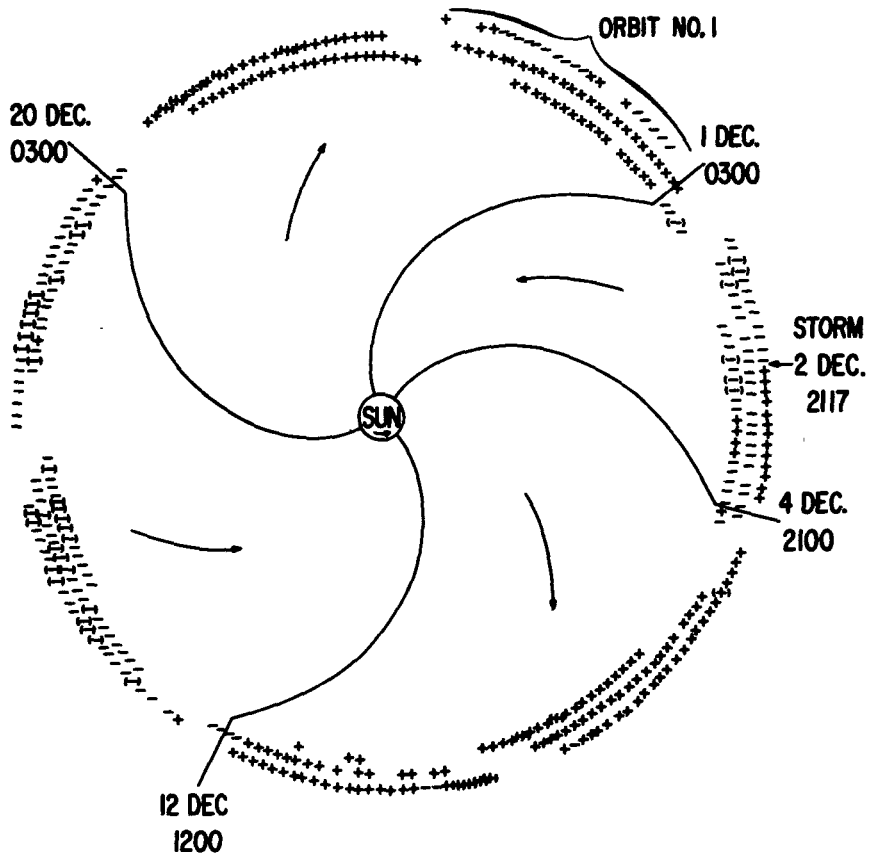


FIGURE 3 Sector structure of the interplanetary magnetic field in the ecliptic plane as observed by IMP-1 in 1963. Positive and negative signs indicate the direction of the measured interplanetary magnetic field away and toward the sun, respectively (Wilcox and Ness[123]).

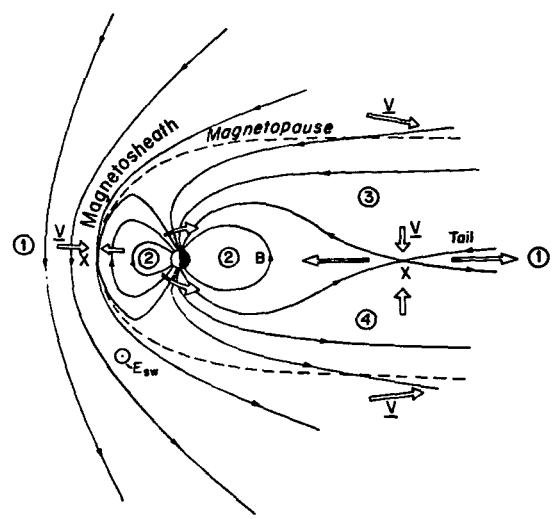


FIGURE 4 The earth's magnetosphere with magnetosheath magnetization due south.

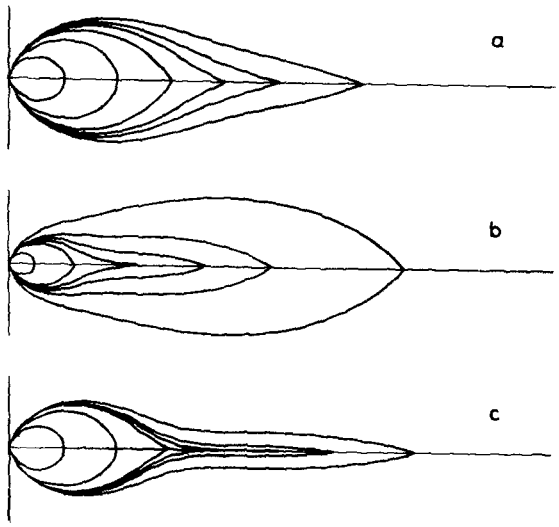


FIGURE 5 Magnetic field configurations for a rapidly spinning magnetosphere containing low-energy plasma (Gleeson and Axford[48]).

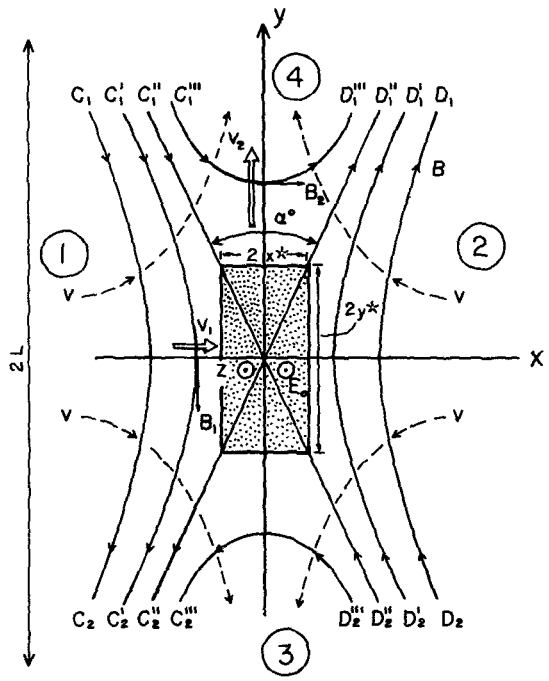


FIGURE 6 Basic plane reconnection configuration. Solid lines are magnetic field lines; dashed lines are streamlines. The shaded region at the center is the diffusion region.

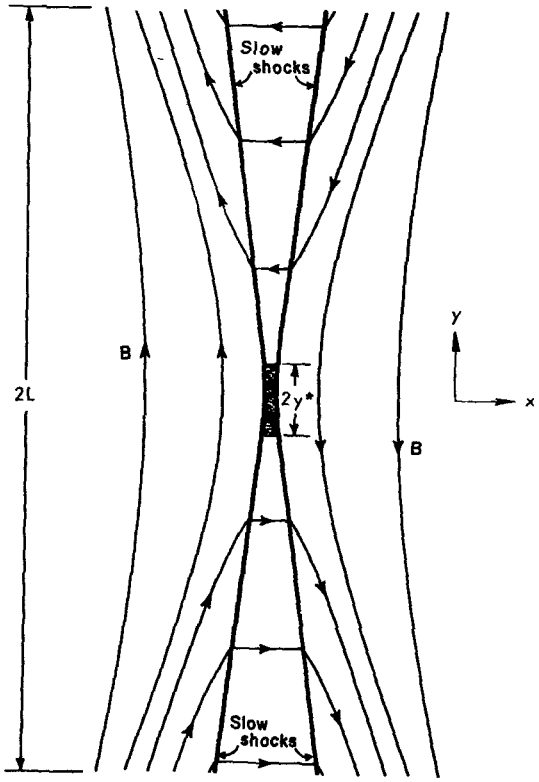


FIGURE 7 Configuration of slow MHD shocks in the reconnection geometry (after Petschek[90]).

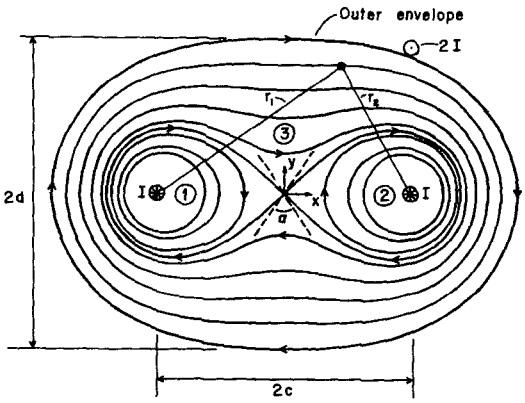


FIGURE 8 Field configuration in double inverse pinch experiment (after Bratenahl and Baum[11]).

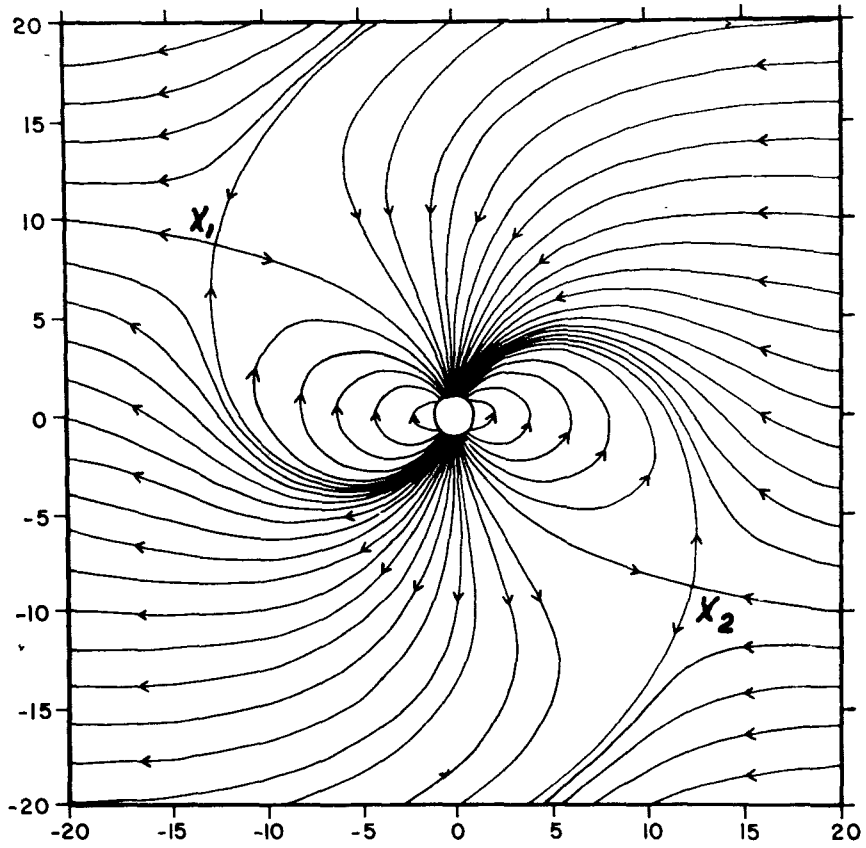


FIGURE 9 Field lines in the plane of the neutral points X_1 and X_2 for a uniform magnetic field and a dipole field (Cowley[24]). Dipole moment vector at right angles to the uniform field.

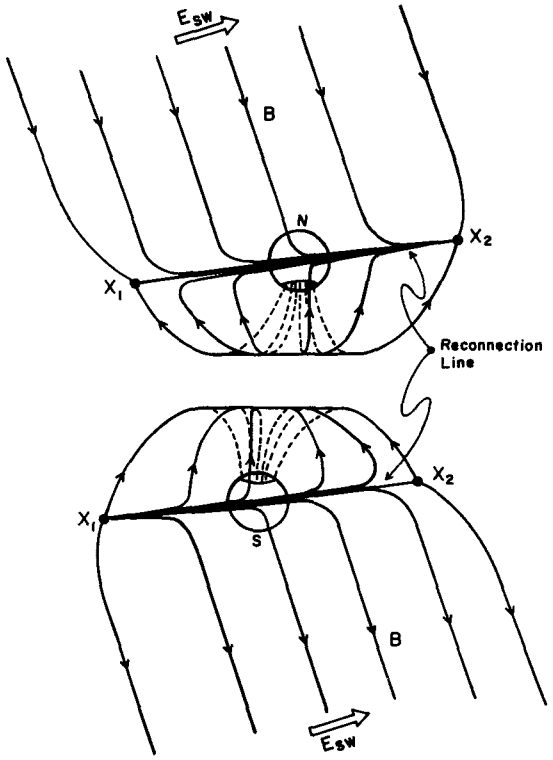


FIGURE 10 Schematic of separatrix surfaces for magnetopause reconnection. Lower figure shows separatrix of the null point X_1; upper figure that of X_2. The two figures are to be superimposed.

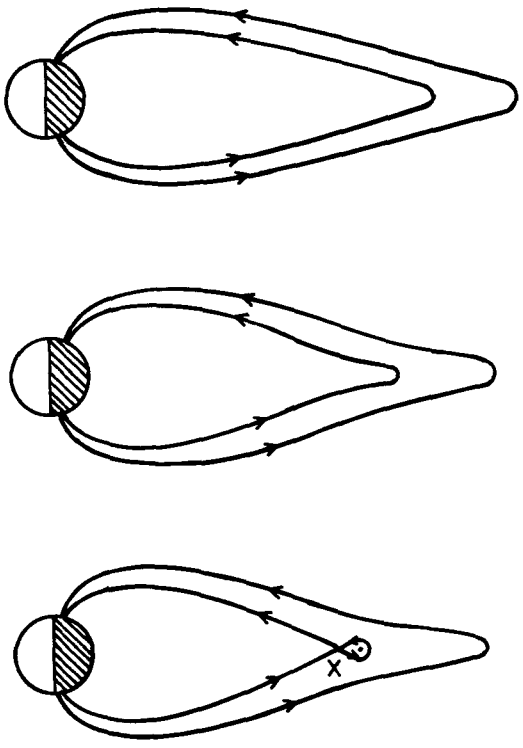


FIGURE 11 Formation of reconnection bubble in the geomagnetic tail. Schematic of field configuration in the noon-midnight meridional plane.

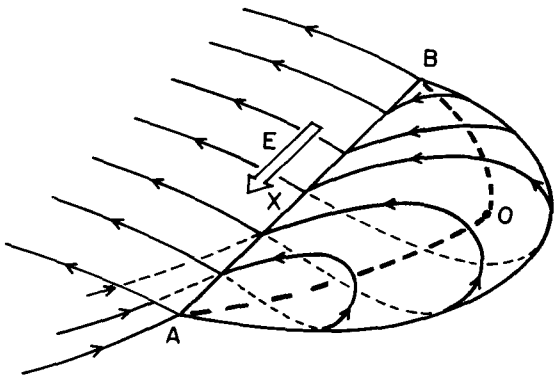


FIGURE 12 Three-dimensional sketch of reconnection bubble with the reconnection line along *AXB* and an *O*-type magnetic null line along *AOB*.

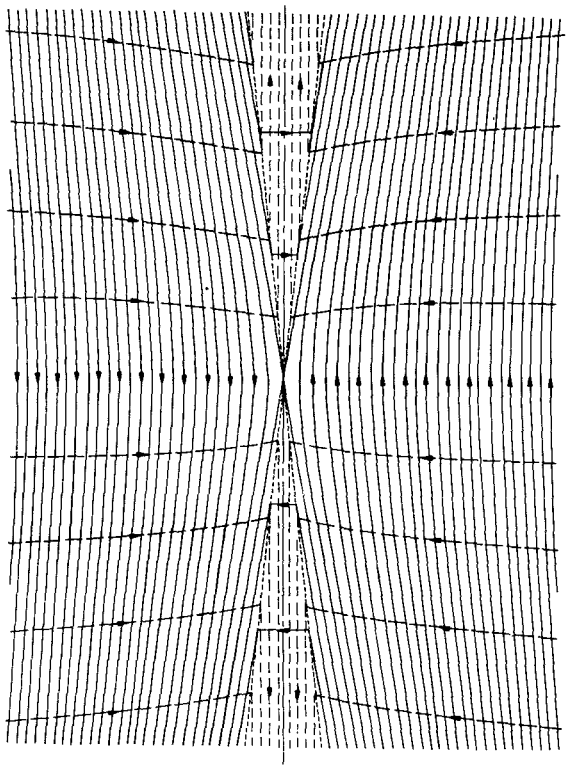


FIGURE 13 Petschek's reconnection model with $M_{A_\infty} = .1$. Magnetic field lines (solid lines) and streamlines (broken lines) are shown (Vasyliunas[118]). Fast-mode expansion in the entry flow.

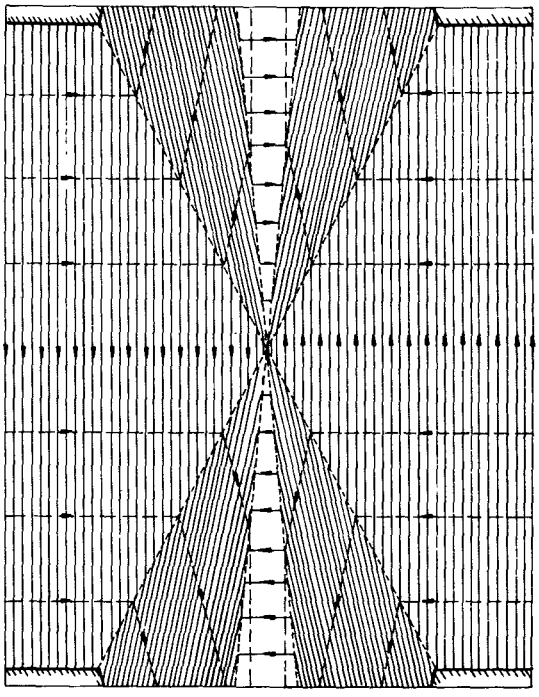


FIGURE 14 Slow-mode reconnection model with $M_{A_1} = .5$. (Vasyliunas[118]). Slow-mode expansion in the inflow is concentrated to waves from four external corners.

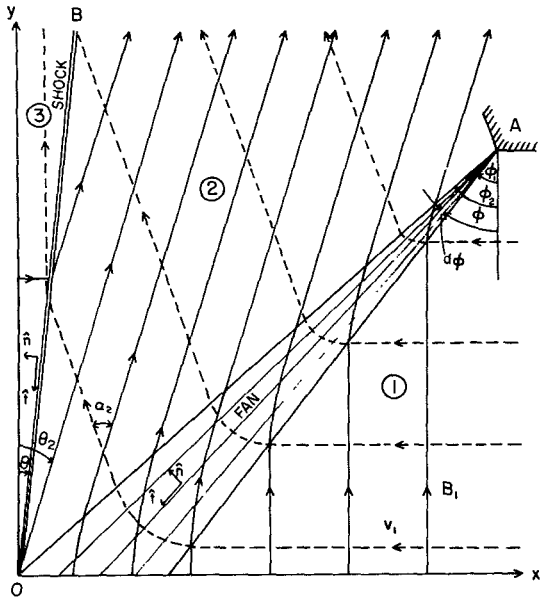


FIGURE 15 One quarter of a symmetric compressible slow-mode magnetic-field reconnection model for $M_{A_1} = 0.7$ and $\beta_1 = 2\mu_0 p_1/B_1^2 = 5$ (Yang and Sonnerup[125]).

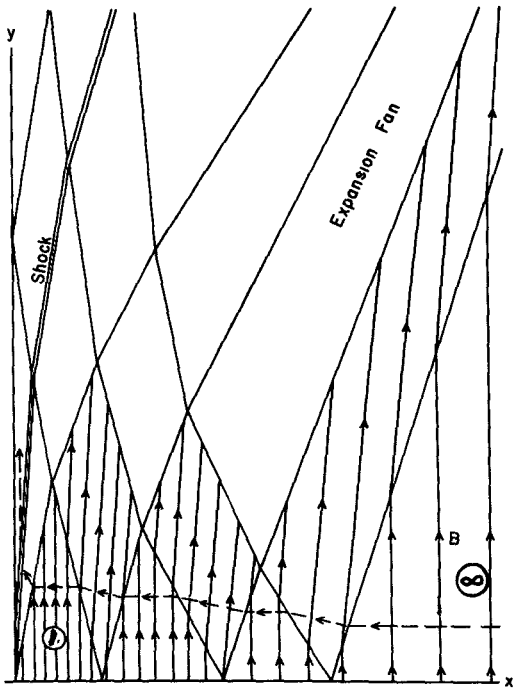


FIGURE 16 Schematic showing "reflection" of slow mode expansion fan in the x axis (Yang and Sonnerup[125]).

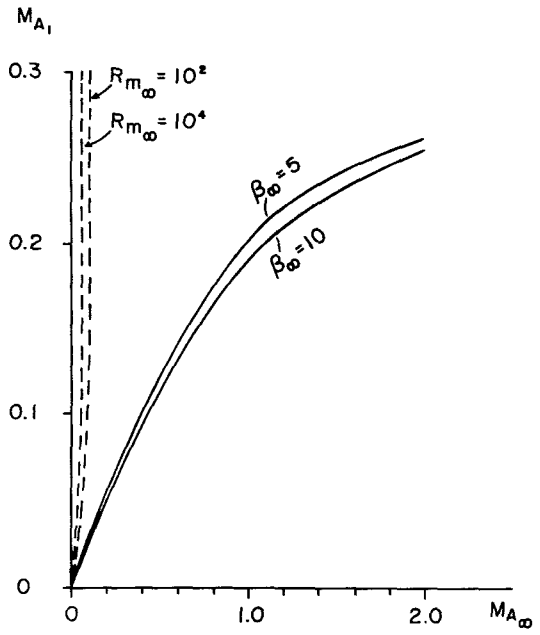


FIGURE 17 Relationship between the inflow Alfvén numbers M_{A_∞}, far upstream, and M_{A_1}, adjacent to the diffusion region. Solid curves refer to slow-mode expansion model,[125] dashed curves to the Soward-Priest[110] analysis of the fast-mode expansion model.

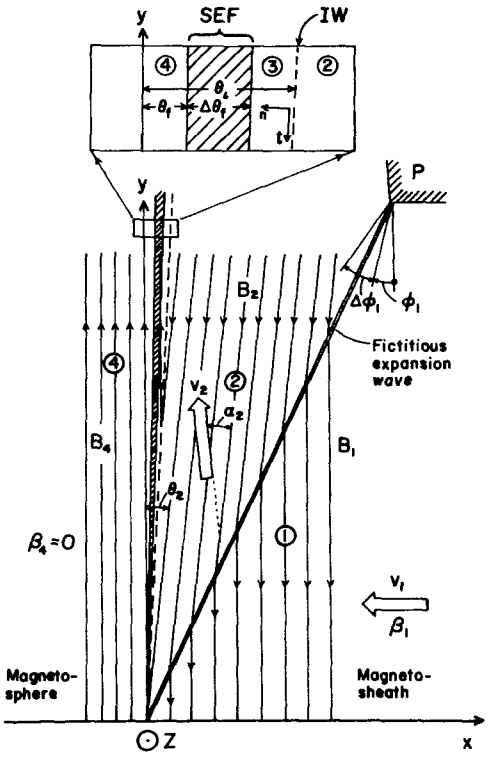


FIGURE 18 Upper half of compressible slow-mode model of magnetopause reconnection for $M_{A_1} = .2$ and $\beta_1 = 2\mu_0 p_1/B_1^2 = 2$. The intermediate wave (IW) marking the magnetopause is shown as a dashed line. The slow mode expansion fan (SEF) is shaded (Yang and Sonnerup[126]).

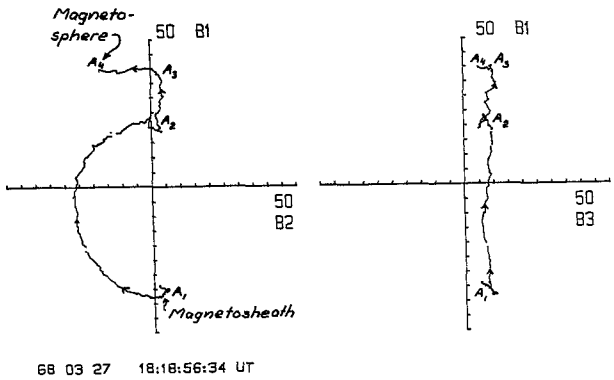


FIGURE 19 Polar plots of magnetic field at the magnetopause. Left hand figure shows the field components B_1 and B_2 tangential to the magnetopause during an OGO-5 crossing; right hand figure shows the nearly constant magnetic-field component B_3 normal to the magnetopause. The field is given in units of γ ($1\gamma = 1$nt). Intermediate wave or rotational discontinuity is segment A_1 - A_2 of the left-hand trace; the slow expansion fan is segment A_2 - A_3. The segment A_3 - A_4 may be caused by a finite gyroradius effect not contained in the MHD model.

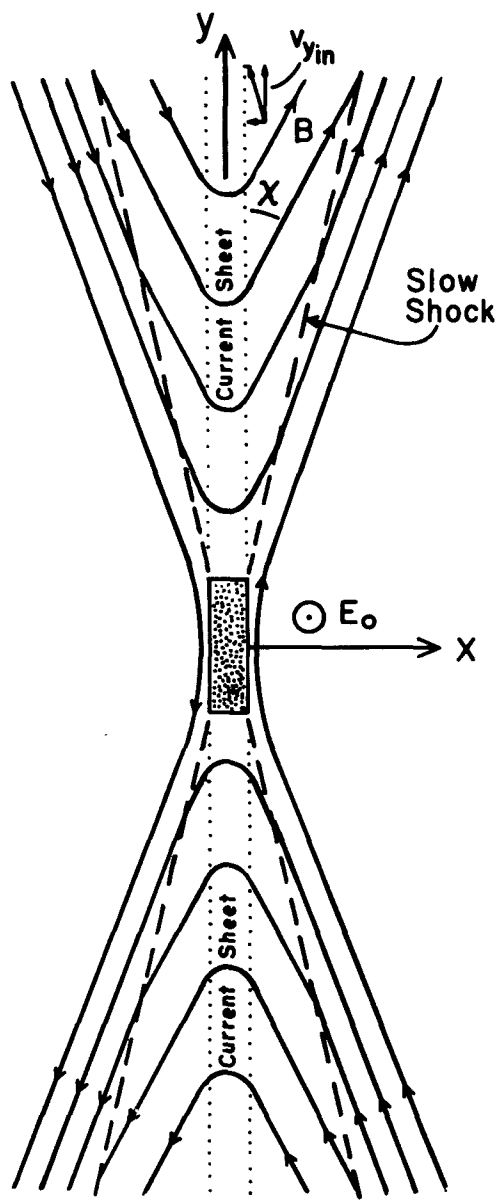


FIGURE 20 Hill's (53) collisionless reconnection model. The magnetic-field change across the slow shock becomes weak for $\beta = 2\mu_o p/B^2 \to 0$ with the principal field reversal occurring in a current layer at $X = 0$.

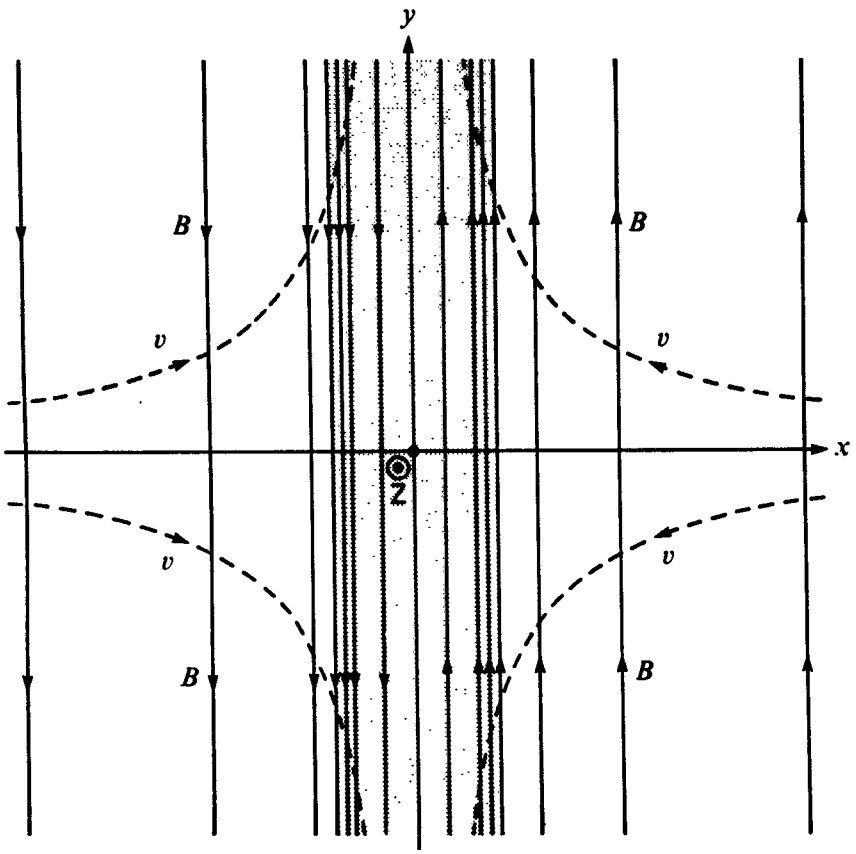


FIGURE 21 Magnetic field lines and streamlines for stagnation point flow, $\underline{v} = (-k_1 x, k_2 y, k_3 z)$, at a current sheet. The diffusion-dominated region is shaded (Sonnerup and Priest[109]).

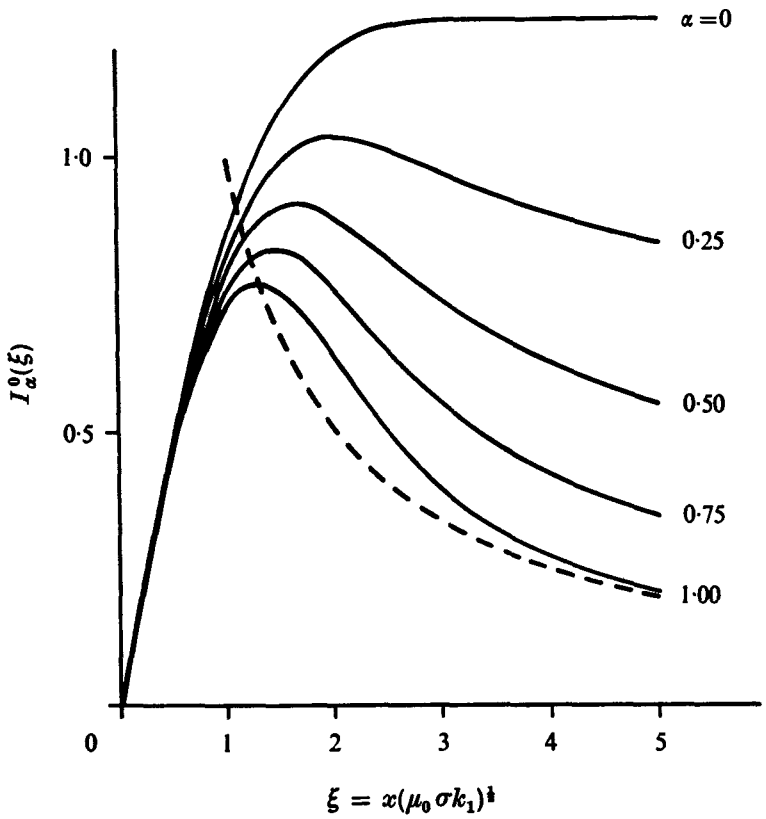


FIGURE 22 Nondimensional magnetic-field profiles $I^0_\alpha = B_y(E_o\sqrt{\mu_o\sigma/k_1})^{-1}$ for the configuration in Figure 21 in the resistive limit ($\alpha = k_2/k_1$; $k_3 = k_1 - k_2$). Plane stagnation point flow for $\alpha = 1$, axisymmetric flow for $\alpha = \frac{1}{2}$. The frozen-field profile for $\alpha = 1$ is shown by dashed line (Sonnerup and Priest[109]).

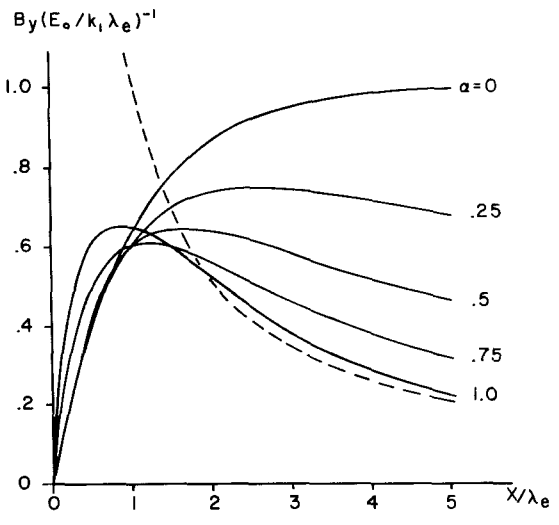


FIGURE 23 Nondimensional magnetic-field profiles for the configuration in Figure 21 in the electron-inertial limit [$\alpha = k_2/k_1$; $k_3 = k_1 - k_2$; $\lambda_e = (m_e/\mu_o ne^2)^{\frac{1}{2}}$]. Plane stagnation-point flow for $\alpha = 1$; axisymmetric flow for $\alpha = \frac{1}{2}$. The frozen-field profile for $\alpha = 1$ is shown by dashed line.

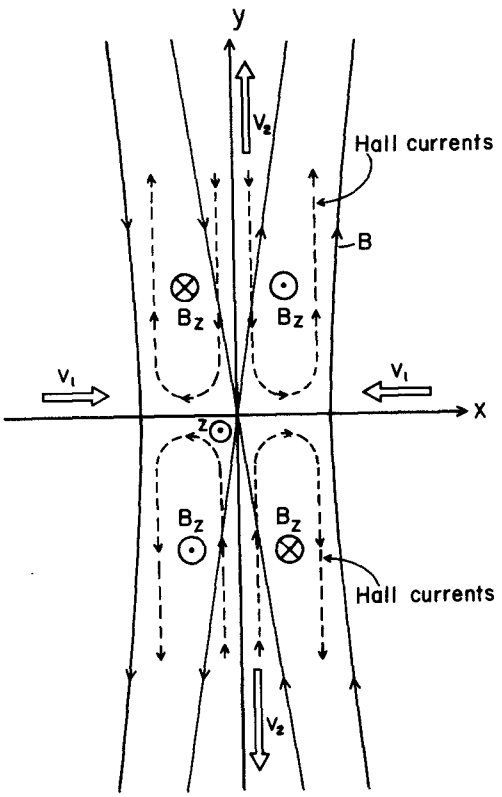


FIGURE 24 Schematic picture of Hall current loops in the diffusion region. Also shown is the transverse magnetic field $B_z(x,y)$ induced by these currents.

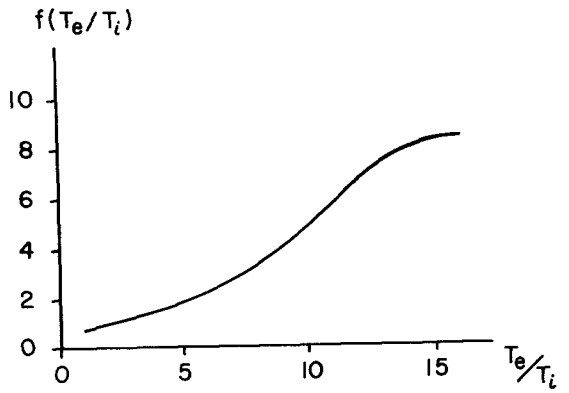


FIGURE 25 Function $f(T_e/T_i) = (kT_i/m_e)^{1/2}/v_c$ where v_c is the critical current velocity for onset of ion-acoustic instability (after Fredricks[40]).

REFERENCES

1 Alfvén, H., Some properties of magnetospheric neutral surfaces, J. Geophys. Res., 73, 4379, 1968.

2 Alfvén, H., Electric current structure of the magnetosphere, in Physics of the Hot Plasma in the Magnetosphere, ed. B. Hultquist and L. Stenflo, New York, 1975.

3 Alfvén, H., On frozen-in field lines and field-line reconnection, J. Geophys. Res., 81, 4019, 1976.

4 Amano, K., and T. Tsuda, Reconnection of magnetic field lines by clouds-in cells plasma model, J. Geomagn. Geoelec. (Japan), 29, 9, 1977.

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

6 Baker, D. N. and E. C. Stone, Observations of energetic electrons in the earth's magnetotail: Plasma sheet and fireball observations, J. Geophys. Res., 82, 1532, 1977.

7 Baker, D. N. and E. C. Stone, The magnetopause electron layer along the distant magnetotail, Geophys. Res. Lett., 4, 133, 1977.

8 Birmingham, T. J., Field line motion in the presence of finite conductivity, in Cosmic Plasma Physics, ed., K. Schindler, Plenum Press, New York, 1972.

9 Bratenahl, A., and C. M. Yeates, Experimental study of magnetic flux transfer at the hyperbolic neutral point, Phys. Fluids, 13, 2696, 1970.

10 Bratenahl, A., and P. J. Baum, Impulsive flux transfer events and solar flares, Geophys. J., Royal. Astr. Soc., 46, 259, 1976.

11 Bratenahl, A., and P. J. Baum, On flares, substorms, and the theory of impulsive flux transfer events, Solar Phys., 47, 345, 1976.

12 Burch, J. L., Observations of interactions between interplanetary and geomagnetic fields, Revs. Geophys. and Space Phys., 12, 363, 1974.

13 Burton, R. K., R. L. McPherron, and C. T. Russell, The terrestrial magnetosphere: a half-wave rectifier of the interplanetary electric field, Science, 189, 717, 1975.

14 Carmichael, H., A process for solar flares, AAS-NASA Symposium on solar flares, W. N. Hess, ed., NASA SP-50, p. 451, 1964.

15 Chapman, S., and P. C. Kendall, Liquid instability and energy transformation near a magnetic neutral line: a soluble non-linear hydromagnetic problem, Proc. Roy. Soc., A, 271, 435, 1963.

16 Coppi, B., Remarks on sequence-of-plasma-instabilities models of solar flares, Astrophys. J., 195, 545, 1975.

17 Coppi, B., G. Laval, and R. Pellat, Dynamics of the geomagnetic tail, Phys. Rev. Lett., 16, 1207, 1966.

18 Coppi, B., and M. N. Rosenbluth, Model for the earth's magnetic tail, In "Stability of Plane Plasmas", Spec. Publ. 36, Eur. Space Res. Org., Paris, 1968.

19 Coppi, B., and A. B. Friedland, Processes of magnetic energy conversion and solar flares, Astrophys. J., 169, 379, 1971.

20 Coroniti, F. V., Laminar wave train structure of collisionless magnetic slow shocks, Nuclear Fusion, 11, 261, 1971.

21 Coroniti, F. V., On the nonlinear evolution of the collisionless tearing mode, UCLA Plasma Physics Group Report PPG-299, May 1977.

22 Coroniti, F. V., and A. Eviatar, Magnetic-field reconnection in a collisionless plasma, Ap. J. Suppl., 33, 189, 1977.

23 Cowley, S. W. H., A self-consistent model of a simple magnetic neutral sheet system surrounded by a cold collisionless plasma, Cosmic. Electrodynamics, 3, 448, 1973a.

24 Cowley, S. W. H., A qualitative study of the reconnection between the earth's magnetic field and an interplanetary field of arbitrary orientation, Radio Sci., 8, 903, 1973b.

25 Cowley, S. W. H., On the possibility of magnetic fields and fluid flows parallel to the *X*-line in a re-connexion geometry, J. Plasma Physics, 12, 319, 1974a.

26 Cowley, S. W. H., Convection-region solutions for the re-connexion of antiparallel magnetic fields of unequal magnitude in an incompressible plasma, J. Plasma Phys, 12, 341, 1974b.

27 Cowley, S. W. H., Magnetic field-line reconnexion in a highly conducting incompressible fluid: Properties of the diffusion region, J. Plasma. Phys., 14, 475, 1975.

28 Cowley, S. W. H., Comments on the merging of nonantiparallel magnetic fields, J. Geophys. Res., 81, 3455, 1976.

29 Crooker, N. U., Explorer 33 entry layer observations, J. Geophys. Res., 82, 515, 1977.

30 Cross, M. A., and G. Van Hoven, High conductivity magnetic tearing instability, Phys. Fluids, 19, 1591, 1976.

31 Dailey, C. L., H. A. Davis, and R. H. Lovberg, An experimental investigation of magnetic field annihilation, Bull. Am. Phys. Soc., 17, 1011, 1972.

32 Dungey, J. W., Interplanetary magnetic field and the auroral zones, Phys. Rev. Letters, 6, 47, 1961.

33 Dungey, J. W., The structure of the exosphere or adventures in velocity space, in Geophys., The Earth's Environment, edited by C. De Witt, J. Hieblot, and L. Le Beau, pp. 503-550, Gordon and Breach, New York, 1963.

34 Dungey, J. W., Waves and particles in the magnetosphere, in Physics of the Magnetosphere, R. L. Carovillano, J. F. McClay, and H. R. Radoski, eds., p. 218, D. Reidel, Dordrecht, Netherlands, 1968.

35 Eastman, T. E., E. W. Hones, Jr., S. J. Bame, and J. R. Asbridge, The magnetospheric boundary layer: Site of plasma, momentum and energy transfer from the magnetosheath into the magnetosphere, Preprint LA-UR 76-1303, Los Alamos Scientific Lab., to appear Geophys. Res. Lett., 1976.

36 Eastwood, J. W., Consistency of fields and particle motion in the Speiser model of the current sheet, Planet. Space Sci., 20, 1555, 1972.

37 Eastwood, J. W., The warm current-sheet model and its implications on the temporal behavior of the geomagnetic tail, Planet. Space Sci., 22, 1641, 1974.

38 Francfort, Ph., and R. Pellat, Magnetic merging in collisionless plasmas, Geophys. Res. Lett., 3, 433, 1976.

39 Frank, L. A., K. L. Ackerson, and R. P. Lepping, On hot tenuous plasmas, fireballs, and boundary layers in the earth's magnetotail, J. Geophys. Res., 81, 5859, 1976.

40 Fredricks, R. W., Electrostatic heating of solar wind ions beyond 0.1 AU, J. Geophys. Res., 74, 2919, 1969.

41 Friedman, M., and S. M. Hamberger, On the neutral point region in Petschek's model of magnetic field annihilation, Ap. J., 152, 667, 1968.

42 Fukao, S., and Tsuda, T., On the reconnection of magnetic lines of force, J. Plasma Phys., 9, 409, 1973a.

43 Fukao, S., and Tsuda, T., Reconnexion of magnetic lines of force: evolution and incompressible MHD fluids, Planet. Space Sci., 21, 1151, 1973b.

44 Furth, H. P., The mirror instability for finite particle gyroradius, Nuclear Fusion Suppl., Pt. 1, 169, 1962.

45 Furth, H. P., J. Killeen, and M. N. Rosenbluth, Finite resistivity instabilities of a sheet pinch, Phys. Fluids, 6, 459, 1963.

46 Galeev, A. A., and L. M. Zeleny, The role of plasma processes in the dynamics of magnetospheric substorms, Paper presented at International Symposium on Solar-Terrestrial Physics, Boulder, Colorado, June, 1976.

47 Galeev, A. A., and L. M. Zeleny, Magnetic reconnection in a space plasma, Academy of Sciences of the USSR, Space Research Institute, Report D-249, Moscow, 1977.

48 Gleeson, L. J., and W. I. Axford, An analytic model illustrating the effects of rotation on a magnetosphere containing low-energy plasma, J. Geophys. Res., 81, 3403, 1976.

49 Gonzalez, W. D., and F. S. Mozer, A quantitative model for the potential resulting from reconnection with an arbitrary interplanetary magnetic field, J. Geophys. Res., 79, 4186, 1974.

50 Haerendel, G., Microscopic plasma processes related to reconnection, J. Atm. Terr. Phys., 39, , 1977. (to appear).

51 Haerendel, G., G. Paschmann, N. Sckopke, and H. Rosenbauer, The front side boundary layer of the magnetosphere and the problem of reconnection, J. Geophys. Res., 83, , 1978. (to appear).

52 Heelis, R. A., W. B. Hanson, and J. L. Burch, Ion convection velocity reversals in the dayside cleft, J. Geophys. Res., 81, 3803, 1976.

53 Heikkila, W. J., Magnetospheric plasma regions and boundaries, In "Physics of the Hot Plasma in the Magnetosphere", B. Hultquist and L. Stenflo, eds., Plenum Publ. Co., New York, p. 69, 1975a.

54 Heikkila, W. J., Is there an electrostatic field tangential to the dayside magnetopause and neutral line?, Geophys. Res. Lett., 2, 154, 1975b.

55 Heppner, J. P., High latitude electric fields and the modulations related to interplanetary magnetic field parameters, Radio Sci., 8, 933, 1973.

56 Heyvaerts, J., and E. R. Priest, Thermal evolution of current sheets and flash phase of solar flares, Solar Phys., 47, 223, 1976.

57 Heyvaerts, J., E. R. Priest, and D. Rust, An emerging flux model for the solar flare phenomenon, Astrophys. J., 216, 123,1977.

58 Hill, T. W., A three-dimensional model of magnetic merging at the magnteopause, Radio Sci., 8, 915, 1973.

59 Hill, T. W., Magnetic merging in a collisionless plasma, J. Geophys. Res., 80, 4689, 1975.

60 Hofmann, I., Resistive tearing modes in a sheet pinch with shear flow, Plasma Phys., 17, 143, 1975.

61 Hoh, F. C., Stability of sheet pinch, Phys. Fluids, 9, 277, 1966.

62 Holzworth, R. H., and C.-I. Meng, Mathematical representation of the auroral oval, Geophys. Res. Lett., 2, 377, 1975.

63 Hones, E. W., Jr., S. J. Bame, and J. R. Asbridge, Proton flow measurements in the magnetotail plasma sheet made with Imp 6, J. Geophys. Res., 81, 227, 1976.

64 Huba, J. D., N. T. Gladd, and K. Papadopoulos, The lower-hybrid-drift

instability as a source of anamalous resistivity for magnetic field line reconnection, Geophys. Res. Lett., 4, 125, 1977.

65 Janicke, L., Adiabatic Invarianz im Mittel bei räumlicher Feldumkehr, thesis, Institut für Plasmaphysic der Kernforchungsanlage, Jülich, 1965.

66 Kan, J. R., and J. K. Chao, Magnetic field reconnection enhanced by interchange stability, to appear J. Geophys. Res., 82, 1977.

67 Kennel, C. F., and F. V. Coroniti, Possible origin of time variability in Jupiter's outer magnetosphere, Geophys. Res. Lett., 4, 215, 1977.

68 Kirsch, E., S. M. Krimigis, E. T. Sarris, R. P. Lepping and T. P. Armstrong Possible evidence for large transient electric fields in the magnetotail from oppositely directed anisotropies of energetic protons and electrons, Geophys. Res. Lett., 4, 137, 1977.

69 Laval, G., R. Pellat, and M. Vuillemin, in Plasma Physics and Controlled Fusion Research-Culham, Vol. II, 259, 1965.

70 Layzer, D., and J. R. Burke, On the acceleration of ultrarelativistic electrons and nuclei in nonthermal radio sources, Ap. J., 157, 1169, 1969.

71 Lees, L., Interaction between the solar plasma wind and the geomagnetic cavity, AIAA J., 2, 1576, 1964.

72 Levy, R. H., H. E. Petschek, and G. L. Siscoe, Aerodynamic aspects of the magnetospheric flow, AIAA J., 2, 2065, 1964.

73 Lin, R. P., K. A. Anderson, J. E. McCoy, and C. T. Russell, Observations of magnetic merging and the formation of the plasma sheet in the earth's magnetotail, J. Geophys. Res., 82, 2761, 1977.

74 Lui, A. T., C.-I. Meng, and S.-I. Akasofu, Search for the magnetic neutral line in the near-earth plasma sheet, I., critical reexamination of earlier

studies on magnetic field observations, J. Geophys. Res., 81, 5934, 1976.

75 Lui, A. T., C.-I. Meng, and S.-I. Akasofu, Search for the magnetic neutral line in the near earth plasma sheet, 2., Systematic study of IMP 6 magnetic field observations, J. Geophys. Res., 82, 1547, 1977.

76 Lynden-Bell, D., Galactic nuclei as collapsed old quasars, Nature, 223, 690, 1969.

77 McDiarmid, I. B., J. R. Burrows, and E. E. Budzinski, Average characteristics of magnetospheric electrons (150eV-200keV) at 1400 km, J. Geophys. Res., 80, 73, 1975.

78 McDiarmid, I. B., J. R. Burrows, and E. E. Budzinski, Particle properties in the day side cleft, J. Geophys. Res., 81, 221, 1976.

79 Maezawa, K., Dependence of the magnetopause position on the southward interplanetary magnetic field, Planet. Space Sci., 22, 1443, 1974.

80 Maezawa, K., Magnetotail boundary motion associated with geomagnetic substorms, J. Geophys. Res., 80, 3543, 1975.

81 Meng, C.-I., and K. A. Anderson, A layer of energetic electrons *(E > 40keV)* near the magnetopause, J. Geophys. Res., 75, 1827, 1970.

82 Mitchell, H. G., Jr., and J. R. Kan, Merging of magnetic fields with field-aligned plasma flow components, submitted to J. Geophys. Res., 1977.

83 Newcomb, W. A., Motion of magnetic lines of force, Ann. Phys., 3, 347, 1958.

84 Ohyabu, N., S. Okamura, and N. Kawashima, Strong ion heating in a magnetic neutral point discharge, Phys. Fluids, 17, 2009, 1974.

85 Overskei, D., and P. A. Politzer, Plasma turbulence in the vicinity of a magnetic neutral point, Phys. Fluids, 19, 683, 1976.

86 Parker, E. N., The solar flare phenomenon and the theory of reconnection and annihilation of magnetic fields, Astrophys. J., Suppl. Series, 8, 177, 1963.

87 Parker, E. N., The origin of magnetic fields, Ap. J., 160, 383, 1970.

88 Parker, E. N., Comments on the reconnection rate of magnetic fields, J. Plasma Phys., 49, 1973a.

89 Parker, E. N., The reconnection rate of magnetic fields, Astrophys. J., 180, 247, 1973b.

90 Petschek, H. E., Magnetic field annihilation, AAS-NASA Symposium on the Physics of Solar Flares, W. N. Hess, ed., NASA SP-50, p. 425, 1964.

91 Petschek, H. E., The mechanism for reconnection of geomagnetic and interplanetary field lines, in The Solar Wind, R. J. Mackin, Jr., and M. Neugebauer, eds., p. 257, Pergamon Press, New York, 1966.

92 Priest, E. R., The solar flare phenomenon, Physics of Solar-Planetary Environments, D. J. Williams, ed., p. 144, Am. Geophys. U., 1976.

93 Priest, E. R., and S. W. H. Cowley, Some comments on magnetic-field reconnection, J. Plasma Phys., 14, 271, 1975.

94 Priest, E. R., and M. A. Raduu, Preflare current sheets in the solar atmosphere, Solar Phys., 43, 177, 1975.

95 Priest, E. R., and B. U. Ö. Sonnerup, Theories of magnetic field annihilation, Geophys. J., R. Astr. Soc., 41, 405, 1975.

96 Rossi, B., and S. Olbert, Introduction to the physics of space, McGraw Hill, New York, 1970.

97 Rutherford, P. H., Nonlinear growth of the tearing mode, Phys. Fluids 16, 1903, 1973.

98 Sarris, E. T., and W. I. Axford, Energetic protons at the plasma sheet boundary, Geophys. Res. Lett., 4, , 1977. (to appear).

99 Schindler, K., A theory of the substorm mechanism, J. Geophys. Res., 79, 2803, 1974.

100 Schindler, K., Large scale plasma processes in the solar system, Physics of Solar Planetary Environments, D. J. Williams, ed., p. 798, Am Geophys. U., 1976.

101 Smith, D. F., and E. R. Priest, Current limitation in solar flares, Astrophys. J., 176, 487, 1972.

102 Smith, D. F., Current instability in reconnecting current sheets, J. Geophys. Res., 82, 704, 1977.

103 Sonnerup, B. U. Ö., Magnetic-field re-connexion in a highly conducting incompressible fluid, J. Plasma Phys., 4, 161-174, 1970.

104 Sonnerup, B. U. Ö., Adiabatic particle orbits in a magnetic null sheet, J. Geophys. Res., 76, 8211, 1971.

105 Sonnerup, B. U. Ö., Magnetic field reconnection and particle acceleration, High Energy Phenomena on the Sun, Symposium Proceedings, pp. 357-376, NASA-GSFC Doc. X-639-73-193, 1973.

106 Sonnerup, B. U. Ö., The reconnecting magnetosphere, in Magnetospheric Physics, B. M. McCormac, ed., p. 23, D. Reidel Publ. Co., Dordrecht-Holland, 1974a.

107 Sonnerup, B. U. Ö., Magnetopause reconnection rate, J. Geophys. Res., 79, 1546, 1974b.

108 Sonnerup, B. U. Ö., Magnetopause and boundary layer, Physics of Solar

Planetary Environments, D. J. Williams, ed., p. 541, Am. Geophys. U., 1976.

109 Sonnerup, B. U. Ö., and E. R. Priest, Resistive MHD stagnation point flows at a current sheet, J. Plasma Phys., 14, 283, 1975.

110 Soward, A. M., and E. R. Priest, Fast magnetic-field reconnection, Phil. Trans., Roy. Soc. London, A, 284, 369, 1977.

111 Speiser, T. W., Conductivity without collisions or noise, Planet. Space Sci., 18, 613, 1970.

112 Spitzer, L., Jr., Physics of fully ionized gases, Interscience Publ., New York, 1964.

113 Sweet, P. A., The neutral point theory of solar flares, in Electromagnetic Phenomena in Cosmical Physics, B. Lehnert ed., Cambridge U. Press, 123, 1958.

114 Syrovatskii, S. I., A. G. Frank, and A. Z. Khodzhaev, Development of a current layer when plasma moves in a magnetic field with a neutral line, Soviet Physics, JETP Lett., 15, 94, 1972.

115 Tamao, T., Anamalous electrical conductivity for the field-line merging current, Report of Ionosphere and Space Res., in Japan, 29, 140, 1975.

116 Tsurutani, B. T., B. E. Goldstein, and A. Bratenahl, Acceleration of energetic particles of the outer regions and planetary magnetospheres: inferences from laboratory and space experiments, Planet. Space Sci., 24, 995, 1976.

117 VanHoven, G., and M. A. Cross, Energy release by magnetic tearing: The nonlinear limit, Phys. Rev., A7, 1347, 1973.

118 Vasyliunas, V. M., Theoretical models of magnetic field line merging, Revs. Geophys. Space Phys., 13, 303, 1975.

119 Vasyliunas, V. M., Private communication, June, 1977.

120 Waddell, B. V., M. N. Rosenbluth, D. A. Monticello, and R. B. White, Nonlinear growth of the m=1 tearing mode, Nuclear Fusion, , , 1977.

121 White, R. B., D. A. Monticello, M. N. Rosenbluth, and B. V. Waddell, Nonlinear tearing modes in tokamaks, Paper presented at Sixth International Conference on Plasma Physics and Controlled Nuclear Fusion, Berchtesgaden, W. Germany, Oct., 1976, Princeton Plasma Physics Laboratory Report PPPL-1282, 1976.

122 White, R. B., D. A. Monticello, M. N. Rosenbluth, and B. V. Waddell, Saturation of the tearing mode, Phys. Fluids, 20, 800, 1977.

123 Wilcox, J. M., and N. F. Ness, Quasi-stationary corotating structure in the interplanetary medium, J. Geophys. Res., 70, 5793, 1965.

124 Yang, C.-K., Compressible magnetic field reconnection, thesis, Dartmouth College, Hanover, NH., June, 1976.

125 Yang, C.-K., and B. U. Ö. Sonnerup, Compressible magnetic field reconnection a slow wave model, Astrophys. J., 206, 570, 1976.

126 Yang, C.-K., and B. U. Ö. Sonnerup, Compressible magnetopause reconnection, J. Geophys. Res., 82, 4, 1977.

127 Yeh, T., Day-side reconnection between a dipolar geomagnetic field and a uniform interplanetary field, J. Geophys. Res., 81, 2140, 1976.

128 Yeh, T., Diffusive hydromagnetic flow in the vicinity of a neutral point, Astrophys. J., 207, 837, 1976a.

129 Yeh, T., Reconnection of magnetic field lines in viscous conducting fluids, J. Geophys. Res., 81, 4524, 1976b.

130 Yeh, T., and M. Dryer, On the reconnection of magnetic field lines in compressible conducting fluids, Astrophys. J., 182, 301-315, 1973.

131 Yeh, T., and W. I. Axford, On the re-connexion of magnetic field lines in conducting fluids, J. Plasma Phys., 4, 207-229, 1970.

132 Zwan, B. J., and R. A. Wolf, Depletion of solar wind plasma near a planetary boundary, J. Geophys. Res., 81, 1636, 1976.

PLASMA PROCESSES IN THE EARTH'S RADIATION BELT

by

L. R. Lyons

I. Introduction

The earth's radiation belts consist of energetic electrons and ions (energies from <1 Kev to >100 MeV) trapped in the geomagnetic field on magnetic field lines with L-values from ∿1.1 to ∿10. (For the purposes here, L can be defined as the equatorial crossing of a magnetic field line measured in earth radii from the center of the earth.) Here we consider the basic plasma processes affecting trapped particles from ∿1 keV to a few MeV. An excellent and more in depth review of theoretical concepts is given by Schulz[37], and a more complete review of research over the past few years is given by West[48]. The very high energy (>30 MeV) trapped protons below L=2, not considered here, have been reviewed by White[50]. Research of the past five years is emphasized in this report. Readers are referred to the comprehensive treatment of the radiation belts by Schulz and Lanzerotti[38] for earlier research .

Trapped particles execute quasi-periodic motion which can be divided into three components, each component associated with an adiabatic invariant. The three components of trapped particle motion are schematically illustrated in Figure 1, along with magnetospheric regions mentioned in this section. The particle gyration about field lines is associated with the first invariant $M=p_{\perp}^2/2mB$, where m is rest mass, p is momentum, and subscripts "$\perp$" and "$\parallel$" refer to the components of vector quantities normal and parallel to the geomagnetic field $\underline{B}$. The gyrofrequency (often called the cyclotron frequency) is $\sim 880/L^3$ kHz for non-relativistic electrons and a factor of ∿2000 less for protons. The bounce motion between mirror points is associated with the second invariant $J=\oint p_{\parallel} ds$, where the line integral is evaluated between mirror points along the field line about

which the particle gyrates. Typical bounce frequencies are ∿0.3-10 Hz for electrons and ∿0.03-3 Hz for protons. The particle drift around the earth is associated with the third invariant Φ which is equal to the magnetic flux enclosed by the drift orbit. Drift frequencies range from ∿0.0001-100 mHz.

Particle collisions and interactions with plasma waves can cause violation of these adiabatic invariants. Generally each interaction causes a small perturbation on a particle's trajectory, so that the net result of a large number of such interactions is diffusion in velocity or in position. Violation of the first two invariants results in particle diffusion in pitch angle and energy. (The pitch angle α is given by $\tan \alpha = v_\perp / v_{||}$, where $\underline{v}$ is velocity. Thus particles with equatorial pitch angles of 90^0 are confined to the geomagnetic equator, while particles with equatorial pitch angles approaching 0^0 and 180^0 mirror at increasingly higher latitudes. Particles with equatorial pitch angles sufficiently close to 0^0 and 180^0 strike the atmosphere and are lost from the radiation belts via collisions. Such particles are said to be in the loss cone. The loss cone is $<1^0$ wide in equatorial pitch angle in the outer regions of the radiation belts, and increases to a few tens of degrees wide at L<2.) Pitch angle diffusion into the loss cone is often an important loss process for radiation belt particles, while diffusion in energy can be an important particle energization process.

Violation of the third invariant causes particles to diffuse to magnetic field lines closer or further from the earth. Such radial diffusion is an important source for much of the inner regions of the

radiation belts. Processes which cause radial diffusion generally conserve the first and second adiabatic invariants, so that particles which diffuse towards the earth are significantly energized due to the increase in B (conservation of M implies $p_\perp^2 \sim B$). In addition to diffusive particle sources and losses, direct energetic particle injection and loss results from processes such as energetic neutron decay and charge exchange with cold neutrals, and particles can be injected directly into the radiation belts from the geomagnetic tail and cusps by electric fields normal to $\underline{B}$ and from the ionosphere by electric fields parallel to $\underline{B}$.

The above plasma processes are all known to affect radiation belt particles. Several of them have been quantitatively studied for specific particle populations, and the results have been directly compared with satellite particle observations. Such studies have generally considered only regions within the location of the plasmapause during periods of relatively low geomagnetic activity.

The plasmapause is the outer boundary of the plasmasphere, a region of high cold plasma density (10^2-10^4 cm^{-3}) extending from the equatorial and mid-latitude ionosphere and terminating relatively abruptly at field lines near L=4 to 6. Cold plasma densities outside the plasmapause are $\lesssim 1$ cm^{-3}. Outside the plasmapause, measured radiation belt particle fluxes vary considerably on time scales on the order of hours to days in response to variations in the level of geomagnetic activity, even during relatively quiet periods. On the other hand, only during large geomagnetic storms (which occur several times a year) does direct injection of particles affect radiation belt fluxes below L=4. Following such storms,

particle fluxes within the plasmasphere gradually return to their pre-storm levels. An example of observed variations of radiation belt particles, in this case electrons at L=2, 3, 4 and 5, is shown in Figure 2 for a one month period which included a relatively quiet period (Dec. 9-16), a large storm (Dec. 17), and a storm recovery (Dec. 17-Jan. 10). Notice the variability of the fluxes at L=5, the stability at the fluxes at L=2 and 3, and the storm flux enhancement at L=3 followed by a recovery to pre-storm levels.

Within the plasmasphere, well-defined plasma processes proceed essentially continually for periods of weeks to months uninterupted by direct injections. This region is thus ideal for comparing radiation belt particle measurements with theoretical discriptions of plasma processes, and fortunately good measurement of the plasma energetic particle distribution function are available from the Explorer 45 satellite. Beyond L=5, good measurements of the particle distribution function are not available. This, together with the variability of the fluxes, has resulted in our understanding of the processes controlling radiation belt particles being for more limited outside the plasmapause than within.

II. General Comments on Particle Diffusion from Resonant Wave-Particle Interaction

Considerable attention has been given to understanding the interaction of radiation belt particles with plasma waves. Such analyses have generally considered waves with frequencies on the order of the particle gyrofrequency, which can cause significant particle pitch angle diffusion, and waves with frequencies on the order of the particle drift frequency, which can cause significant radial diffusion. Resonance with waves on the order of the bounce frequency[35, 42] has received considerably less attention.

The conditions for resonance between waves of frequency ω and particles of gyrofrequency Ω is given by:

$$\omega - k_{||} v_{||} = n\Omega \; ; \; n = o \; \pm 1, \; \pm 2, \; \pm 3$$

where $\underline{k}$ and $\underline{v}$ are the wave vector and the particle velocity. The dispersion relation for the wave mode under consideration relates $\underline{k}$ to ω. The cyclotron harmonic resonances have the doppler shifted wave frequency equaling a harmonic of the particle gyrofrequency and are given by $n=\pm 1, \pm 2, \ldots$, and the $n=o$ Landau resonance has the wave parallel phase velocity $\omega/k_{||}=v_{||}$. All resonances result in particle diffusion in both pitch angle and energy. The cyclotron resonances can cause significant diffusion into the loss cone, while the Landau resonance results in diffusion solely in $v_{||}$ so that diffusion is primarily in energy at pitch angles near the loss cone.

For simplicity, assume $\omega << \Omega$. Then the resonant parallel velocity for each cyclotron harmonic resonance $v_{||,n}=n\Omega/k_{||}$aand for the Landau resonance $v_{||,o}=\omega/k_{||} << |v_{||,n}|$. Assuming waves propagate both up and down field lines, resonance at each harmonic occurs for both $v_{||,n}$ and $-v_{||,n}$. Figure 3 illustrates the regions of cyclotron resonance in velocity space for waves distributed over some band of $k_{||}$'s. Relativistic effects are not included. Resonance at each cyclotron harmonic occurs over a band of $v_{||}$'s and no cyclotron interaction occurs for $v_{||}$ less than a minimum $v_{||,min}$. That no cyclotron resonance occurs for $v_{||}$ less than a $v_{||,min}$ does not depend on the assumption $\omega << \Omega$; however the value of $v_{||,min}$ depends upon the band of $k_{||}$'s over which wave energy is distributed.

The simple picture in Figure 3 is modified somewhat in the radiation belts because the geomagnetic field is not homogeneous, and wave energy is generally distributed over some range of geomagnetic latitudes.

As a particle moves away from the equator along its bounce trajectory, the increasing geomagnetic field strength causes both the particle pitch angle and the parallel velocities $v_{||,n}$ for resonance with the waves (for most waves modes) to increase. The effects of off-equatorial interactions can be represented by Figure 3 by removing the upper $v_{||}$ bounds of the resonant regions and re-labeling the axes as equatorial velocity. Thus cyclotron resonance occurs for all equatorial parallel particle energies $E_{||}=1/2mv_{||}^2$ greater than some minimum value $E_{||,min}$, so that particle diffusion into the loss cone occurs for all particle energies $E>E_{||,min}$, but not for $E<E_{||,min}$.

As shall be seen in the example discussed later, equatorial particle measurements can show marked effects from cyclotron resonance at $E_{||}>E_{||,min}$ but not for $E_{||}<E_{||,min}$, and such observations have given strong support to specific wave-particle interactions which have been proposed as the dominant loss mechanism for specific populations of trapped particles. Given a proposed distribution of wave energy in the radiation belts, quasi-linear diffusion theory [20] can be applied to obtain pitch angle and energy diffusion coefficients[23, 24]. These diffusion coefficients can then be used to calculate the distribution of trapped particles as a function of pitch angle[19, 22], and these calculations can be directly compared with particle measurements.

The above approach for study resonant wave-particle interactions requires knowledge of the wave distribution either from measurements directly within the radiation belts or from theoretical predictions. The more general approach of self-consistently calculating the wave distribution and

particle diffusion rates is generally unreasonably difficult for the radiation belts. Specific cases have been attempted[10, 21], but sufficient sophistication to allow definitive comparison with particle observations has not been achieved.

Fluctuations of large-scale magnetospheric electric fields with frequencies equal to a harmonic of the particle drift frequency can cause radial diffusion. Such fluctuating electric fields are generally divided into two classes[12]; 1) Electric fields induced by fluctuations in the geomagnetic field ($\nabla \times \underline{E} \neq 0$), 2) Fluctuations of the magnetospheric potential electric field for which $\nabla \times \underline{E} = 0$. By attempting to make reasonable assumptions about the distribution of fluctuating fields, expressions for radial diffusion coefficients have been derived (e.g. Nakada and Mead [33], for magnetic fluctuations; Cornwall [2], for electric field fluctuations).

Given radial diffusion coefficients and loss rates calculated from pitch angle diffusion coefficients, and including other processes such as those listed in Section I when relevant, radiation belt fluxes as a function of radial distance can be calculated. Such calculations can be compared directly, and as shall be seen, some of these comparisons have been remarkably successful.

III. Plasma Processes and Comparisons with Observations

1. Quiet-time electrons

Valid measurements of radiation belt electrons within the plasmasphere are available for electron energies from ∿30 keV to ∿2 MeV. The interaction of these electrons with naturally occurring whistler-mode waves is a well understood example of wave-particle interactions in the radiation belts. During geomagnetically quiet times, the electrons are distributed in two zones, inner zone fluxes peak near L=1.2-2, outer zone fluxes peak near L=4-6, and the region of low

fluxes in between is called the electron slot. Electron interactions with whistler-mode waves have been studied in detail to explain the electron losses required to form the slot [28].

Satellite measurements have shown that a band of whistler-mode waves (called hiss) centered near a few hundred Hz exists nearly continually and are probably the dominant wave mode throughout the plasmasphere [36, 45]. These waves, which can readily propagate across magnetic field lines [26], are believed to be generated by radiation belt electrons in the outer region of the plasmasphere [45, 46]. The generated wave energy then propagates across field lines to all regions of the plasmasphere.

Pitch angle diffusion coefficients were calculated [28] for electrons resonant with the observed distribution of plasmaspheric hiss. Using the calculated diffusion coefficients, prediction of electron pitch angle distributions and loss rates from diffusion into the loss cone were obtained.

Examples of the calculated pitch-angle diffusion coefficients and equatorial pitch-angle distributions at L=4 are shown in Figure 4 for 20 keV, 200 keV, and 2 MeV electrons. Cyclotron resonance occurs for equatorial pitch angles up to $\sim 55^0$ at 20 keV. Cyclotron resonance extends to increasing pitch angles with increasing electron energy, since cyclotron resonance occurs for all $E_{||} \geq E_{||,min}$ independent of the electron energy. Landau resonance occurs only at pitch angles near 90^0, but off-equatorial interactions cause the range of equatorial pitch angles affected by Landau resonance to overlap the pitch-

angle range of cyclotron resonance. Note that the pitch-angle diffusion rates are relatively low at values of $E_{||}$ just below $E_{||,min}$. This slow diffusion manifests itself in the equatorial pitch-angle distributions as a region of large slope. Thus the equatorial distributions develop bumps surrounding 90^0 pitch angles, and this bump decreases in size and pitch angle extent with increasing electron energy. Since $E_{||,min} \sim B^2/N \sim L^{-2}$ (for an equatorial electron density $N \sim L^{-4}$), these bumps are also expected to decrease with increasing L-value for a constant electron energy.

Such pitch-angle distributions, with the expected variation with energy and L, were first observed on OGO-5 [49], and are now known to be a general feature of quiet-time radiation belt electrons [29]. Examples of West et al.'s observations are shown in Figure 5 (from Lyons et al. [28]. The theoretically predicted pitch angle distributions are shown as solid lines at the appropriate L-values and energies. The calculations overestimate the magnitude of the 90^0 pitch angle bumps, but this could be corrected by reasonable alterations of the distribution of wave energy. However, no reasonable alteration of the wave distribution could change how the bumps vary with L and energy.

Since $E_{||,min}$ increases with decreasing L, cyclotron resonance with the band-limited hiss does not occur within the inner zone for the energies considered here. Coulomb collisions with the cold plasmaspheric plasma thus become the dominant pitch angle diffusion mechanism within the inner zone, becoming dominant at $L \lesssim 3$ for 50 keV electrons and $L \lesssim 2$ for 350 keV electrons. Equatorial pitch angle distributions obtained by Explorer 45 clearly show the transition from pitch angle

diffusion dominated by wave-particle interactions at higher L to domination by Coulomb collisions at lower L [29].

Using the calculations of diffusion rates for interactions with the hiss, and including the effects of Coulomb collisions in the inner zone, we have realistic predictions of particle loss rates as a function of electron energy and L throughout the plasmashphere. Lyons and Thorne [27] balanced the calculated electron loss rates with radial diffusion to obtain an equilibrium structure of radiation belt electrons throughout the plasmasphere. Radial diffusion was assumed to be driven by fluctuations of the magnetospheric potential electric field as modeled by Cornwall [2].

The resulting particle distribution function f for equatorially mirroring electrons as a function of L at constant first adiabatic invariant M is shown in the left-hand panel of Figure 6. Curves for the various values of M were all normalized to the same value at the plasmapause, here taken to be L=5.5. Since f must vanish at L=1, all the curves monotonically decrease with decreasing radial distance, and there is no sign of a two-zone structure. However, the curves at constant M were then used to obtain the differential unidirectional particle flux $j=fp^2$ (p is momentum) as a function of L at fixed electron energy. A flux versus energy spectrum at L=5.5 based on observations was used as the required boundary condition. The results are shown in the right panel of Figure 6, and a two-zone structure appears as a consequence of radial diffusion from the plasmapause to a sink at L=1 and of the increase in energy as an electron diffuses inwards (energy $\sim 1/L^3$ for non-relativistic,

equatorially-mirroring electrons). The calculated slot moves inward and becomes increasingly pronounced with increasing energy and the outer zone peak mover inward within the plasmasphere above 1 MeV in agreement with observation. A comparison between the calculated equilibrium electron structure and quiet-time observations of Pfitzer et al.[34] is given in Figure 7.

During large geomagnetic storms, electrons are injected within the quiet time location of the plasmapause, and the equilibrium pitch angle distributions and radial structure are destroyed. However, Explorer 45 observations have shown that the pitch angle distribution and radial profiles gradually return to their pre-storm structure over a period of a few weeks [30]. It thus appears that the dominant, long-term-averaged, quiet-time source and loss processes for radiation belt electrons within the plasmasphere have been correctly identified and quantitatively evaluated. Variations of waves distributions and resulting diffusion rates over time scales of a day or less are not as well understood.

2. Quiet-time ions

Ions undergo cyclotron resonance with ion-cyclotron waves in much the same manner as do electrons with whistler-mode waves. While whistler waves have frequencies below the electron gyrofrequency and generally above the proton gyrofrequency, ion-cyclotron wave frequencies are below the proton gyrofrequency. However, our knowledge of the distribution of ion cyclotron waves within the plasmasphere is far less than our knowledge of the distribution of whistler-mode wave. This is because satellite instrumentation has been less sensitive at ion-cyclotron wave frequencies (generally $\lesssim$10 Hz) than at whistler-mode frequencies (generally $\gtrsim$100 Hz).

Low altitude (several hundred km) satellite observations within the plasmasphere have found ions precipitating into the ionosphere during quiet times and during storms, implying that trapped ions are pitch angle diffused into the loss cone [13, 16]. However, the importance of such diffusion in determining radiation belt structure has been a subject of continuing research for the past several years.

A significant advance in our understanding of the quiet time distribution of trapped protons has recently been achieved by Spjeldvik [43]. He followed the approach of Cornwall[3] and balanced radial diffusion as modelled by Cornwall [2] with proton losses from charge exchange with the ambient neutral hydrogen geocorona and Coulomb collisions. The outer boundary condition was taken from the observed ion fluxes from ATS-6 at L=6.6. His results are displayed in Figure 8 as energy spectra at various L-values for equatorially mirroring ions. The circles and triangles at $L \lesssim 5.25$ are equatorial ion observations from Explorer 45. Unfortunately, the lack of valid data prohibits comparison below 25 keV at L=4 to 5 and below 150 keV at L<4. However, the remarkable agreement between the calculations and the observations implies that the dominant processes affecting the protons have been identified over the energy range of valid observations, and that wave-particle interactions are probably not important for equatorially mirroring protons at these energies. Spjeldvik considered only equatorially mirroring protons, and whether or not his results holds for non-equatorially mirroring protons is yet to be determined.

Since the ion measurements can not distinguish between protons and other ions, the agreements in Figure 8 imply that the measured

ion fluxes were indeed dominated by protons, since the calculated radial diffusion and loss rates were obtained by assuming the particles were protons and the rates are significantly different for other ion species. However, the calculated proton spectra at $L<4$ fall off markedly below 100 keV due to the fast charge exchange loss rates below 100 keV. It is unfortunate that valid measurements are not available below 100 keV, since Spjeldvik's calculations imply that any significant ion fluxes below 100 keV may not be dominated by protons. This range of energies at $L \lesssim 4$ is discussed in more detail in the following section on the decay of the stormtime ring current.

Spjeldvik and Fritz [44] performed a similar calculation for He^{+} and He^{++}. They compared their results to the very limited Explorer 45 He^{++} observations at 1-3 MeV. The comparison was fairly good, but sufficient observations for a definitive comparison are simply not yet available.

Note that Lyons and Thorne[27], Spjeldvik[43], and Spjeldvik and Fritz[44] all used radial diffusion rates as formulated by Cornwall[2] for fluctuations of the magnetospheric potential electric field. In addition, the magnitude of Cornwall's diffusion rates appear to be valid for the > 30 MeV protons below L=2 [1, 7]. It thus appears that Cornwall's formulation, which was based on very limited electric field observations, is remarkably accurate over a wide range of particle energies, L-values, and particle species. Without such a valid estimate of radial diffusion rates, attempts to explain the quiet-time structures of radiation belt particles within the plasmasphere would have been far less successful.

3. Decay of stormtime ring current ions

During storms, electrons and ions are injected to L-values as low as 2 so that trapped particle fluxes become significantly enhanced over their quiet-time equilibrium values within the location of the quiet-time plasmapause. The drift of these particles around the earth forms a current, called the ring current, which causes a significant (50-400γ) depression of the magnetic field at the earth's surface. While the decay of the electrons is understood in terms of interactions with whistler-mode waves, the decay of the stormtime ring current ions is presently an active area of research. Even the composition of these ions, until recently believed to be mostly protons, is now being seriously questioned.

Enhanced ion fluxes are observed from $\sim$1 keV to $\sim$200 keV during storms, and fortunately these fluxes are above the Explorer 45 background levels so the valid measurements of equatorial pitch angle distributions and energy spectra are available over most of this energy range from L=2.5 to 5.

Figure 9 shows Explorer 45 observations of the distribution function f of equatorially mirroring ions throughout the period of the large storm on Dec. 17, 1971. The quantity $2mf=j/E$ is shown versus time for L=2.5, 3.0, ...5.0 for 8 representative ion energies where j is the measured ion flux. Dst, a measure of the magnetic depression on the ground, is also shown versus time. Every valid data point is shown from both inbound and outbound portions of Explorer 45 orbits that are within $\sim 8^0$ of the geomagnetic equator. The prestorm intensities at energies $\lesssim$26 keV are upper limits to the true intensities because of an undetermined background count rate; however,

these data are included to emphasize the stormtime increase in particle intensities.

In association with the December 17 storm main phase (the large decrease in Dst), the data show intensity increases at the lower ion energies and, at $L \gtrsim 4$, intensity decreases at the higher energies. Some response to the sudden commencement on December 16 is also evident. Following the storm, the particle intensities gradually return towards their pre-storm values. By December 21, Dst had nearly recovered, though some of the particle intensities remained near their stormtime values (e.g. 164 keV protons at 3.5 R_e). Small intensity variations at low E and high L occurred on December 22 along with a small decrease in Dst.

The decrease in f at the higher energies simply results from the conservation of M (M=E/B for equatorially mirroring ions). During the storm, B decreases so that particle energies decrease in order to conserve M. The decrease in particle energies, together with the overall decrease in f with increasing E, results in a decrease in f when measured at a fixed value of E. However no actual loss of high energy particles occurs [31]. On the other hand, the observed increase in f at the lower energies represents a net increase in the trapped ion population, and this increase is a significant part of the stormtime ring current.

Two processes are believed to be important in the decay of the stormtime ring current: charge exchange with neutral hydrogen [9] and resonant interactions with ion-cyclotron waves [5]. The importance of

both of these processes has been tested by using Explorer 45 measurements of equatorial pitch angle distributions.

3.1 Interaction with ion-cyclotron waves

Figures 10 and 11 show the equatorial pitch angle distributions, obtained ∿16 hours (orbit 103) and ∿24 hours (orbit 104) after the minimum Dst of the Dec. 17, 1971 storm. Observations from selected ion energy channels are shown every 0.4 in L from L=3.0 to 5.0. Notice the transition from relatively isotropic pitch angle distributions at the lower energies to rounded distributions, peaked at 90^0 pitch angle, at the higher energies. The isotropic distributions show significant flux decreases whenever the pitch angle scan of a measurement reaches the loss cone, implying the loss cones are nearly empty of particles. Such isotropic distributions with empty loss cones indicate a stably trapped particle populations undergoing negligible pitch angle diffusion into the loss cone. As the storm recovery progresses, the transition from flat to rounded pitch angle distributions shifts to lower energies, so that flat distributions become rounded. Note, for example, the pitch angle distributions in Figure 10 and 11 for 14 keV at L=3.8, 26 keV at L=4.2, and 42 keV at L=4.6. Such rounding of the pitch angle distributions represents a loss of non-equatorially mirroring particles, and rounded distributions with fluxes monotonically increasing towards 90^0 pitch angle are expected under conditions of pitch angle diffusion into the loss cone.

For each L-value shown in Figures 10 and 11, the energy of the highest Explorer 45 energy channel showing a nearly isotropic pitch angle distribution was assumed to be $E_{||,min}$ (the minimum parallel ion energy subject to cyclotron resonance with a band of waves), and the pitch angles corresponding to the chosen value of $E_{||}$ was calculated for all higher energy channels. These pitch angles have been indicated by vertical ticks on the distributions. Distributions with ticks at 0^0 and 180^0 give the chosen value of $E_{||,min}$, though these distributions are not shown at all L-values in the figures because of the use of only selected energy channels. Notice that to within the accuracy of the pitch angle measurements (which is 22^0 for $\leq$104 keV and 33^0 for E>104 keV), the pitch angle distributions are nearly isotropic between the ticks, i.e. for $E_{||}<E_{||,min}$. At larger $E_{||}$ (pitch angles approaching 0^0 and 180^0), the distributions become rounded. For all L-values shown, the ticks quite accurately separate regions of nearly isotropic distributions at lower values of $E_{||}$ from regions of rounded distributions at higher values of $E_{||}$.

In order for an $E_{||,min}$ to be evident in the observed pitch angle distributions, pitch angle diffusion driven by resonant wave-particle interactions must have been the dominant loss process responsible for the rounding of the distributions. Such interactions are governed by the parallel particle velocity, while other loss processes are not organized by the parallel

velocity. Thus another loss process, if dominant, would have masked the ability to determine $E_{||,min}$. In addition, the observed values of $E_{||,min}$ satisfy the condition for cyclotron resonance with ion-cyclotron wave when realistic cold plasma densities are assumed for the outer region of the plasmasphere during a storm recovery phase [18, 51, 52]. Landau resonant diffusion of the ions is probably of little importance for these waves, since wave energy at large wave normal angles is needed for such diffusion and such wave energy is probably damped by low energy electrons[6].

Williams and Lyons [51,52] concluded that the ion-cyclotron waves were amplified by the ring current particles as proposed by Cornwall et al. [5], and that the evolution of isotropic distributions to rounded distributions occurred as the cold plasma density increased during the storm recovery phase. (The parallel energies for cyclotron resonance with ion-cyclotron waves $E_{||,r}$ decrease with increasing cold plasma density, and during storms the cold plasma in the outer regions of the plasmasphere is severely depleted. This plasma is replenished from the ionosphere over a several day period following storms.) However, Williams and Lyons' argument does not explain how the isotropic distributions, which apparently were not undergoing diffusion, could become resonant with unstable waves. A sufficient pitch angle anisotropy is required to grow ion-cyclotron waves, and the isotropic distribution, even with the empty loss cones, have too small an anisotropy to grow waves.

Joselyn and Lyons [18] suggested a resolution of this difficulty by by showing that waves growing off the equator can propagate towards the equator and resonate with the isotropic distributions. $E_{\parallel,r}$ increases markedly away from the equator along field lines, so that we can find a frequency which resonates with rounded distributions off the equator (where $E_{\parallel,r} > E_{\parallel,min}$) but which resonates with isotropic distributions near the equator (where $E_{\parallel,r} < E_{\parallel,min}$). Joselyn and Lyons showed that as the plasma densities continually increase during the storm recovery, wave growth will occur at frequencies so that the minimum $E_{\parallel,r}$ at the equator will continually decrease at the equator. Thus their calculations imply that $E_{\parallel,min}$ should continually decrease as is observed.

3.2 Inconsistency with proton charge exchange

Tinsley [47] has noted that the charge exchange lifetimes for equatorially mirroring protons at energies $\lesssim$30 keV should be on the order of hours. The loss rates from charge exchange should increase significantly with increasing mirror latitude for particles of a given energy on a given L-shell. This results from the decrease in mirror altitude with increasing mirror latitude together with the increase in neutral hydrogen density with decreasing altitude.

Lyons and Evans [25] have investigated the question of how the nearly isotropic pitch angle distributions at the lower energies in Figures 10 and 11 can remain isotropic when charge

exchange with neutral hydrogen should cause the loss rates to increase markedly for equatorial pitch angles increasing or decreasing from 90^0. In Figures 12 and 13 we compare the observed pitch angle distributions at L=3.0 and 3.5 with those expected to evolve from proton charge exchange with neutral hydrogen. Observations are shown from the inbound portion of orbits 101, 102, ..., 106, which are $\sim$8 hours apart, with the observations from orbit 101 being less than one hour following the minimum of Dst.

The evolving pitch angle distributions predicted from charge exchange were obtained by neglecting possible sources and assuming an isotropic proton distribution at the time of the orbit 101 observations, which is taken to be t=0. Fluxes j at subsequent times were calculated from $j \sim \exp(-t/\tau_{ce})$, where the charge exchange lifetimes τ_{ce} are given by Tinsley [47]. These lifetimes were obtained from a recent neutral hydrogen density model using parameters for Dec., 1971.

At both L-values, the charge exchange calculations predict that the pitch angle distributions for 2 and 10 keV protons will become greatly anisotropic in <8 hours. This dramatic rounding of the pitch angle distribution is not observed! At L=3.5, the observed distributions remain essentially isotropic, while at L=3.0 some rounding of the distributions occurs. This rounding is apparently the result of resonant interactions with ion-cyclotron waves. $\tau_{ce} \sim 1$ hour at L=3.5 for 20^0 pitch angle. Thus in order for there to be essentially no rounding of the pitch angle distribution in 40 hours, as is observed, a lifetime $\gtrsim$40 hours is required.

In addition to the disagreement between the shapes of the pitch angle distributions, charge exchange predicts that even the 90^0 pitch angle fluxes should decay much more rapidly than is observed. This discrepancy is particularly dramatic at L=3. The comparisons presented here show that the ring current ions do not decay in a manner consistent with proton charge exchange with neutral hydrogen. Charge exchange decay rates are far too rapid, especially for $L \lesssim 3.5$ and for non-equatorially mirroring particles. Some form of pitch angle diffusion cannot account for the discrepancy in the shapes of the pitch angle distributions, since this would imply total loss rates even greater than those from charge exchange alone.

Lyons and Evans [25] concluded that neither a strong, continual proton source between L=3 and 4 during the storm recovery nor a large error in the neutral hydrogen densities were a reasonable explanation for the observed discrepancy between the observations and the charge exchange predictions. They concluded that the most likely explanation was that the ring current at particle energies $\lesssim 50$ keV and $L \lesssim 4$ was dominated by some ion other than protons during the storm recovery phase. Such ions must have much longer lifetimes for charge exchange with hydrogen than do protons (a factor $\gtrsim 40$ appears to be required) in order to resolve the gross disagreement between the predicted and observed pitch angle distributions. A candidate ion, which has sufficiently long charge exchange lieftimes, is He^+. Additional arguments that He^+ may be the dominant ring current ion during recovery phase, and calculations of charge exchange lifetimes using recent hydrogen models, are given by

Tinsley [47]. At energies $\gtrsim$50 keV, charge exchange lifetimes for He^+ become longer than for H^+, so that the above argument does not apply above 50 keV. Lyons and Evans' arguments also imply that quiet time ring current ions below 50 keV at $L \lesssim 4$ should not be dominated by protons.

Sharp et al.[39, 40] have presented mass spectrometer observations from a low altitude satellite showing significant O^+ and H^+ precipitating from the outer region of the ring current ($L \lesssim 4$) during the recovery of the Dec. 17, 1971 storm; however, precipitating fluxes below L=4 were generally undetectable. In addition, Smith and Bewtra [41] have found that charge exchange can account for some features of the decay of equatorially-mirroring, ring current ions if a multi-component (H^+, He^+, and O^+) plasma is assumed. However, the relative composition of the ring ions, and the relative roles of wave-particle interactions and charge exchange as loss mechanisms, is far from being resolved as a function of particle energy, L-value and time throughout a storm recovery.

IV. Summary and Future Directions

While the dominant particle source and loss processes have yet to be understood throughout most of the magnetosphere, significant quantitative understanding of the quiet-time structure of radiation belt electrons ($\gtrsim$30 keV) and equatorially-mirroring protons ($\gtrsim$100 keV) within the plasmasphere has been obtained over the past several years. The dominant plasma processes affecting these particles apparently have been correctly identified and quantitatively evaluated, though exact quantitative agreements have yet to be obtained. In fact, for

the electrons we can now explain both the quiet-time structure and the post-storm decay be invoking the same processes, though the cause of the stormtime injections is still unknown. Such understanding has only been possible because of the availability of valid measurements of the equatorial pitch angle distributions as a function of radial distance and particle energy. Most of these observations have come from the equatorially orbiting Explorer 45 satellite, with some useful observations also obtained from OGO-5. Electrons with energies <30 keV have not yet been adequately studied, and the suggestion that harmonic waves from ground power distribution networks may affect these electrons[15] needs to be evaluated.

The Explorer 45 equatorial pitch angle distributions of ions, in conjunction with some low altitude observations of precipitating ions, have also given us significant information on the post-storm decay of ring current ions. Effects of wave-particle interactions have been identified and strong evidence has been obtained that protons are not the dominant ion species throughout the ring current. However, presently available observations will probably not allow the ring current decay problem to be completely solved. The distribution of wave energy is unknown from either observations or theory, since instrumentation has not been sufficiently sensitive to measure most of the wave energy below the proton gyrogrequency and we do not know the relative ion composition of the ring current. Calculations of wave growth rates assuming a single ion species have yielded useful information; however the relative densities of different ion species are needed since one species may be unstable to wave growth over a range of frequencies while another

species may damp the same wave frequencies [4]. Even the effects of charge exchange, in principle a simple process to analyze, is currently a subject of debate.

An extremely serious lack in our understanding of the radiation belts exists outside the plasmasphere at $L \gtrsim 5$. The most intense ion and electron precipitation into the ionosphere occurs outside the quiet time plasmasphere at energies from $\sim$1 to $\sim$100 keV (e.g. Hultqvist [17]). Some of this precipitation is associated with bright auroral features and may result from acceleration by low altitude electric fields parallel to $\underline{B}$ [11]. However, the vast majority of the observed particle precipitation outside the plasmasphere appears to have been pitch angle scattered out of a trapped particle population [17]. However, no equatorial measurements of the pitch angle distributions as a function of particle energy and radial distance exist for these particles.

Energy spectra at one pitch angle and pitch angle distributions at one energy have been obtained at one L-value, 6.6, from the ATS satellites at synchronous orbit. These observations have shown that the radiation belt fluxes vary in association with geomagnetic activity [8] and that electric fields along $\underline{B}$ may at times accelerate particles out of the ionosphere into the radiation belts [32]. However, identification and quantitative evaluation of the dominant source and loss processes outside the plasmapause will be impossible until equatorial pitch angle distributions as a function of particle energy and radial distance become available. Hopefully, some of the concepts that have been successfully applied to the inner region of the radiation belts, particularly

those concepts associated with wave-particle interactions, will also be useful in understanding the region beyond L=5. However, available wave observations from outside L=5 indicate that more attention will have to be given to non-linear interactions. Detailed non-linear phenomena, such as the VLF wave-electron interactions observed by Helliwell and Katsufrakes[14] may be important.

The processes discussed here are probably not unique to the Earth's radiation belts, and attempts have been made to apply them to the Jovian radiation belts. We now have made crude measurements of Jovian radiation belt fluxes. However, only measurements of the total particle flux above several fixed energies are available (which necessitates numerical differentiation of data to obtain a crude energy spectra). Until pitch angle distributions and good energy spectra become available from Jupiter, we will be unable to definitively identify which processes govern the distribution of Jovian trapped particles and to quantitatively analyze these processes.

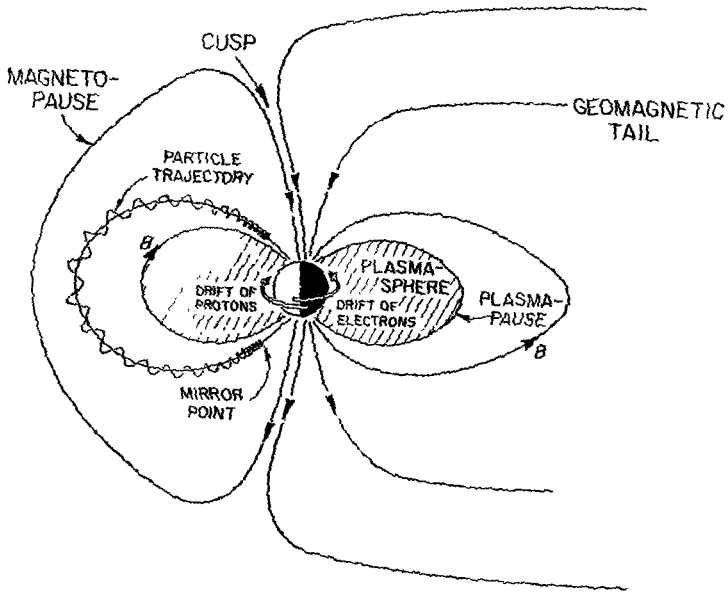


FIGURE 1 Schematic illustration of the three components of trapped particle motion and relevant magnetospheric regions.

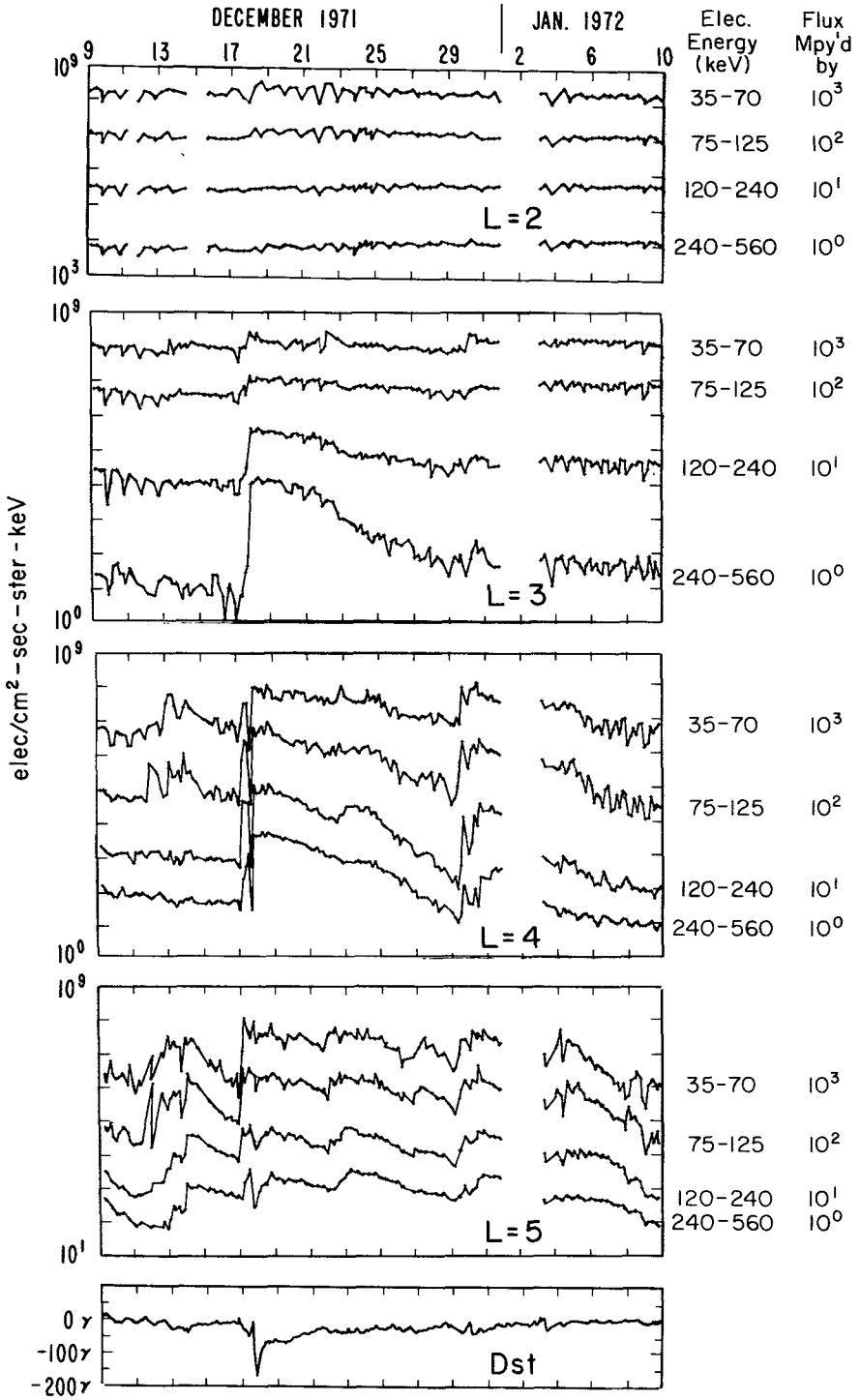


FIGURE 2 Fluxes of equatorially mirroring electrons versus universal time for the period Dec. 9, 1971-Jan. 9, 1972. All available data points from both inbound and outbound portions of the Explorer 45 are shown. Each panel shows the observations at the indicated L-value for the four energy channels. The 120-240 keV, 75-125 keV, and 35-70 keV fluxes have been multiplied by 10^1, 10^2, 10^3, respectively. Dst is also shown (from Lyons and Williams, 1975b).

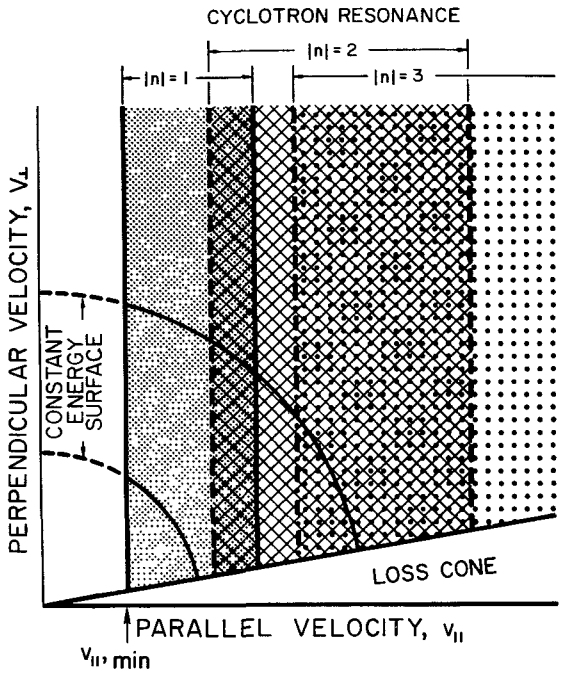


FIGURE 3 Regions in the $(v_\perp, v_\parallel)$-plane of resonance at the $|n| = 1, 2$, and 3 cyclotron harmonics are indicated by different shadings. Wave energy is assumed to be distributed over a bank of $k_\parallel$'s with $\omega \ll \Omega$ for all frequencies in the wave distribution. The minimum parallel velocity for cyclotron resonance $v_{\parallel,min}$ is indicated.

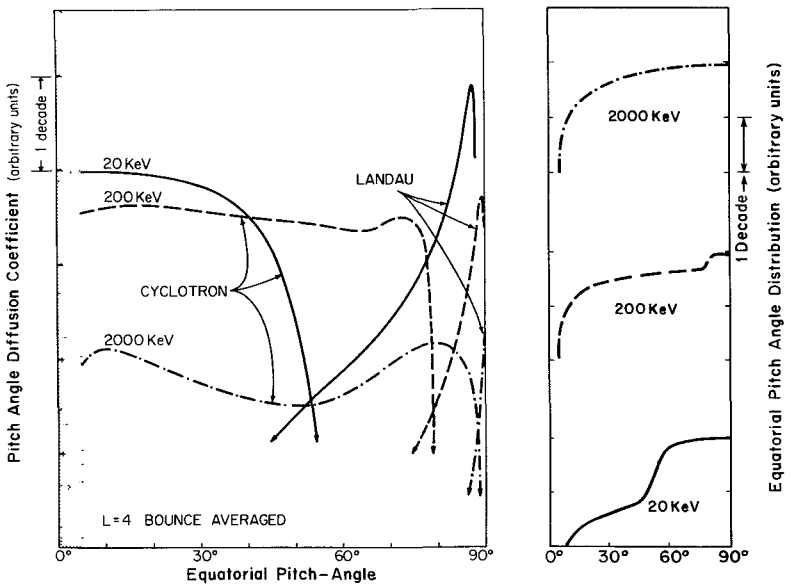


FIGURE 4 Bounce-orbit averaged cyclotron and Landau resonant pitch-angle diffusion coefficients as a function of equatorial pitch angle at L=4 for 20, 200, and 2000 keV electrons. For each energy, the line coding used for the Landau resonant diffusion coefficients corresponds to that for the sum of the cyclotron-harmonic resonances, where the particle energy is indicated. Equatorial pitch-angle distributions, calculated using the diffusion coefficients shown in the figure, are displayed on the right-hand side (from Lyons et al., 1972).

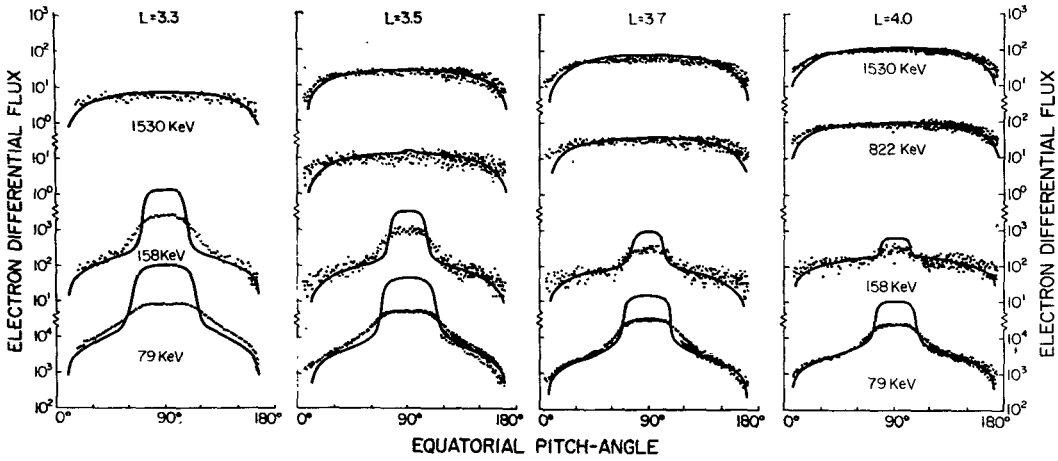


FIGURE 5 Comparison of calculated equatorial pitch-angle distributions with equatorial distributions observed within the electron slot. The data points (electrons/cm^2-sec-ster-keV) are plotted as a function of pitch angle at the L values and energies indicated on the figure. The vertical placement of the corresponding calculated distributions, shown by solid lines, are arbitrary on a logarithmic scale and have therefore been adjusted so as to best fit the data (from Lyons et al., 1972).

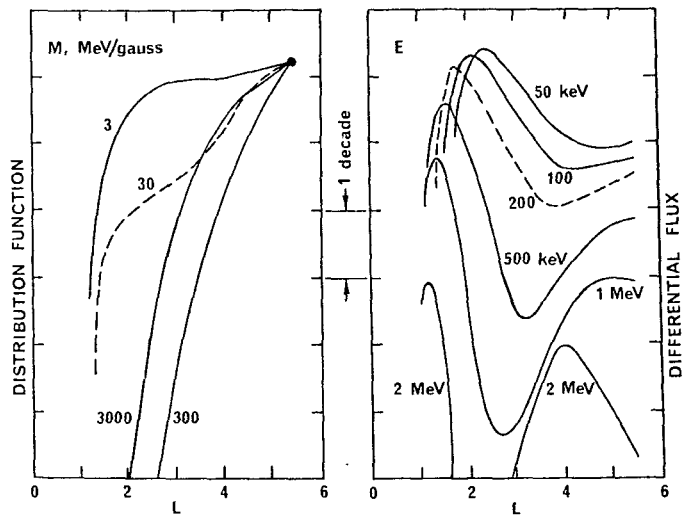


FIGURE 6 Equilibrium distribution function f of equatorially mirroring electrons versus L at constant first adiabatic invariant M, with each curve normalized to the same value at L=5.5 (left panel), and equilibrium differential flux j versus L at constant electron energy using the curves in the left panel and a prescribed energy spectrum at L=5.5 (right panel). (From Lyons and Thorne, 1973 as modified by Schulz, 1975.)

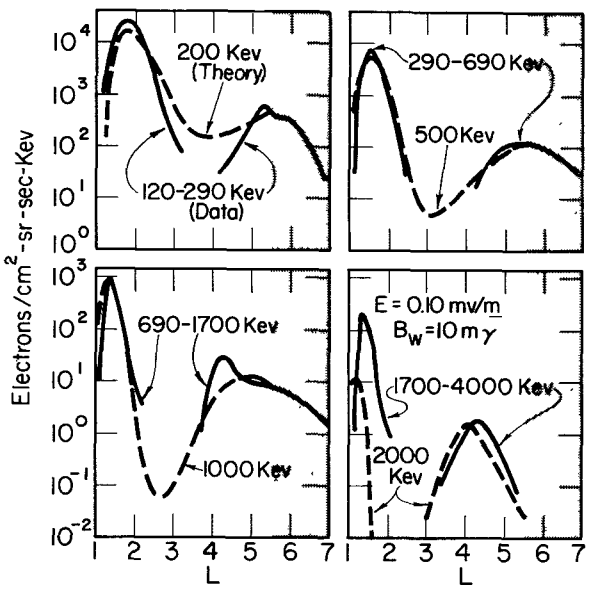


FIGURE 7 Comparison between calculated equilibrium electron structure (dashed lines) and quiet time observations of Pfitzer et al. (1966) (from Lyons and Thorne, 1973).

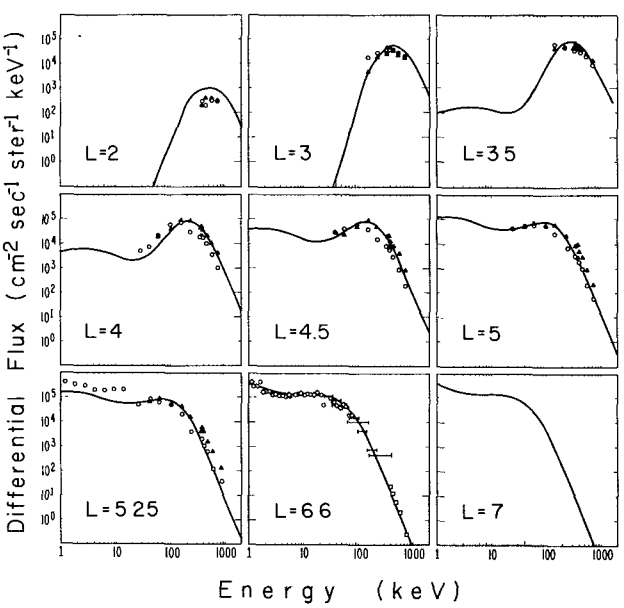


FIGURE 8 Comparison of theoretical and observed proton energy spectra in the kiloelectronvolt energy range at L-values of L=2, 3, 3.5, 4, 4.5, 5, 5.25, 6.6, and 7. The spectrum at L=6.6 constitutes the adopted boundary condition on the theoretical calculations and is taken from ATS-6 observations. The data at the lower L-values are from Explorer 45 orbits 97 (circles) and 667 (triangles) (from Spjeldvik, 1976).

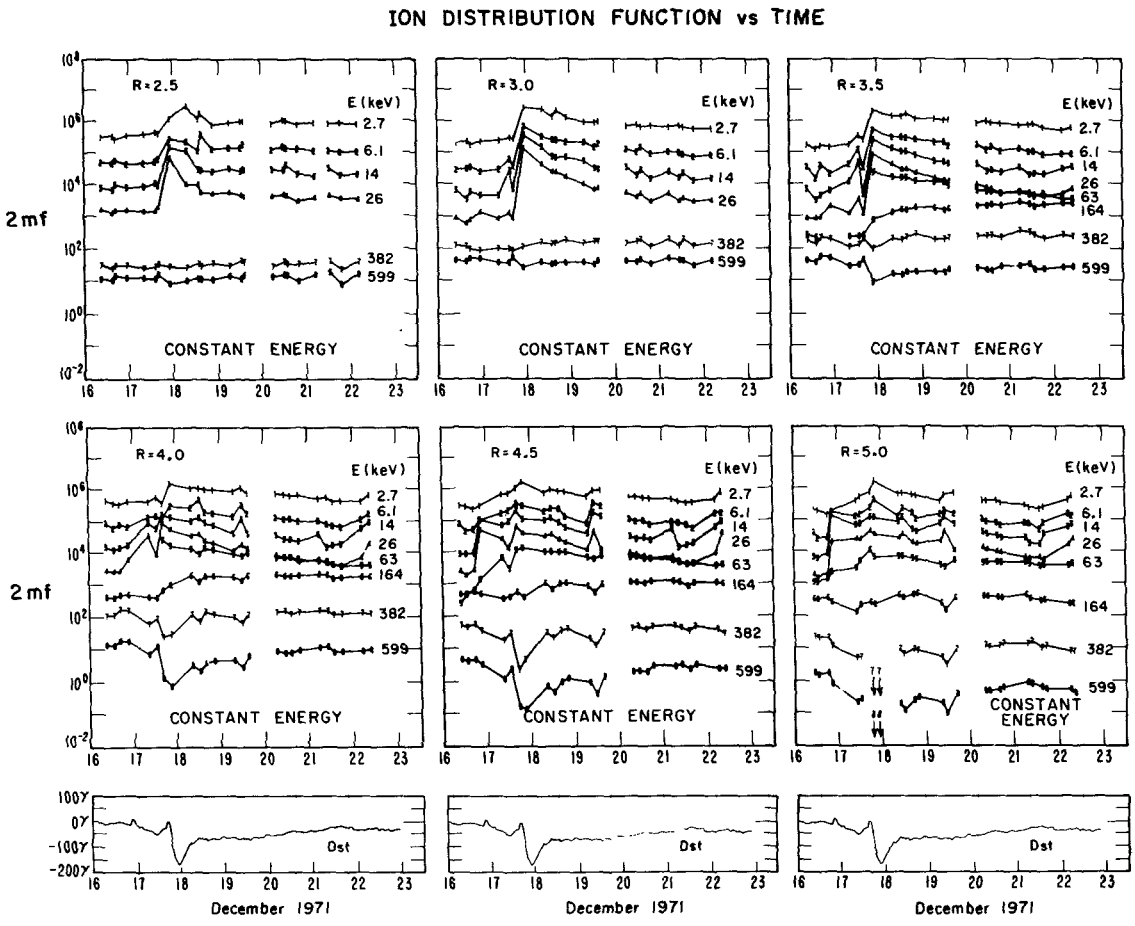


FIGURE 9 Distribution function, multiplied by 2 times the ion mass, for equatorially mirroring ions as a function of time throughout the period of the December 17, 1971 storm for 8 Explorer 45 energy channels. Dst is also shown. Radial distances R=2.5, 3.0, . . ., 5.0 are shown. *Note that the pre-storm intensities for energies ⩽ 26 keV are upper limits to the true intensities* (from Lyons and Williams, 1976).

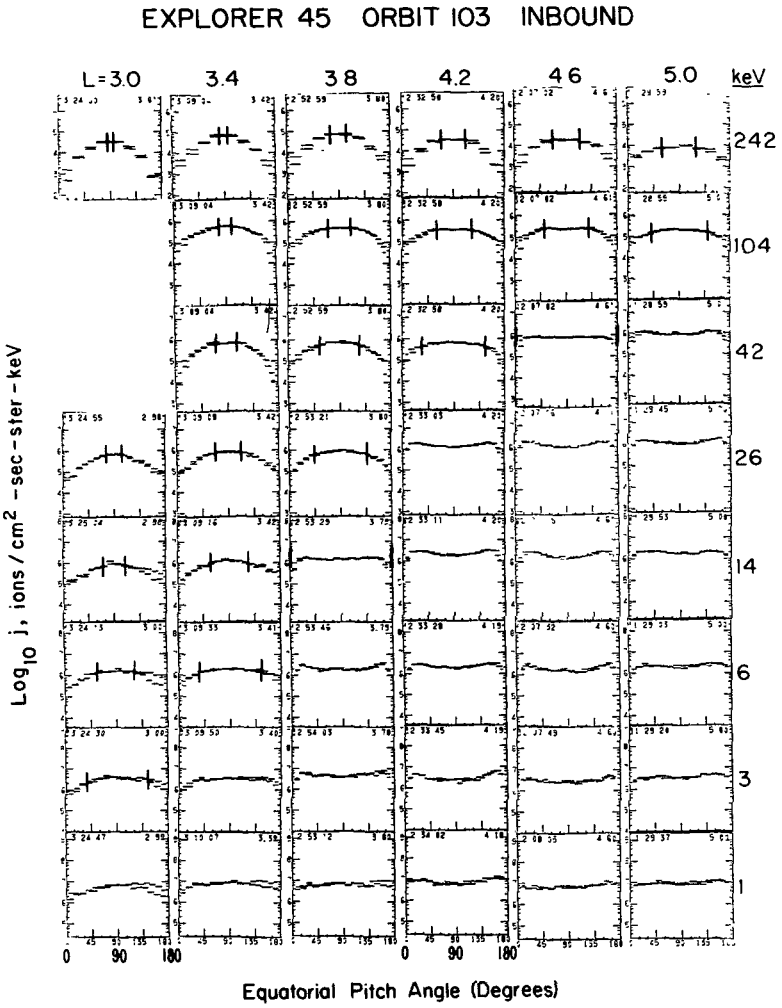


FIGURE 10 Equatorial ion pitch-angle distributions observed on Explorer 45 orbit 103 inbound, ~16 h after the minimum Dst of the December 17, 1971 storm main phase. Distributions are shown every 0.4 in L from L=3 to L=5, and selected ion energy channels are stacked vertically at each L. Elevated fluxes at pitch angles of 90-180° for energies ⩽14 keV are due to reflected sunlight. Ticks are at constant values of $E_\parallel$ for each L, with the chosen value of $E_\parallel$ at each L being equal to the energy of one of the Explorer 45 channels (from Joselyn and Lyons, 1976).

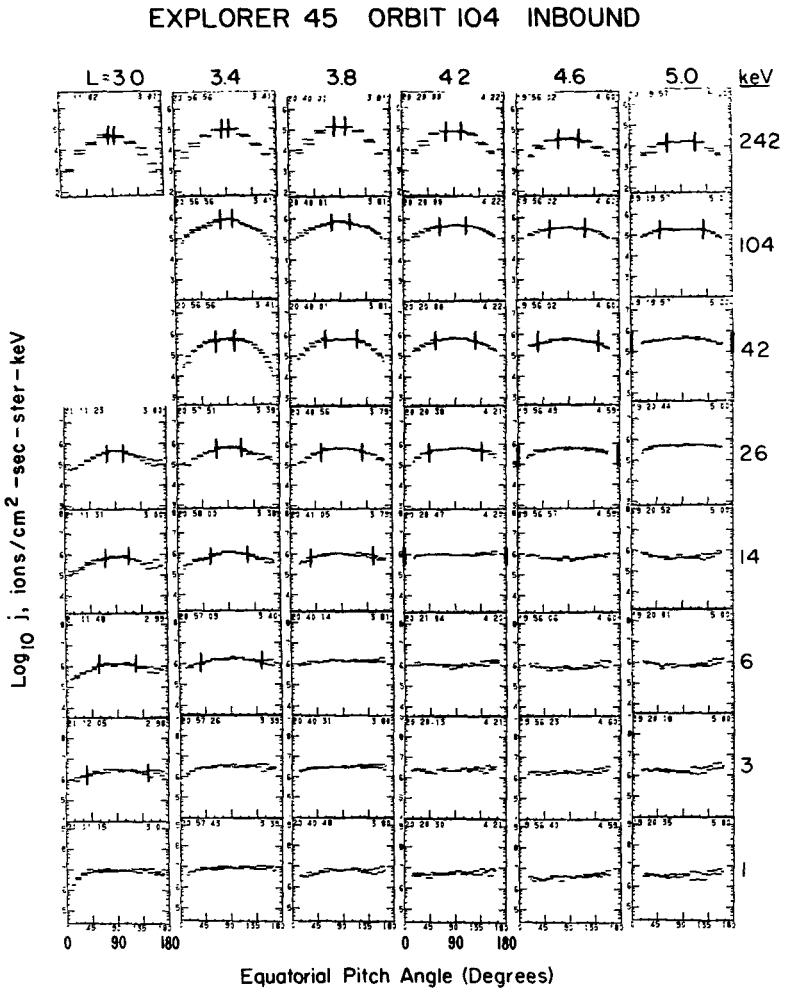


FIGURE 11 Same as Figure 10, except for Explorer 45 orbit 104 inbound, ~24 h after the minimum Dst of the storm main phase (from Joselyn and Lyons, 1976).

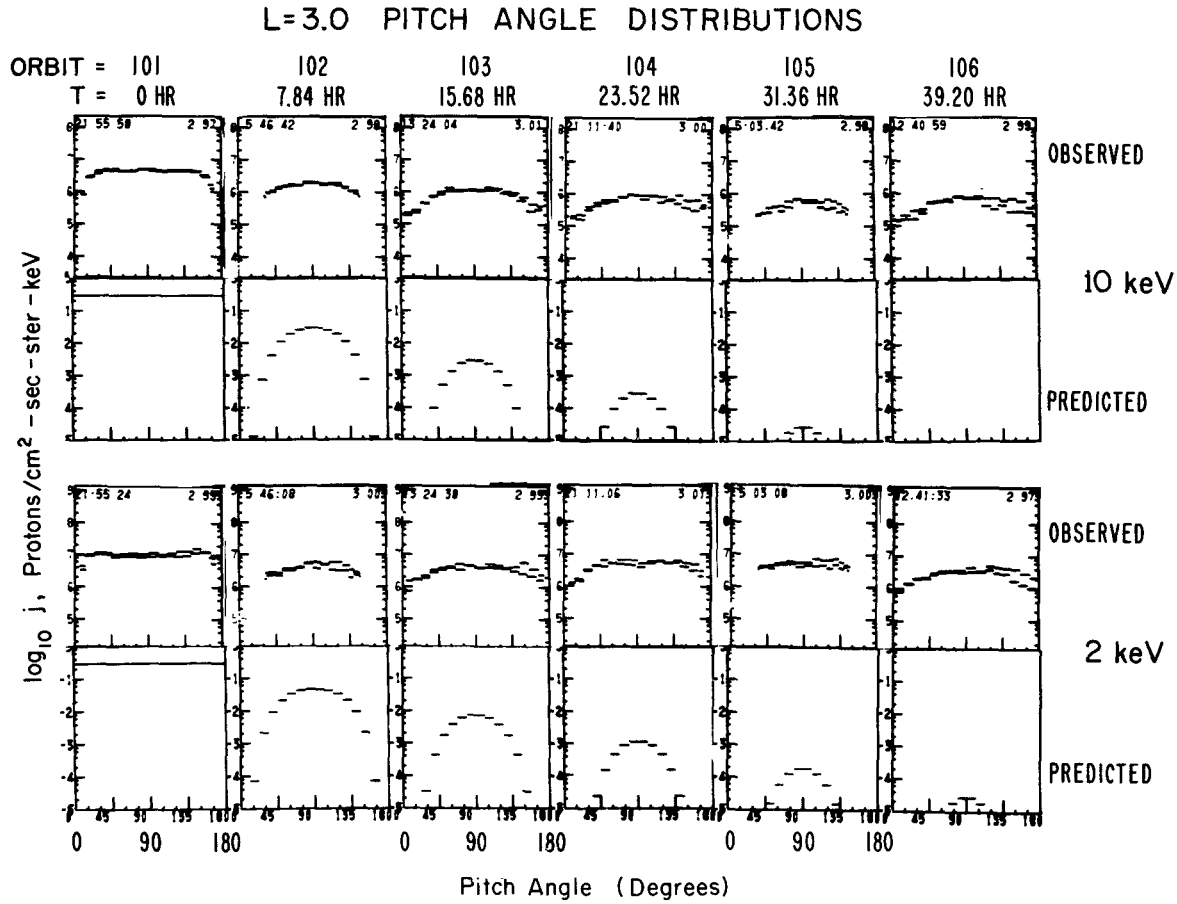


FIGURE 12 Equatorial pitch angle distributions at L=3 observed during the recovery phase of the Dec. 17, 1971 storm are compared with those expected to evolve from proton charge exchange with neutral hydrogen. Observations are shown from the inbound portion of Explorer 45 orbits 101, 102, . . ., 106 with the observations from orbit 101 being less than 1 hour following the minimum of Dst. The appearance of elevated fluxes over the pitch angle range of 90° to 180° is an instrumental error resulting from reflected sunlight. The pitch angle distributions expected to evolve from charge exchange were obtained by neglecting possible sources, assuming an isotropic proton pitch angle distribution at the time of the orbit 101 observations (t=0), and using the charge exchange lifetimes given by (Tinsley, 1976). The initial fluxes for the calculations were arbitrarily normalized to approximately $s \times 10^{-1}$ (from Lyons and Evans, 1976).

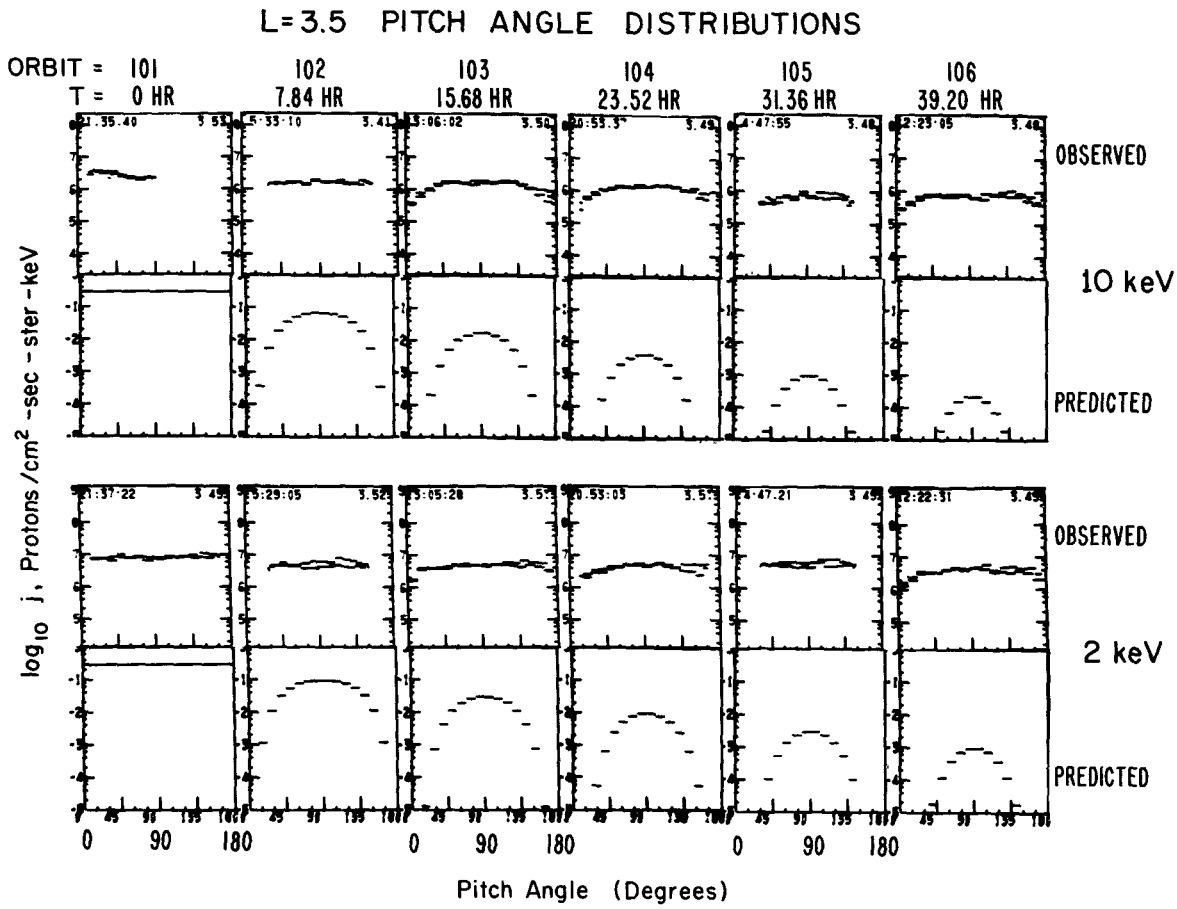


FIGURE 13 Same as Figure 12, except for L=3.5 (from Lyons and Evans, 1976).

REFERENCES

1. Chaflin, E. S., and R. S. White, A study of equatorial inner belt protons from 2 to 200 MeV, JGR, 79, 959, 1974.

2. Cornwall, J. M., Diffusion processes influenced by conjugate point wave phenomena, Radio Science, 3, 740, 1968.

3. Cornwall, J. M., Radial diffusion of ionized helium and protons: a probe for magnetosphere dynamics, JGR, 77, 1756, 1972.

4. Cornwall, J. M., On the role of charge exchange in generating unstable waves in the ring current, JGR, 1976 (in press).

5. Cornwall, J. M., F. V. Coroniti, and R. M. Thorne, Turbulent loss of ring current protons, JGR, 75, 4699, 1970.

6. Cornwall, J. M., F. V. Coronite, and R. M. Thorne, A unified theory for SAR arc formation at the plasmapause, JGR, 76, 4428, 1971.

7. Croley, D. R., Jr., M. Schulz, and J. B. Blake, Radial diffusion of inner zone protons: observational and variational analysis, 81, 585, 1976.

8. DeForest, S. E. and C. E. McIlwain, Plasma clouds in the magnetosphere, JGR, 76, 3587, 1971.

9. Dessler, A. J. and E. N. Parker, Hydromagnetic Theory of Geomagnetic Storms, JGR, 64, 2239, 1959.

10. Etcheto, J., R. Gendrin, J. Solomon, and A. Roux, A self-consistent theory of magnetospheric ELF hiss, JGR, 78, 8150, 1973.

11. Evans, D. S., Evidence for the low altitude acceleration of auroral particles, in Physics of the Hot Plasma in the Magnetosphere, Ed. by B. Hultqvist and L. Stenflo, Plenum Press, New York, 319, 1975.

12. Fälthammar, C.-G., Effects of time-dependent electric fields on geomagnetically trapped radiation, JGR, 70, 2503, 1965.

13. Hauge, R., and F. Söraas, Precipitation of >115 keV protons in the evening and forenoon sectors in relation to the magnetic activity, Planet. Space Sci., 23, 1141, 1975.

14. Helliwell, R. A., and J. P. Katsufrakes, VLF wave injection into the magnetosphere from Siple Station, Antarctica, JGR, 79, 2511, 1974.

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

16. Hultqvist, B., The ring current and particle precipitation near the plasmapause, Ann. Geophys., 31, 111, 1975a.

17. Hultqvist, B., Some experimentally determined characteristics of the turbulence in the magnetosphere, in Physics of the Hot Plasma in the Magnetosphere, ed. by B. Hultqvist and L. Stenflo, Plenum Press, New York, 291, 1975b.

18. Joselyn, J. A., and L. R. Lyons, Ion cyclotron wave growth calculated from observations of the proton ring current during storm recovery, JGR, 81, 2275, 1976.

19. Kennel, C. F., Consequences of a magnetospheric plasma, Rev. Geophys. 7, 379, 1969.

20. Kennel, C. F. and F. Engelmann, Velocity space diffusion from weak plasma turbulence in a magnetic field, Phys. Fluids, 9, 2377, 1966.

21. Kennel, C. F., and H. E. Petschek, Limit on stably trapped particle fluxes, JGR, 71, 1, 1966.

22. Lyons, L. R., Comments on pitch-angle diffusion in the radiation belts, JGR, 78, 6793, 1973.

23. Lyons, L. R., General relations for particle diffusion in pitch angle and energy, J. Plasma Phys., 12, 45, 1974a.

24. Lyons, L. R., Pitch angle and energy diffusion coefficients from resonant interactions with ion-cyclotron and whistler waves, J. Plasma Phys., 12, 417, 1974b.

25. Lyons, L. R. and D. S. Evans, The inconsistency between proton charge exchange and the observed ring current decay, JGR, 1976 (in press).

26. Lyons, L. R. and R. M. Thorne, The magnetospheric reflection of whistlers, Planet. Space Sci., 18, 1753, 1970.

27. Lyons, L. R. and R. M. Thorne, Equilibrium structure of radiation belt electrons, JGR, 78, 2142, 1973.

28. Lyons, L. R., R. M. Thorne, and C. F. Kennel, Pitch-angle diffusion of radiation belt electrons within the plasmasphere, JGR, 77, 3455, 1972.

29. Lyons, L. R., and D. J. Williams, The quiet time structure of energetic (35-560 keV) radiation belt electrons JGR, 80, 943, 1975a.

30. Lyons, L. R., and D. J. Williams, The storm and post-storm evolution of energetic (35-560 keV) radiation belt electrons, JGR, 80, 3985, 1975b.

31. Lyons, L. R. and D. J. Williams, Storm associated variations of equatorially mirroring ring current protons, 1-800 keV, at constant first adiabatic invariant, JGR, 81, 216, 1976.

32. McIlwain, C. E., Auroral electron beams near the magnetic equator, in Physics of the Hot Plasma in the Magnetosphere, ed. by B. Hultqvist and L. Stenflo, Plenum Press, New York, 91, 1975.

33. Nakata, N. P. and G. D. Mead, Diffusion of protons in the outer radiation belt, JGR, 70, 4777, 1965.

34. Pfitzer, K., S. Kane and J. R. Winckler, The spectra and intensity of electrons in the radiation belts, Space Res., 6, 702, 1966.

35. Roberts, C. S. and M. Schulz, Bounce resonant scattering of particles trapped in the earth's magnetic field, JGR, 73, 7361, 1961.

36. Russell, C. T., R. E. Holzer, and E. J. Smith, Observations of ELF noise in the magnetosphere, 1, Spatial extent and frequency of occurrence, JGR, 74, 755, 1969).

37. Schulz, M., Geomagnetically trapped radiation, Space Sci. Reviews, 17, 481, 1975.

38. Schulz, M. and L. J. Lanzerotti, Particle Diffusion in the Radiation Belts, Springer, Heidelberg, 1974.

39. Sharp, R. D., R. G. Johnson, and E. G. Shelley, The morphology of energetic O^+ ions during two magnetic storms: Temporal variations, JGR, 81, 3283, 1976a.

40. Sharp, R. D., R. G. Johnson, and E. G. Shelley, The morphology of energetic O^+ ions during two magnetic storms: Latitudinal variations, JGR, 81, 3292, 1976b.

41. Smith, P. H, and N. B. Bewtra, Role of charge exchange decay in the energy coupling between the magnetosphere and the ionosphere, presented at ISSTP Symposium, Boulder, Colorado, 1976.

42. Southwood, D. J., J. W. Dungey, and R. J. Etherington, Bounce resonant interaction between pulsations and trapped particles, Planet. Space Sci., 17, 349, 1969.

43. Spjeldvik, W. N., Equilibrium structure of equatorially mirroring radiation belt protons, submitted to JGR, 1976.

44. Spjeldvik, W. N., and T. A. Fritz, Energetic ionized helium in the quiet time radiation belts, (in preparation) 1976.

45. Thorne, R. M., E. J. Smith, R. K. Burton, and R. E. Holzer, Plasmaspheric Hiss, JGR, 78, 1581, 1973.

46. Thorne, R. M., E. J. Smith, K. J. Fiske, and S. R. Church, Intensity variations of ELF hiss and chorus during isolated substorms, Geophys. Research Letters, 1, 193, 1974.

47. Tinsley, B. A., Evidence that the recovery phase ring current consists of helium ions, JGR, 1976 (in press).

48. West, H. I., Jr., Advances in magnetospheric physics: energetic particles, Rev. Geophys, 13, 943, 1975.

49. West, H. I., Jr., R. M. Buck, and J. R. Walton, Electron pitch angle distributions throughout the magnetosphere as observed on OGO-5, JGR, 78, 1064, 1973.

50. White, R. S., High-energy proton radiation belt, Rev. Geophys., 11, 595, 1973.

51. Williams, D. J., and L. R. Lyons, The proton ring current and its interaction with the plasmapause: storm recovery phase, JGR, 79, 4195, 1974a.

52. Williams, D. J. and L. R. Lyons, Further aspects of the proton ring current interaction with the plasmapause: main and recovery phases, JGR, 19, 4791, 1974b.

MAGNETOSPHERIC PLASMA WAVES

by

Stanley D. Shawhan

ABSTRACT

A brief history of plasma wave observations in the Earth's magnetosphere is recounted and a classification of the identified plasma wave phenomena is presented. The existence of plasma waves is discussed in terms of the characteristic frequencies of the plasma, the energetic particle populations and the proposed generation mechanisms. Examples are given for which plasma waves have provided information about the plasma parameters and particle characteristics once a reasonable theory has been developed. Observational evidence and arguments by analogy to the observed Earth plasma wave processes are used to identify plasma waves that may be significant in other planetary magnetospheres. The similarities between the observed characteristics of the terrestrial kilometric radiation and radio bursts from Jupiter, Saturn and possibly Uranus are stressed. Important scientific problems concerning plasma wave processes in the solar system and beyond are identified and discussed. Models for solar flares, flare star radio outbursts and pulsars include elements which are also common to the models for magnetospheric radio bursts. Finally, a listing of the research and development in terms of instruments, missions, laboratory experiments, theory and computer simulations needed to make meaningful progress on the outstanding scientific problems of plasma wave research is given.

1. INTRODUCTION

The term plasma wave is used to denote all waves which are generated in a plasma or which have their wave characteristics significantly modified by the presence of the plasma. These waves may be electromagnetic, electrostatic or magnetosonic and are generated by the conversion of plasma electrical and particle kinetic energy into wave energy through a variety of plasma-particle processes. In turn, these waves may interact with the particles and modify the particle populations within the plasma.

Plasma waves are generated in many natural plasma systems: planetary ionospheres and magnetospheres, the solar wind, the solar atmosphere, interstellar space, stellar atmospheres, pulsars, flare stars, galaxies and quasars. An upper limit to the frequency range of plasma waves generated in these different systems is given approximately by the maximum frequency for which the plasma can respond to the presence of the wave fields -- the electron plasma frequency. Within the solar system this electron plasma frequency is 10 kHz for the solar wind, ~ 10 MHz for planetary ionospheres, and 10 GHz for the solar corona. In laboratory plasma devices it may range from 10 MHz to 10 GHz. Radio astronomy techniques are used to study plasma waves that have escaped solar and cosmic plasma systems from above the solar wind plasma frequency ~ 10 kHz to the millimeter wavelength range ~ 1000 GHz. Space

plasma wave instruments measure plasma waves generated locally in the plasma below the electron plasma frequency.

The emphasis of this paper is placed on the review and interpretation of plasma wave observations made in the Earth's magnetosphere. Observations of plasma waves from other planetary magnetospheres and cosmic plasma systems are discussed and interpreted in analogy to the Earth's processes. In situ space instruments have provided a survey of the wave phenomena, plasma parameters and energetic particle characteristics throughout most of the Earth's magnetosphere. With this information it has been possible to classify plasma wave phenomena as to frequency range, wave mode, and region of occurrence within the magnetosphere for both trapped and escaping waves. Some of the associated plasma and particle information with which to develop detailed theories of wave generation and propagation and of wave-particle interactions is also available. Gaps in the wave and particle information are to be filled in part by the ISEE Missions and the proposed Dynamics Explorer Missions. Less extensive in situ measurements are becoming available for the solar wind and the other planetary magnetospheres. Presently, however, most deductions about plasma wave processes in solar system plasmas other than the Earth's magnetosphere and in cosmic plasma systems are based on observations of escaping plasma waves and characteristics of IR, optical, UV, x-ray and γ-ray emissions. Therefore, we take the approach that the interpretation of plasma wave phenomena in the Earth's magnetosphere together with existing experimental evidence, laboratory experiments, computer plasma simulations, and theoretical extensions can be used to evaluate plasma wave phenomena in other plasma systems. Also,

the magnetosphere is itself a convenient plasma laboratory of cosmic scale in which to perform experiments, active and passive, related to plasma wave processes.

In this paper a brief history of plasma wave observations in the Earth's magnetosphere is recounted and a classification of the identified plasma wave phenomena is presented. The existence of plasma waves is discussed in terms of the characteristic frequencies of the plasma, the energetic particle populations, and the proposed generation mechanisms. It is pointed out that plasma waves can provide information about the plasma parameters and particle characteristics once a reasonable theory has been developed for the plasma wave process. Observational evidence and arguments by analogy to Earth are used to identify plasma wave processes that may be significant in other planetary magnetospheres and cosmic plasma systems. Important scientific problems concerning plasma wave processes in the solar system and beyond are identified and discussed. Finally, a listing of the research and development in terms of instruments, missions, laboratory experiments, theory and computer simulations needed to make meaningful progress on the outstanding scientific problems is given.

2. THE MENAGERIE OF PLASMA WAVES

2.1. Brief History of Observations

Naturally occurring plasma waves associated with the Earth's ionosphere and magnetosphere have been observed and studied with ground-based instrumentation since the late 1800's.[47, 51] These phenomena include whistlers, hiss, chorus, discrete emissions and geomagnetic micropulsations (see Table 1). In 1894 Preece reported hearing what may have been whistlers and chorus associated with an auroral display by using a telephone receiver connected to a telegraph line. Rapid variations in the Earth's magnetic field, as measured with a magnetometer, were reported by Terada in 1917. Interpretation of these early observations lead to the realization that the Earth had an extensive magnetic field, significant ionization to great distances and a population of energetic particles (at least after enhanced solar activity).

In the years encompassing the International Geophysical Year (IGY), networks of magnetometers and of receivers in the VLF (3-30 kHz) frequency range were set up at various places in the world to make systematic measurements of these wave phenomena in terms of local time, latitude, magnetic conjugacy, solar activity and geomagnetic disturbances. From this intensive effort the detailed morphologies of the observed micropulsations, whistlers and VLF emissions were established and theories for interpreting these plasma waves were developed. Once the frequency-time dispersion relationship for whistlers was understood,

for example, it was possible to use whistlers to determine the electron density in the equatorial plane. From whistler measurements it was found that the radial electron density profile exhibited an abrupt order of magnitude decrease in the range of 3-6 earth radii. This region of density decrease is now identified as the plasmapause (PP in Figure 1) and the high density region at lower radial distances as the plasmasphere (PS in Figure 1).

After the IGY it was well known that whistlers were generated by lightning discharges in the Earth's atmosphere and the dispersion characteristics were a result of propagation through the magnetospheric plasma approximately along magnetic field lines. The source mechanisms for micropulsations and the other VLF waves were not well understood but these waves were thought to be generated by energetic particles (electrons and ions) interacting with the plasma and plasma waves in the magnetosphere. With the advent of scientific sounding rockets and of scientific satellites in 1957 it became possible to make measurements of the plasma waves (frequency spectrum, wave mode, wave vector), the plasma (density, ion species, temperature) and the energetic particles (energy spectrum, pitch angle distribution, ion species) within the magnetosphere itself. From these measurements it became possible to develop better theories of wave generation, propagation and interaction with particles.

Vanguard III in 1959 carried a magnetometer with frequency response to 6 kHz with which the first satellite observations of whistlers and other VLF waves were made.[47] Since that time a variety of magnetometers and plasma wave receivers have been carried to nearly all regions

of the Earth's magnetosphere and into the solar wind between 0.3 and 1 astronomical units.[83] Most of these instruments have measured the spectral characteristics of the waves using a single axis sensor - search coil, magnetic loop or electric dipole antenna. From the spectral information and the change of orientation of the sensor with respect to the magnetic field, it has been possible to classify the observed waves as to frequency range and occurrence probability, and to deduce if the wave vector is nearly along the geomagnetic direction or not. On a few satellites (i.e., FR1, OGO 5, Injun 5, IMP 6) wave measurements have been obtained for more than one wave component (of the possible 3 electric and 3 magnetic) simultaneously. From these measurements it has been possible to specify the wave normal vector direction, the wave polarization or the component of the Poynting vector along the geomagnetic field to distinguish between upgoing and downgoing waves.[93] These methods have been applied particularly to micropulsations, hiss, chorus, saucers, upstream whistler mode waves and whistlers. For the non-thermal continuum and kilometric radiation which are not constrained to propagate along the geomagnetic field direction, it has been possible to use the nulls in the antenna pattern rotating with the spacecraft (i.e., RAE 1, IMP 6, IMP 7, Hawkeye 1) and the lunar occultation technique[1] (RAE 2) to actually locate the source regions.

2.2 Locations and Characteristics

Based on the extensive rocket and satellite measurements, distinct plasma wave types have been identified in addition to those observed on the ground. Figure 1 gives a noon-midnight cross-sectional

view of the magnetosphere with the magnetospheric regions identified. A detailed discussion of the magnetospheric configuration is given by Juan Roederer ("The Earth's Magnetosphere") in this report. Indicated in Figure 1 are the regions having the highest probability for observing the various types of plasma waves. In general, the noise types are similar in the north and south half of the magnetosphere (although not shown in Figure 1), however, the types of noise that occur in the different magnetospheric regions are somewhat distinct. This distinction is understandable because a particular plasma wave type is generated by a particular particle population under particular plasma conditions and these populations and conditions vary between regions in the magnetosphere.

More detailed information about each type of plasma wave identified in the Earth's magnetosphere is given in Tables 1, 2, 3. Each table lists the noise type, its location in the magnetosphere, the frequency range for which it is observed, the characteristic frequency of the plasma associated with the waves, key literature references, a brief description of the important wave characteristics and a statement of the proposed source mechanism. Table 1 includes the electromagnetic plasma waves -- waves with a detectable magnetic field component. The geomagnetic micropulsations category includes a large class of phenomena -- both magnetohydrodynamic waves and transient disturbances -- associated with variations in the geomagnetic field. Because the wave periods and wavelengths are long, these phenomena are best observed from the Earth. However, the ATS geostationary series of spacecraft have yielded meaningful observations near the equatorial plane. All the other wave types, except the trapped continuum and UHR noise are classified as whistler-

mode waves because the wave characteristics can be described by the whistler dispersion relation for a plasma with a magnetic field. These waves tend to propagate along the magnetic field lines although in certain frequency ranges the waves can propagate across the field to populate regions remote from the source. Many of these waves can be observed only within the magnetosphere itself. The electron whistler classification includes a variety of types such as nose whistlers, multiple hop ducted whistlers, ion-cutoff whistlers, subprotonospheric whistlers, magnetically reflected whistlers and walking trace whistlers.[83] These are distinguished by their frequency-time characteristic due to propagation differences.

In Table 2 are listed the identified electrostatic plasma waves -- waves with an electric field but without a detectable magnetic field component.[43] These types of plasma waves are generally found in regions where collisions between the plasma and the neutral gas are important (Farley instabilities) or where there is a mixture of hot and cold plasma. These waves are entirely trapped within the plasma and generally are observed close to their region of generation because they generally have a low group velocity and are strongly damped. Two types of plasma waves have been observed to escape from the magnetosphere. These waves, listed in Table 3, are generated near the electron plasma frequency and are observed to be propagating in an electromagnetic mode (perhaps by coupling from an electrostatic mode) that can escape the plasma. Because of these waves, the Earth exists as a significant (10^9 watts) source of radio emission within the solar system[44].

Not all the waves observed in the magnetosphere are natural. Three types of waves given in Table 4 are man-made. The power system harmonic radiation is accidental but it apparently leads to wave growth and increased precipitation of trapped electrons within the magnetosphere. These induced effects may have some long term consequences as discussed by P. A. Sturrock in this report ("Impacts of Solar System Environment on Man and Man in the Environment"). The other waves are stimulated by a VLF transmitter and by a rocket-borne electron accelerator in order to study wave-particle interactions in the magnetosphere under somewhat controlled conditions.

3. GENERATION, PROPAGATION AND INTERACTION OF PLASMA WAVES

Interpretation of simultaneous *in situ* plasma wave, plasma and particle measurements in terms of the particular plasma wave processes is complicated by a number of factors:

i. The observed plasma waves may be a combination of waves locally produced and of waves that have propagated from remote source regions,

ii. Waves that have propagated to the observation point may have wave characteristics different from those in the source region due to reflection, refraction, wave-wave interactions and wave-particle interactions,

iii. Waves and particles in the source region may be interacting strongly at the saturation limit of the instability, so that the particle distribution function has been modified into a nearly stable state. It is therefore difficult to associate the final state with the wave mode and the particle distribution that caused the instability during the linear growth phase,

iv. The observed particle distributions may be a superposition of the final state distribution and the distribution of particles newly injected into the source region,

v. In any series of observations with a single spacecraft it is usually difficult to separate the temporal and spatial variations of the measured parameters, and

iv. A given satellite may not be suitably instrumented to measure all the desired parameters with sufficient dynamic range and time resolution.

However, by modeling the plasma to solve for the possible wave modes and propagation characteristics, by collecting the data from a variety of instruments on various spacecraft and by comparing the measured features to theoretical predictions, computer simulations and laboratory experiments, it has been possible to identify certain wave modes, to classify the most likely wave generation mechanism, and to interpret the observed wave-particle interactions.

3.1 Characteristic Plasma Wave Frequencies

A plasma having electrons, ions and neutrals of finite temperatures and permeated by a magnetic field can support a variety of electromagnetic, electrostatic and magnetosonic wave modes that cannot exist in free-space. The range of frequencies for which these various modes are observed to exist can be understood in terms of the frequencies characteristic of the plasma. To a first approximation these frequencies can be obtained from solutions to the wave equation for an infinite, homogeneous plasma with cold electrons and ions and a static magnetic field assuming plane waves of the form exp $(i\ (\underline{k} \cdot \underline{x} - \omega t))$.[12, 102]

A representation of this solution is given in Figure 2 in terms of the index of refraction of the plasma

$$n = c|\underline{k}|/\omega$$

where $\underline{k}$ is the wave vector and ω is the angular wave frequency ($\omega = 2\pi f$) for the wave frequency f. With a magnetic field, the plasma is anisotropic so that the solution depends also on the angle between the wave vector $\underline{k}$ and the magnetic field which is called the wave normal angle. Those waves below the plasma frequency with the wave normal angle near 0°, have the wave energy primarily in the wave magnetic field and the waves are circularly polarized. Waves with wave normal angles near 90° tend to have their energy primarily in the electric field component and to be linearly polarized. The various resonance ($n^2 \rightarrow \infty$) and cutoff ($n^2 \rightarrow 0$) frequencies can be expressed in terms of the plasma wave frequencies related to the electron density (N), ion species (mass = M_i) and magnetic field strength (B):[12, 37, 102]

f_p^- , electron plasma frequency $\simeq 9 \times 10^3$ (N, electrons/cm^3)$^{1/2}$ Hz

f_p^+ , ion plasma frequency $\simeq f_p^- (M_e/M_i)^{1/2}$ Hz

f_g^- , electron gyrofrequency $\simeq 2.8 \times 10^6$ (B, gauss) Hz

f_g^+ , ion gyrofrequency $\simeq f_g^- (M_e/M_i)$ Hz

f_{LHR} , low hybrid resonance frequency =

$$\left[\frac{M_i}{M_e}\left(\frac{1}{(f_p^-)^2}+\frac{1}{(f_g^-)^2}\right)\right]^{-1/2}$$

f_{UHR} , upper hybrid resonance frequency =

$$\left[(f_p^-)^2+(f_g^-)^2\right]^{1/2}$$

Representative curves for the variations of some of these characteristic frequencies with radial distance in the Earth's magnetosphere are shown in Figure 3 for an auroral field line and for a cut along the equatorial plane. Within the magnetosphere these characteristic frequencies can vary over three orders of magnitude. This variation accounts for the wide frequency range observed for a given plasma wave phenomenon.

The cold plasma dispersion relation has been successful in describing the propagation characteristics of the plasma wave types listed in Tables 1 and 3. These are also listed on Figure 2. However, the magnetospheric background plasma has a finite temperature which may be different for electrons and ions and may range from 500°K to 10^7°K. When the thermal velocity becomes comparable to the wave phase velocity (c/n), the cold plasma description is no longer valid. Introducing thermal effects into the dispersion relation allows the possibility of electrostatic modes which occur near harmonics of the electron and ion gyrofrequencies and near the ion plasma frequency. Waves related to these modes are included in Table 2 and are listed in parentheses on Figure 2.

3.2 Generation Mechanisms

A number of excellent books and articles exist which review our present deductions concerning wave generation mechanisms.[33, 34, 43, 47, 51, 88, 89, 107] Also, for each of the wave phenomena listed in Tables 1, 2 and 3 the proposed generation mechanism is given along with references to representative observational and theoretical papers.

The energy source for ion cyclotron and electron whistlers is well identified with lightning discharges in the upper atmosphere. This energy is derived from solar heating of the atmosphere and the Earth's rotational motion which drives the wind systems to effect the differential charging between clouds and the Earth's surface. The momentary discharge current has frequency components extending to tens of megahertz in frequency. Under certain conditions some of this energy couples into the magnetosphere through the ionosphere and propagates in plasma wave modes associated with the electrons and with the ions. These whistler waves may be amplified or damped through electron and ion cyclotron resonances with the energetic particle populations.

The energy for plasma waves generated internal to the magnetosphere comes ultimately from the solar wind flow past the Earth and the Earth's rotation. This energy is dispersed throughout the magnetosphere in terms of electric fields and particle kinetic energy and converted to wave energy through changes in the particle distribution functions. Particle distributions unstable to wave growth may exist at a given location in the magnetosphere due to processes such as plasma diffusion and convection across magnetic field lines, selective pitch

angle scattering into the atmospheric loss cone, energy selective gradient drifts, energy selective acceleration, resistive energy dissipation or mixture of hot and cold ionosphere and solar wind plasmas.[88] Such processes are discussed more in detail in the papers by Roederer, Lyons, Haerendel and Sonnerup (this report).

The resulting non-Maxwellian distributions may be continuously unstable or perturbed into an unstable state under geomagnetic storm conditions. For the purpose of discussion in this paper we divide the proposed plasma wave generation mechanisms into two general categories: gyroresonance and streaming.

3.2.1 Gyroresonance Instabilities

Particle distributions that contain free energy in the component transverse to the magnetic field direction (anisotropic temperature distributions $T_\perp > T_{||}$ or loss cone distributions, for example) can give up energy to waves (and vice-versa) under the condition[33]

$$f - k_{||} v_{||}/2\pi + sf_g^{\pm} = 0, \quad s = \pm 1, \pm 2, \ldots.$$

In this expression $k_{||}$ and $v_{||}$ are the wave vector and component of particle velocity parallel to the magnetic field direction, f is the wave frequency and f_g the ion or electron gyrofrequency around the field line, and s gives the harmonic number. Interacting particles undergo pitch angle diffusion which may cause them to be precipitated into the atmosphere (see Lyons) or energy diffusion which results in a harder spectrum for locally trapped particles.

Waves associated with the ion cyclotron frequency (also called gyrofrequency) are thought to be created just inside the plasmapause boundary where energetic ions from the magnetotail are injected into the cold plasma of the plasmasphere. Even though these waves have not been definitely observed indirect identifications have been made, for example, from ATS data.[9] Strong theoretical arguments for their existence can also be made.[56]. It may be that the higher frequency micropulsations (Pc1 and Pi1) are also generated by this process just outside the plasmapause boundary.[23, 51] Lion's roar seems to be explained by this ion cyclotron resonance due to 10 keV protons streaming through the magnetosheath particularly during geomagnetic storm periods.[98] Also ion cyclotron waves associated with streaming ions were observed in the polar cusp region with OGO 5.[30]

ELF hiss, chorus and discrete VLF emissions are electromagnetic wave phenomena which seem to result from the electron gyroresonance but under different conditions. ELF hiss is observed nearly continuously throughout the plasmasphere but is thought to be generated by the trapped electron population in the outer radiation zone just inside the plasmapause.[106] Chorus, however, is generated in the region between the plasmapause and the magnetopause principally on the dayside of the Earth during increased magnetic activity by electrons in the range of 5-150 keV[13] and very near the equatorial plane on the nightside.[109] The sporadic occurrence of the chorus tones suggests that the necessary generation conditions are met sporadically in time or that the propagation paths are spatially distributed. The large

amplitude and burst-like nature of chorus suggests a non-linear treatment of the emission process. Discrete emissions are rather rare and their generation seems to be similar to chorus but to require special conditions so that particle bunches moving along the field lines can interact with the waves for an extended period and then repeat the interaction periodically.[47] Both ELF hiss and chorus are associated with electron precipitation process (see Lyons).

At the plasmapause boundary and beyond, electrostatic plasma waves with frequencies near $3/2\ f_g^-$, $5/2\ f_g^-$, etc. have been observed particularly during magnetic storm time. Also, similar electrostatic electron cyclotron harmonic waves have been occasionally observed near the plasma sheet in the tail.[45] These waves can occur in regions for which the particle distribution function increases with the perpendicular velocity. This condition can be met by a loss-cone pitch angle distribution, by a beam or by large temperature anisotropy and a ratio of $f_p^-/f_g^- > 1$. It may be that these waves control the precipitation of moderate energy (1-10 keV) electrons in the outer magnetosphere.[33] Direction finding measurements on the nonthermal continuum radiation found throughout the outer magnetosphere suggests that the source region is near the plasmapause boundary on the morning side. In some cases it is closely associated with the electrostatic electron cyclotron noise suggesting that the continuum electromagnetic noise has coupled from this electrostatic noise. It has also been suggested, however, that this noise is due to gyrosynchrotron radiation from the high energy radiation belt electrons (in analogy to the Jovian decimetric radiation).[42]

A doppler shifted electron cyclotron emission mechanism[71] has been proposed for the kilometric radiation observed to be emitted along auroral field lines.[1, 41] This emission is thought to result from the same energetic electrons which cause the auroral optical emissions. Both emission and amplification of the waves is required to explain the observed power levels representing 10% of the particle energy flux.[41] Waves emitted by this process would be righthand polarized.[71]

3.2.2 Streaming Instabilities

Under this category we include the types of instabilities which are classed as current driven, beam, two-stream and drift wave. Common to these instabilities is a distribution function for which at least one component of the plasma is streaming through the other thermal and energetic components. As an example, the distribution may have a "bump on the tail" -- this distribution function, compared to a Maxwellian, has an enhanced population at some velocity different from the average velocity. This sort of distribution is unstable, according to the Penrose criterion, if there is a minimum in the distribution function $F(v)$ at a velocity v^* such that

$$\int_{-\infty}^{+\infty} \frac{F(v)-F(v^*)}{(v-v^*)^2}\, dv > 0 .$$

If this distribution is unstable then energy is extracted from the enhanced population with $v > v^*$. In general, plasma waves are created with phase velocities near v^*, the streaming velocity:

$$v_p = c/n \approx v^* .$$

Note that for frequencies near resonances in the plasma (electron and ion cyclotron frequency, lower and upper hybrid frequency and plasma frequency) the index of refraction goes toward infinity giving a low phase velocity making possible the interaction with low energy particles.

a) Current Driven. At the bow shock and magnetosheath, in the boundary layer of the tail plasma sheet and along auroral field lines, the few magnetometer and plasma particle probe measurements indicate the presence of field aligned currents and the presence of streaming ions (which may be carrying the current). The observed bow shock turbulence, tail broadband electrostatic noise and auroral field line turbulence have common features and may all plausibly be generated by an electrostatic ion cyclotron or an ion acoustic instability associated with the streaming ions. The electromagnetic bow shock plasma waves and the magnetic noise bursts observed on auroral field lines and near the neutral sheet in the tail are intimately associated with the presence of the electrostatic noise. Either this electromagnetic noise is coupled from the electrostatic noise through a wave-wave interaction or this noise is generated directly by an electromagnetic instability associated with the field aligned currents.[33, 43, 80, 88, 89]

Instabilities related to currents moving across the field lines are possible in the ionosphere in the presence of collisions. As an example, collisions in the E-region provide sufficient electron

conductivity to support transverse electric fields and to cause the electrons to stream at an E/B drift velocity with respect to the neutrals (and ions). When this drift exceeds the ion-acoustic velocity, an instability occurs producing observable electrostatic waves,[77] and corresponding density fluctuations which are seen as echos on backscatter radars (see chapter by Farley).

b) Beam. Particularly in the auroral regions, the precipitating energetic electrons ($\leq$ 15 keV) often constitute a field-aligned electron beam. This beam can emit plasma waves by the incoherent Cerenkov and cyclotron processes and the emitted waves can be further amplified by a coherent beam process. Noise bands are observed in the magnetosphere at the lower hybrid frequency f_{LHR}[64] and the upper hybrid frequency f_{UHR}[74].

It seems that these noise phenomena can be understood as incoherent Cerenkov radiation due to rather low energy electrons.[74, 105] Also, it may be that at least the LHR noise band is due in part to trapped and dispersed electron whistler waves.[21]

VLF hiss and saucers (see Table 1) are found in nearly the same region of the low altitude (~ 2000 km) auroral zone as illustrated with data from Injun 5 in Figure 4. The VLF hiss is associated with an intense flux of precipitating auroral electrons with energies in excess of 10 keV and also with an ELF noise band which is similar to lion's roar.[38] The saucers, however, have no apparent association with particles above 5 eV.[52] With Injun 5 the VLF hiss is found to be downward propagating waves whereas the saucers are upward. It has been shown

that the VLF hiss spectral characteristics and power levels can be explained by amplified Cerenkov radiation due to this auroral electron beam at or below the region where the electrons are accelerated (3000 - 20,000 km).[69] Saucers can be explained by assuming a source of suprathermal electrons of < 5 eV beamed upward. This beam produces amplified Cerenkov radiation at altitudes above 1000 km.[52]

The kilometric radiation, emitted from auroral field lines[1], may also result from the auroral electron beam and current system. A beam driven electromagnetic instability which operates in the presence of ion wave turbulence has been proposed. The instability saturation levels are consistent with the large observed power levels. Waves generated by this mechanism would be lefthand polarized.[79]

c) Two Stream. At the bow shock, solar wind ions and electrons are reflected to produce a counter-streaming (two-stream) distribution. These reflected ions may be the source of the whistler mode noise observed just below the ion gyrofrequency upstream of the bow shock in the solar wind. The noise band just below the electron gyrofrequency is identified with the reflected electrons as is the electrostatic oscillations at the electron plasma frequency.[36] The dynamics of the bow shock is discussed in detail by Greenstadt and Fredricks in this report.

d) Drift Wave. Plasma can be given directed energy in regions where strong temperature or density gradients exist. Ion cyclotron waves are reported in the polar cusp region during a magnetic storm. They may be caused by a current driven ion cyclotron instability or by a density gradient drift instability. Fredricks[33] suggests that the theoretical

work on electrostatic drift waves beyond the plasmapause leads to a source mechanism for Pc5 micropulsations. The Kelvin-Helmholtz instability, resulting from a shear in the ion velocity parallel to the magnetic field lines at the cusp boundaries, may generate micropulsations in the .07 to 30 Hz range.[24] Farley (this report) discusses the role of gradient drift instabilities in the ionosphere.

3.3 Wave Propagation Characteristics

Plasma waves observed in the magnetosphere may have propagated away from the source region. The path of propagation depends on the detailed plasma and magnetic field distributions, on the initial wave normal direction $\underline{k}$, on the wave mode and on the wave frequency.[55] To illustrate the characteristics of plasma wave propagation, ray paths (direction of energy flow) for components of a magnetospherically reflected electron whistler are shown in Figure 5a and for continuum radiation trapped between the plasmapause and the magnetosheath in Figure 5b. From these figures it is seen that waves are not necessarily constrained to follow the magnetic field lines as energetic particles do. (However, whistler-mode waves may be trapped in field-aligned ionization ducts.) Therefore, waves generated by one particle population in one source region can propagate to an entirely different particle population region where they can interact to modify that population.

To obtain the wave characteristics in the generation and wave-particle interaction regions it is desirable to measure the wave characteristics at the point of observation and to trace the rays

forward and backward. The wave characteristics for simple waves have been determined from the amplitudes and phases of the various wave components.[75, 93] These characteristics include the wave normal vector $\underline{k}$ and Poynting vector $\underline{S}$. However, in most cases the wave field is a mixture of many different waves so that it is necessary to obtain a wave distribution function $F(\underline{k}, \underline{S}, \omega)$ to fully describe the waves in analogy to the particle distribution functions. So far the problem has only been formulated but not developed to a useful stage.[103]

3.4 Consequences of Wave-Particle Interactions

Detailed considerations of the magnetospheric plasma wave interactions with the particle populations is given in the papers by Lyons, Greenstadt and Fredricks, Haerendel and Sonnerup in this report and in a number of review articles.[33, 34, 47, 88, 89, 107] Scarf and Russell[89] have developed a list of magnetospheric plasma processes which are known to involve or which probably involve plasma waves. This list is presented here to illustrate the variety and significance of wave-particle interactions:

1. Mechanisms for ring-current decay and for precipitation of electrons and ions to form the diffuse aurora,
2. Energy transfer and heating at the collisionless bow shock and at the field-aligned current regions within the magnetosphere,
3. Mechanisms that provide microscopic coherence and lead to the enormous electromagnetic radiation levels from regions above auroral arcs,

4. Dissipation mechanism in field-line merging regions,
5. Source of anomalous resistivity or instabilities leading to potential double layers which support field-aligned electric fields responsible for the auroral particle acceleration process to form bright auroral arcs,
6. Cause of viscous interactions that support the formation of the geomagnetic tail and the overall plasma convection pattern,
7. Formation of SAR arcs,
8. Cause of electron loss to form the energetic electron slot between the trapped radiation zones, and
9. Scattering of particles from open cusp field lines to closed magnetic field lines.

An example of one wave generation and wave-particle interaction system is illustrated in Figure 6. During geomagnetic storm times a hot ion ring current plasma is injected from the tail into the cold plasmaspheric plasma. Ion cyclotron waves may be generated causing some of the ions to be precipitated into the atmosphere. The ion waves are damped as they propagate downward giving rise to a downward heat flux which may produce a stable auroral red (SAR) arc in the lower ionosphere.

In order to further test the collection of possibilities for wave generation and wave interaction processes, controlled experiments within the magnetosphere are being carried out. A few details about the Siple VLF transmitter[48, 49] and the Echo rocket electron gun and argon ion source[20, 72] experiments are given in Table 4. The scheme for carrying out the Siple transmitter experiment is illustrated in

Figure 7 including the use of satellites such as the Dynamics Explorer pair. Modulated waves in the range of 2 to 16 kHz are transmitted from Siple, Antarctica. These waves enter the ionosphere and propagate along flux tubes between L = 3 and L = 5. If an energetic electron population exists in the equatorial region, the waves may be amplified or may organize the electrons to emit a new wave at a slightly different frequency which changes with time. Wave growth can exceed 30 dB at the rate of 100 dB sec^{-1}. The interacting electrons may be scattered into the loss cone and precipitate into the atmosphere. Satellites along the field line can be used to make careful measurements of the wave characteristics and of the particle distribution functions to provide details of the interaction.

4. DIAGNOSTICS WITH PLASMA WAVES

Observations of naturally occurring plasma waves resulting from processes that are reasonably well understood have been used to obtain information on local plasma parameters and on remotely occurring plasma processes which is complementary to that obtained from other instrumentation. Measurements of ion and electron whistlers and of LHR and UHR noise have provided estimates of plasma density, ion species and temperature. Micropulsations are used as an indicator for changes in the structure of the plasmapause and magnetopause. The source regions for saucers, VLF hiss and kilometric radiation are thought to be associated with the regions of auroral particle acceleration and the presence of kilometric radiation is a good index of bright auroral arc activity. In addition active wave experiments such as performed with the topside sounders, Siple and Echo, have been used to determine plasma parameters and to study wave-particle interactions. More sophisticated active wave experiments are planned.

4.1 Plasma Parameters

Compared to techniques using instruments such as Langmuir probes and mass spectrometers, determination of plasma parameters using wave techniques are less affected by satellite potential and sheath effects since the wave characteristics are established over a volume on the order of wavelength cubed which exceeds 1 km^3. However, one must wait

for the wave phenomenon to occur or must stimulate and receive the appropriate waves.

4.1.1 Ion and Electron Whistlers

Ion cyclotron whistlers are observed to occur in the midlatitude ionosphere. From the frequency of coupling between the electron and ion whistler trace and the observed ion gyrofrequency it is possible to obtain the fractional concentration of the ions. From the shape of the ion whistler trace it is possible to estimate the ion number density and the ion temperature.[3,92] The time of occurrence with respect to the causative lightning discharge and the frequency of the "nose" for the electron nose whistlers observed on the ground at mid- to high latitudes are used to drive the electron density distribution in the equatorial plane. This technique is useful for locating the plasmapause as a function of local time nearly continuously. An example of the plasmapause location for three days in July 1963 is shown in Figure 8.[15]

4.1.2 Magnetohydrodynamic Waves

Recently it has been demonstrated that measurement of the period and identification of the field line for geomagnetic micropulsations (at frequencies below the ion cyclotron frequency, see Table 1) can also give an estimate of the equatorial electron density near the plasmapause. It is found that an external driving force (possibly the solar wind) can cause the field lines to resonate and that this resonance frequency is dependent on the electron density distribution along the field line. The deduced density values are in good agreement with those from simultaneous nose whistler measurements.[112]

4.1.3 LHR and UHR Noise Bands

The lower hybrid resonance frequency depends on the effective ion mass, f_p^- and f_g^-. LHR noise bands are found at altitudes ~ 1000 km in the mid- to high latitudes. With measurements or estimates for f_p^- and f_g^-, the mean ion mass, which depends on the fraction ion concentration, has been derived over a wide latitude range.[64, 92] The non-thermal continuum noise may be generated at the UHR frequency[42]. The observed low frequency cut-off of this noise is identified as the local plasma frequency f_p^-. From measurements of this cutoff with IMP 6

it has been possible to obtain the total electron density in the hot plasma region outside the plasmapause. In fact the comparison of total density from the plasma wave measurements and the suprathermal ion density (88 to 38000 eV) shown in Figure 9 indicate that the plasma is almost entirely suprathermal.[40]

4.1.4 Wave Sounders

By driving an antenna with a sweep-frequency signal and measuring the amplitude and time delay of the received waves it is possible to excite resonances in the local plasma and to observe waves reflected from distant density gradients. From this information, the local density and temperature as well as the density profile toward increasing density can be determined. Such topside sounder experiments have been performed with Alouette and ISIS, for example.[3] Similar experiments are planned for the ISEE Mission and for AMPS Shuttle payloads.[86]

4.2 Plasma Processes

Observations of propagating plasma waves can provide information on plasma processes remote to the point of observation:

4.2.1 Micropulsations

The period, amplitude and polarization of geomagnetic micropulsations have been used to deduce information about the interaction of the solar wind with the magnetosphere, the dimensions of the magnetosphere, and about streaming particles within the magnetosphere. Micropulsations in the Pc1 range (.02-5 sec. periods) are probably caused by cyclotron instabilities outside of the plasmapause. Micropulsations in the Pc2 to Pc5 range (5-600 sec. periods) depend both on particle instabilities

and field line resonance conditions in the vicinity of the plasmapause.[108] For example, the location of the plasmapause boundary can be determined from a chain of micropulsation stations by measuring the polarization and period of Pc3 and Pc4 pulsations. The period increases with the increased plasmasphere extent because the length of the source field line is increased giving a longer resonant period for these waves.[66, 78]

4.2.2 Electron Whistlers

By observing changes in the characteristics of nose whistlers as observed in Antartica and Canada on the time scale of ~ 15 minutes, the cross-L plasma motions near the plasmapause have been inferred. Plasma velocities of ~ 100 m/sec corresponding to electric fields of ~ 0.1 mV/m are obtained.[17]

4.2.3 Saucers and VLF Hiss

Both of these phenomena have a V-shaped frequency-time characteristic as seen in Figure 4 and illustrated in Figure 10a. Ray paths between the source and the satellite that explain this shape are shown in Figure 10b. By modeling the auroral region and calculating ray paths similar to those in Figure 10b it has been possible to determine the approximate saucer source altitude and source region dimensions (10 km vertically by 0.5 km horizontally). The altitude distribution is seen to be above 1000 km and up to the satellite altitude.[52] A similar analysis of many VLF hiss events has not been carried out but one case from Injun 5 indicated a source region near 4000 km altitude.[38] Since these waves are thought to be due to locally accelerated electrons, it is tempting to associate the acceleration region

with the noise source region. Saucers and VLF hiss are indicators of the acceleration process for which the acceleration region must be below or at the saucer source and above or at the hiss source altitude.

4.2.4 Kilometric Radiation

The presence of auroral kilometric radiation on the nightside of the earth has been shown to have a high correlation with bright auroral arc activity[41] (but not with diffuse aurora) as exhibited in Figure 11. Also, the apparent source position of the kilometric radiation as observed by lunar occultations with RAE2[1] is illustrated in Figure 12. These source locations trace out an auroral field line region with an altitude range of 1 to 8 earth radii. Since the source mechanism for kilometric radiation may also involve the auroral accelerated electrons,[71, 79] the presence of kilometric radiation may be used as an indicator of the acceleration region.

4.2.5 Wave-Particle Experiments

The Siple VLF transmitter and Echo particle gun experiments (see Table 4 and Section 3.4) are representative of active experiments which generate waves that carry information about the wave-particle plasma process. For example, the amplitude, growth rate and frequency changes of the waves stimulated by the Siple transmitter contain information about the energetic electron distribution function and bunching along the field line.[48]

4.3 Plasma Waves Associated with Auroral Particle Acceleration

As an example of plasma waves associated with significant plasma processes, we consider the auroral region. Based on the source region determinations for VLF hiss, saucers and kilometric radiations as well as their observed associations with field aligned currents, field aligned fluxes of keV electrons, and bright auroral arcs in the nightside auroral region, these phenomena may all be interpreted as part of the auroral plasma system as depicted in Figure 13. One interpretation is as follows: An emf drives a current along these auroral field lines which closes through the ionosphere. On the field lines for which electrons are driven downward, an electric field parallel to the field line may develop due to one or more processes[8] -- anomalous resistivity, potential double layers, hot/cold plasma mixture or simply due to magnetic mirror forces on the electrons. These parallel electric fields can accelerate a fraction of the low energy electron population to keV energies which are observed below 3000 km and which cause the auroral arcs (see Haerendel, this report). Waves generated as part of the acceleration mechanism or by the electron beam escape upward as kilometric radiation and downward as VLF hiss. The return current ($\sim 10^6$ amps) is carried by large numbers of suprathermal electrons ($\lesssim 5$ eV) on adjacent field lines and this low energy beam generates the upward propagating saucer waves. Further verification of this model may be carried out with the Dynamics Explorer spacecraft -- one at low altitudes at the foot of the field lines and the other moving approximately along a field line through the acceleration/generation regions.

The most significant plasma wave emission from this auroral process is the kilometric radiation. This radiation covers a frequency range of 20 kHz to 2 MHz which escapes the magnetosphere with a maximum at ~ 200 kHz. This frequency of maximum emission is approximately the electron gyrofrequency at the observed source region (see Figure 3a). The radiation is beamed into a solid angle about the field line of ~ 3 steradians with a total power of ~ 10^9 watts which is ~ 10% of the electron beam energy flux and ~ 1% of the energy flux supplied by the solar wind. The noise is emitted as a superposition of short bursts in noise storms lasting up to several hours[41] (see Figure 11). As will be discussed in Section 5.1.2 many characteristics of the kilometric radiation are similar to those of the Jovian decametric radiation.

5. PLASMA WAVES IN OTHER MAGNETOSPHERES AND COSMIC SYSTEMS

Plasma wave instruments have been carried beyond the Earth's magnetosphere into the solar wind in the vicinity of the Earth by many spacecraft and to a distance of 0.3 AU by Helios 1 and 2. However, no plasma wave instruments have been included on either US or Soviet fly-by and lander missions to the planets (Mercury, Venus, Mars and Jupiter). Plasma wave instruments are included in the forthcoming Pioneer Venus and MJS (Jupiter, Saturn and possibly Uranus) Missions and are suggested for the Jupiter Orbiter Mission. At present the nature of plasma waves trapped in the magnetospheres of other planets as well as those is cosmic plasma systems must be inferred from the escaping waves observed by radio astronomy techniques and from plasma parameters that can be deduced from IR, optical, UV, x-ray and γ-ray emissions. For the Jovian magnetosphere, we do have plasma energetic particle and magnetometer measurements performed with Pioneer 10 and 11. Using the detailed interpretations of plasma wave processes in the Earth's magnetosphere along with the existing observational data, and results of laboratory experiments, computer simulations and theoretical developments it is possible to make predictions about plasma wave phenomena in other planetary magnetospheres and cosmic plasma systems.

5.1 Jupiter

Jupiter is an intense radio emitter in two frequency ranges: the decimetric range (100 MHz - 10,000 MHz) due to synchrotron radiation from the trapped relativistic electrons and the decametric/hectometric range (500 kHz - 40 MHz) due in part to the motion of the moon Io through the Jovian magnetosphere probably stimulating emission similar to the terrestrial kilometric radiation. The existence of trapped plasma waves is inferred from those at the Earth and from the observed energetic particle characteristics.

5.1.1 Decimetric Radiation

A radiointerferometer map of the Jovian radiation at 21 cm wavelength is shown in Figure 14. This emission pattern is consistent with this emission being due to synchrotron radiation by relativistic electrons trapped in a dipole magnetic field.[6] Northrop and Birmingham[76] used the electron fluxes and magnetic field directly observed by Pioneer 10 for the first time to calculate the expected synchrotron flux. This result agrees with the observed flux to within a factor of two. Klein[62], however, demonstrates that the synchrotron flux is not constant but varies by ~ 30% over a fifteen year period and the flux variations at two wavelengths are different, as seen in Figure 15.

The variations also do not follow the 10.7 cm solar flux variations (indicator of solar activity) nor the square of Jupiter's distance from the sun suggesting that the relativistic electron flux is not controlled by the solar wind directly but probably is determined by

plasma processes within the magnetosphere itself. Some of these processes may involve plasma waves[90, 110] (see Section 5.1.3).

5.1.2 Decametric Radiation

Ground-based observations of the Jovian decametric (5-40 MHz) radio bursts cover a period of 20 years. Recently satellite-borne receivers have extended the frequency range down to hectometric wavelengths (0.5-5 MHz).[19] The many observational characteristics of this radio noise phenomenon have been discussed by a number of authors.[10, 18, 19, 111] Only a summary of the most prominent features is given here with some interpretation.

Bursts of radio noise from Jupiter are observed over the frequency range of 0.5 to 40 MHz with a peak at ~ 9 MHz, as shown in Figure 16, although each burst may have an instantaneous bandwidth as narrow as 100 kHz. Bursts at different frequencies are not generally correlated. These bursts have durations in the tens of seconds range (L-bursts) and in the millisecond to tenths of seconds range (S-bursts) although burst structure on the microsecond scale has been observed. The L-burst duration is thought to be due to interplanetary scintillations (due to solar wind density irregularities) whereas the S-burst structure may represent the coherence time for the plasma instability or the time for a beam from the source to cross the Earth. Probably the most striking feature is that the probability of observing an intense burst (with a power of $\geq 10^9$ watts) depends jointly on the longitude of Jupiter facing the Earth (Central Meridian Longitude, λ_{III}) and the position of the moon Io in its orbit around Jupiter with respect to the Earth

(φ_{Io}, where 180° places Io between the Earth and Jupiter). Figure 17 is a plot of this probability which indicates the pronounced Io-associated emission for the regions $\varphi \sim 90°$, $\lambda \sim 110°$ (Source B); $\varphi \sim 240°$, $\lambda \sim 240°$ (Source A); and $\varphi \sim 240°$, $\lambda \sim 330°$ (source C). Also a significant, but less probable, Io-independent source at $\lambda \sim 250°$ is apparent.[10] It has been shown that the gross frequency-time character of a noise storm (lasting several hours) is surprisingly reproduceable for a given set of Φ and λ over time periods of ~ 12 years. This reproduceability, the frequency range of emission and the pronounced circular polarization of the bursts (up to 80% righthand polarized above 20 MHz, changing to lefthand below) is used to argue that the observed frequency is related to the electron gyrofrequency in the the source region. Magnetic field models based on the Pioneer 10 and 11 measurements of the Jovian magnetic field are consistent with these deductions[97]: the surface field in the northern hemisphere where the field line through Io intersects the ionosphere is ~ 14 gauss giving a gyrofrequency of ~ 40 MHz and the polarity of the magnetic dipole (opposite to that of the Earth) can explain the polarization sense. Radiointerferometer measurements of the apparent source size indicate an extremely small source region: ~ 400 km if spatially incoherent and ~ 4000 km if spatially coherent (like an antenna) compared to the Jovian radius of 70,000 km. Whatever the emission mechanism, it must be a collective process.

Most effort has been expended in interpreting these observational characteristics for the most intense Io-related bursts. It is generally

accepted that the Io-related emission is generated by Io's motion through the Jovian magnetosphere at a frequency related to the electron gyrofrequency (perhaps a harmonic). This radiation is generated as bursts in small source regions along the field lines connecting Io to Jupiter and is beamed by the source or by the allowed propagation paths (which may include wave-wave coupling) so that in only certain Io-Jupiter-Earth geometries are the bursts observable at earth. Smith[99] has recently reviewed the many models suggested for the Jovian decametric emission in light of the Pioneer 10 and 11 measurements and the possible analogy to the terrestrial kilometric radiation. These models for the coupling of energy from Io's orbital motion to plasma waves include mechanisms by which Io generates large amplitude Alfvén waves or whistlers, accelerates particles along the Io flux tube or sweeps up the existing energetic trapped particles. Smith concludes that at present no comprehensive theory exists but that acceleration of particles by plasma sheaths near Io[95] appears to be a promising coupling mechanism and that most emission properties could be explained by an electron cyclotron instability[71] or parametric instability[79] due to the highly anisotropic distribution of these streaming electrons. Also Scarf[87] has suggested that the emissions could be analogous to the $3/2\ f_g^-$ emissions seen in the Earth's magnetosphere. Many of the fine scale features may be attributable to the propagation paths determined by the structure of the multipole magnetic field and the plasma distribution in the magnetosphere.

A pictorial representation of the basic Io sheath-acceleration model[94] is shown in Figure 18. Because of rapid rotation of Jupiter's

magnetosphere (~ 10 hour period past Io) and the proximity of Io to Jupiter (6 R_J) it may be that a motional emf of up to 570 kV is developed across Io's ionosphere. (An ionosphere on Io was detected with the Pioneer 10 radio occultation experiment). This emf is thought to drive a current down the field lines connecting Io to the Jovian ionosphere and back up to Io for closure in the Io ionosphere. Plasma sheaths may form at or near Io (perhaps similar to the double layers described by Block and Fälthammar[8]) to separate the plasma moving with Io from that moving with the Jovian magnetic field. A significant fraction of the motional emf may be dropped across these sheaths leading to the acceleration of electrons and ions to ~ 100 keV energies. The streaming electrons carry up to 10^{13} watts of power compared to 10^{9} watts for a large decametric burst. This model seems plausible based on the Pioneer 10 and 11 observations of order of magnitude increases in the flux of 100-500 keV electrons only on the field lines associated with Io. On comparing the models depicted in Figure 18 and Figure 13 and the observed characteristics of the Jovian decametric and terrestrial kilometric emissions, the similarities are striking. These similarities are the basis for suggesting that the mechanisms may be the same and that perhaps this process is common in plasma systems.[57, 60, 71, 95]

5.1.3 Trapped Plasma Waves

No direct measurements of trapped plasma waves have yet been made in the Jovian magnetosphere. However, the measurements of Pioneer 10 and 11 revealed sources of free energy in the plasma and energetic

particles which must lead to the generation of a variety of plasma waves similar to those at the Earth. The frequency range for these waves would be scaled by the relative plasma densities and magnetic field magnitude. Some examples of likely types of plasma waves are mentioned here.

Greenstadt and Fredricks (in this report) have discussed the bow shock structure expected for Jupiter and the associated bow shock waves. If a current system of high energy electrons is driven by Io as suggested in Figure 18 then it is reasonable to expect whistler mode noise similar to VLF hiss and saucers from instabilities or amplification along the field lines in analogy to Figure 13. Sentman et al.[90] report significant pitch angle anisotropies of the relativisitic electrons (> 20 MeV) which are apprently recirculating through the Jovian magnetosphere. They have assumed that whistler mode turbulence maintains this anisotropy in the range of 3 to 5 R_J and deduced the required cold plasma density in good agreement with the values derived from Pioneer 10 plasma measurements.[110] Barbosa and Coroniti[5] considered this same instability and arrived at upper limits to the stably trapped flux in agreement with observed values. Also they estimated the lower limit amplitude for this whistler mode noise (probably like ELF hiss) of $\sim 0.5\ m\gamma\ L^{-1/2}$ where $L \sim 4\ R_J$. Another estimate for this noise amplitude based on particle data from Pioneers 10 and 11 gives $\sim 2\ m\gamma$ for a typical amplitude [Sentman and Goertz, private communication, 1977]. Maximum values at the Earth reach 30 mγ.[106] Outside of Io's orbit where the cold plasma density decreases below 10 cm^{-3} it is very likely that electrostatic waves are generated at the upper hybrid resonance

frequency and near odd-half harmonics of the electron gyrofrequency ($3/2\ f_g^-$, $5/2\ f_g^-$, ...) according to Ashour-Abdalla and Kennel.[4] At the Earth, the corresponding waves may cause particle precipitation leading to diffuse aurora. Also, the non-thermal continuum radiation may result from coupling of this electrostatic noise to an electromagnetic mode. The escaping component of such a process is a likely explanation for the non-Io-associated decametric radiation. Results from the plasma wave instrument on MJS are anxiously awaited to test these ideas and to identify other significant plasma wave processes in the Jovian magnetosphere.

5.2 Other Planets

Direct spacecraft measurements indicate that Venus and Mars do not have a magnetic dipole moment large enough to produce a magnetosphere. Siscoe (in this report) considers the nature of the interaction of the solar wind with these bodies. Although plasma waves may be generated as part of the interaction processes, measurements have not yet been made and the problem has been given little attention theoretically. Arguments that Saturn, Uranus and Neptune may have a magnetic moment significant enough to produce an extensive magnetosphere are reviewed by Siscoe (in this report) and by Van Allen.[110] These arguments are based primarily on similarity of the planets to Jupiter in terms of rapid rotation and size and on the detection of hectometric bursts from Saturn and possibly Uranus. So far synchrotron emission has not been detected from these outer planets. Both magnetic field and energetic particle measurements at Saturn are forthcoming from Pioneer 11

on 1 September 1979 and MJS and possibly at Uranus with MJS. At present descriptions of plasma wave processes at these major planets are entirely based on analogies to the Earth assuming an extensive magnetosphere with trapped plasma and energetic particle populations.

5.2.1 Synchrotron Radiation

In the case of Jupiter, detection and mapping of the decimetric synchrotron radiation provided definitive evidence for an extensive magnetosphere with a trapped energetic electron (> 10 MeV) population. Detections of synchrotron radiation from Saturn, Uranus or Neptune have not been reported. These measurements are difficult because the expected lower energy electrons and weaker magnetic fields imply that the emission frequencies would fall in the 10^2 to 10^3 MHz range where the galactic background noise level is high and the radiotelescope resolution is poor. Models for synchrotron emission from Saturn have been studied by Luthey[68] and for Uranus and Neptune by Kavanagh[58]. An example of the situation for Uranus is shown in Figure 19. Measurements above 1000 MHz trace out the 180°K black-body atmospheric emission curve whereas the synchrotron emission is expected to be significant below 1000 MHz. The 34 MHz presently achievable upper limit point is almost an order of magnitude above the model synchrotron flux levels.[96] Both Saturn and Uranus are uniquely different from Jupiter in characteristics which may affect the emission of detectable synchrotron radiation: Saturn has a system of particulate rings which may preclude the possibility of intense radiation belts close to the planet and Uranus may have a magnetic axis nearly aligned with its rotational axis which makes

an angle of 98° with the ecliptic plane normal. Since the spin axis of Uranus is currently pointing at ~ 40° to the direction of the sun and the synchrotron emission would be beamed in the equatorial plane (perpendicular to the magnetic axis), if emitted, it may not be beamed toward the Earth. Improved radio-astronomical techniques or measurements closer to the planets themselves are required to possibly detect this radiation.

5.2.2 Hectometric Radiation

Using the spinning antenna on IMP 6, sporadic bursts of hectometric radiation attributable to Saturn[57] and possibly to Uranus[11] have been reported. Typical spectra for these bursts are shown in Figure 16 along with those of the terrestrial kilometric and the Jovian decametric emissions. By assuming that the Saturn and Uranus peak emission frequencies (1.1 and 0.5 MHz respectively) scale as the relative magnetic field strengths (as they do for the Earth and Jupiter) an estimate of the planetary surface magnetic field can be obtained. These values are compared in Table 5A to values predicted from a magnetic Bode's law[60, 110] with reasonable agreement. Kennel and Maggs[60] have also used the observed power flux at Earth to compute the total power emitted assuming that it is beamed into a hemisphere (2π steradians). This power is then compared to an estimate of the power intercepted from the solar wind by the planetary magnetosphere and in Table 5B the ratio is calculated. Included, also, are the recent data for Uranus.[11] Calculated in this manner, the power ratios (conversion efficiency) increase from 0.1% at the Earth to 25% at Uranus! Following this trend, Neptune would

be predicted to emit nearly 100% of the power received from the solar wind which would produce a power flux of 10^{-20} Wm^{-2} Hz^{-1} at the Earth. However, the radiation may not be beamed into 2π steradians as assumed by Kennel and Maggs. For the Earth the emission solid angle is ~ 2 steradians and for Jupiter it is ~ 10^{-2} steradians leading to lower estimates of the emitted power as indicated in Table 5B. Using the Jovian efficiency of .02%, Neptune would produce a power flux of 2 x 10^{-24} Wm^{-2} Hz^{-1} at Earth. Of course, it may be that the power source comes from the rotational energy loss of the planets and not from the solar wind.

Detection of these hectometric bursts is significant in indicating the presence of magnetic fields nearly in agreement with the predicted values scaled from Jupiter and the existence of significant free energy from the plasma or particle populations. Because of the nature of the bursts it is tempting to conclude that the emission process is the same as for the Earth and Jupiter but the mechanism for coupling energy into this process is unresolved -- is it Earth-like coming from the solar wind flow past the magnetosphere or Jupiter-like coming from the planetary rotational energy via a close-in moon? Both Saturn and Uranus do have moons of appreciable size relatively close-in. Hopefully, measurements with MJS will resolve this question.

5.3 Cosmic Systems

Radio emissions from the solar chromosphere and corona and from cosmic plasma systems beyond the solar system are the result of plasma wave processes with the characteristic plasma frequencies scaled up

for the plasma number density and the magnetic field magnitude. Interpretation of these processes must depend on the properties of processes studied in space and the laboratory with computer simulations and theories taking into account the different parameter regimes and relativistic effects. Three plasma processes, for example, have escaping plasma waves with characteristics that are somewhat similar to the escaping waves from the planetary magnetospheres. These wave phenomena are solar radio bursts, flare star radio outbursts and pulsar radio emissions.

5.3.1 Solar Radio Bursts

The energy for most solar radio emissions that are non-thermal (not due to thermal Bremsstrahlung) is thought to come from energetic electrons associated with solar flares. The most intense bursts and noise storms occur after some large optical flares in which electrons are found to be accelerated to > 10 MeV energies. One suggested acceleration mechanism by Carlqvist[14] is that of a double layer which is formed on interruption of the current flowing along the magnetic field lines linking sunspot regions on the solar surface. Stored magnetic energy is released across this potential region which accelerates particles to high energies causing electromagnetic emissions from x-ray to radio wavelengths, and particle as well as shock wave injections into the solar corona. In this model the double layer is thought to be similar to that possibly existing in the terrestrial auroral regions and in the vicinity of Io. Radio noise is produced by a variety of processes ranging from gyrosynchrotron emission from the highest energy

electrons in the chromosphere for the microwave bursts to plasma oscillations driven by the outward streaming electrons in the corona for type III bursts.[63]

5.3.2 Flare Star Radio Outbursts

A class of relatively close stars (dMe stars, dwarf, M-type with emission spectra) have been observed to include members which exhibit increases in optical brightness, especially in the UV, lasting for minutes. A few of these optical flare stars also have detectable radio outbursts which are associated with the optical flare within ± 5 minutes.[101] An example from Wolf 424 is shown in Figure 20. These outbursts last from tens of seconds to tens of minutes and are highly polarized (60%) with both circular and linear components. In morphology, these outbursts are similar to solar microwave and type III bursts but the total energy released in an outburst is over two orders of magnitude greater than for a solar radio burst from a star that is about half the solar diameter. Some features of these flare stars may help explain the radio outburst characteristics: they are rapid rotators (several day period), they have large surface magnetic fields of $\sim$ 20,000 gauss and many are part of binary star systems. A coherent synchrotron emission mechanism is likely. The short duration of the emission is suggestive of a rapid electron acceleration process similar to that for solar flares (with 10^2 more energy released) and/or to that for planetary particle acceleration associated with the radio bursts. A double layer acceleration process could result from a current driven in the stellar magnetosphere by the rapid rotation or by the rotation with respect to the companion star.

5.3.3 Pulsar Emission

Pulsars are generally thought to be rapidly rotating (periods .03 to 10 seconds) neutron stars (10 km radius) with surface magnetic fields as high as 10^{12} gauss. In a model for pulsar emissions due to Sturrock[104] field-aligned currents are driven along polar field lines due to the interaction of the rapidly rotating pulsar magnetosphere with the surrounding nebula. This current is carried by beams of electrons which have been accelerated to 10^{14} eV energies by a plasma sheath at the star surface. These electrons emit γ-radiation which annihilates to produce an electron-positron plasma. The resulting high energy plasma may lead to a two-stream instability which breaks the current into current sheets. These current sheets may then radiate in the radio spectral range like single particles, with a large effective charge, moving along curved magnetic field lines. Observation of the radiation in the form of complex pulses depends on the relative orientation of the rotation and magnetic axes with respect to the line of sight. This model, shown in Figure 21 (from Kennel[59]) is also capable of explaining the γ- and x-ray bursts observed from the Crab pulsar (CP 1919). Of interest for this pulsar emission model are the elements of sheath particle acceleration and current driven plasma wave instabilities which may be common to solar system processes as well.

6. RECOMMENDATIONS FOR CONTINUED PLASMA WAVE RESEARCH

In the previous sections the characteristics of observed plasma wave processes and the current interpretations regarding these processes have been reviewed. The wide variety of plasma wave types in the Earth's magnetosphere was emphasized since our information is most complete on these, but, known characteristics and suggested mechanisms for planetary and cosmic escaping plasma waves were also discussed. In this final section we summarize the major advances in plasma wave research during the past several years, identify the most significant problem areas to be emphasized and make recommendations for the future direction of this research.

6.1 Recent Major Advances

During the past several years, major advances in plasma wave research included the following:

6.1.1 Development of a theory which describes the dynamics of the outer electron radiation zone in terms of diffusion rates, amplitude of ELF hiss, cold plasma density and precipitation fluxes.

6.1.2 Observations of electrostatic noise at the bow shock, magnetosheath, cusp, tail, plasmapause, auroral field lines and the ionosphere -- all regions of significant wave-particle interactions related to many of the fundamental plasma processes listed in the next section.

6.1.3 Discovery of the terrestrial kilometric radiation and detailed measurements on location, beaming, and association with bright auroral arcs.

6.1.4 Description of VLF hiss and saucers as coherent beam amplification processes driven by the intense auroral field-aligned currents.

6.1.5 Discovery that power system harmonics of 60 Hz can leak into the magnetosphere and cause electron precipitation and that the active study of similar wave-particle and wave-wave interaction using the Siple VLF transmitter and electron beams injected from rockets can be carried out.

6.1.6 Detection of hectometric radio bursts from Saturn and Uranus and realization that these bursts may result from processes similar to the Jovian decametric and terrestrial kilometric bursts which may also be similar to solar, flare star and pulsar emission mechanisms.

6.2 Significant Research Problems

In my opinion research concerning the following plasma processes should be emphasized during the next decade. Examples of unanswered questions on the involvement of plasma waves in these processes are given:

6.2.1 Energy transfer from the solar wind to the magnetosphere and heating at the bow shock.

Greenstadt and Fredricks (this report) emphasize the role of plasma waves in the collisionless bow shock structure. It seems that

ion-acoustic waves are responsible for thermalizing the directed solar wind electron flow at the shock but the waves responsible for the ion heating have not been identified. It is suggested that electrostatic waves near the lower hybrid resonance frequency produced by a counterstreaming ion distribution may be responsible. Measurements of waves and particles in the bow shock region must be done by mother-daughter spacecraft pairs (like ISEE) to account for the rapid spatial variations.

6.2.2 Precipitation of electrons and ions from the trapped radiation regions to cause diffuse aurora.

Lyons (this report) thoroughly discusses the dynamics of the radiation belts. Electron diffusion and precipitation can be understood in terms of ELF hiss and to some extent chorus. Electrostatic electron cyclotron noise (3/2, 5/2 f_g^-) is thought to be important also but its spatial distribution and conditions for generation are not yet well enough documented, along with adequate particle distribution function measurements, to assess its significance. Ion cyclotron noise has long been thought to control ion precipitation but as yet it has not been conclusively detected in the interaction region. More specialized instrumentation may be required for future missions.

6.2.3 Acceleration of electrons along auroral field lines causing bright auroral arcs and intense kilometric radio emissions.

The evidence for field-aligned currents and field-aligned electric field regions is discussed by Haerendel (this report). Many mechanisms have been proposed to account for the intense beams of electrons

which carry most of this current at low altitudes and which cause bright auroral arcs. However, so far it has not been possible to identify the responsible mechanism because observations are required in the magnetosphere and because the various possibilities need more study through theoretical analysis, computer plasma simulation and laboratory plasma experiments. If anomalous resistivity is the cause then it is thought to be electrostatic ion cyclotron or ion acoustic plasma waves which cause the resistivity. If it is double layers then waves may be important to the formation but not the persistence of these layers and if it is a magnetic mirror effect then plasma waves may play only a secondary role.

A good case has been made that the VLF hiss is Cerenkov emission amplified by the auroral electron beam and that the saucer emissions are due to the same process for a return current lower energy electron beam which has not yet been observed. Several completely different mechanisms (amplified electron cyclotron and parametric wave decay for example) have been suggested as the cause for the high power kilometric radiation. Measurements of the wave characteristics (especially polarization), plasma parameters and energetic particle distribution function in or near the source region are not yet available with which to identify the specific plasma wave process.

6.2.4 Energy dissipation in the magnetotail allowing field line merging which determines the magnetosphere shape and the plasma convection patterns.

For the process of field line merging between the solar wind and magnetosphere magnetic fields at the bow shock and along the magnetosheath

and for the process of field line reconnection in the tail region (see Sonnerup, this report), an energy dissipation process is required. One likely suggestion is that energy is dissipated through the creation of electrostatic waves. Both broadband electrostatic noise and electrostatic electron cyclotron emissions are observed in the regions where field line merging should be taking place. The detailed associations between field-aligned currents, changes in the magnetic field vectors, particle energization and the amplitudes of these waves have not yet been made to determine if the waves can provide the required localized resistivity. ISEE should provide some of these measurements.

6.2.5 Energy transfer between the magnetosphere and atmosphere and heating in the ionosphere.

Energy is transferred directly from the magnetosphere through the ionosphere into the atmosphere through precipitating particles, field-aligned and transverse current systems and by electric fields. Banks (this report) discusses a number of these coupling mechanisms. As with many plasma wave processes it is not known whether waves are instrumental in the particle acceleration, current resistivity or electric field penetration mechanisms or whether they are just by-products. Many of these questions should be answered by the continued Atmospheric Explorer and the Dynamics Explorer Missions. Sturrock (this volume) has mentioned several processes by which this energy coupling could affect the atmospheric weather systems.

6.2.6 Energization, diffusion and precipitation of energetic particles in the magnetospheres of other planets.

Some evidence for wave-particle interactions in the Jovian magnetosphere was presented in Sections 5.1.1 and 5.1.3 and other characteristic are discussed by Lyons, Haerendel and by Kennel and Coroniti (this report). There is no doubt that trapped plasma waves exist in the Jovian magnetosphere with much the same variety as for the Earth. However, because of differences in the plasma distribution, density and temperature (~ 100 eV) and because the particles are mildly relativisitic, different wave processes may dominate. MJS'77 will give the first clues on plasma waves and the wave-particle interactions at Jupiter, Saturn and possibly Uranus. However, missions such as the Jupiter Orbiter are required to obtain any detailed understanding of the plasma wave processes.

6.2.7 Emissions of the intense radio bursts from Jupiter, Saturn and Uranus.

As indicated in Section 5, it is tempting to conclude that the mechanisms for planetary radio emissions are similar because they seem to scale with the supposed magnetic field in frequency and with the solar wind energy input in intensity. But, in fact, the details for the emission process of terrestrial kilometric radiation are unknown. Direct measurements with MJS(U) of beaming, polarization and intensity are definitely required along with measurements of any associated trapped plasma waves similar to saucers and VLF hiss, broadband electrostatic noise and magnetic noise bursts as well as evidence for electron

acceleration, electron beams, and field-aligned currents. It is desirable to determine if the current systems are driven by solar wind convection or by planetary rotation (with respect to a moon?). For example, if one argues that the Jupiter-Io situation is unique and that each planet should have a solar wind current system which causes radiation in addition, as at the Earth, then Jupiter may have a night-time decametric source of 10^{11}-10^{13} watts which could explain the observed non-Io associated emissions by leakage to the dayside.

6.2.8 Cascading and transport of energy within the solar wind.

Observations of solar wind density, velocity and temperature parameters are reported by Feldman (this report). Observations of waves within the solar wind have been reported from the early Pioneers and IMP's near the Earth (1 AU). Pioneer 10 and 11 have extended measurements in the magnetometer frequency range to beyond 7 AU and Helios 1 and 2 to 0.3 AU for a wide frequency range. Observed waves cover the expected spectrum (as in Figure 2) from long-period Alfvén waves (like some micropulsations and ion cyclotron waves) to waves at twice the electron plasma frequency generated by electron beams from solar flares. However, it remains to identify the role of these waves in the transport of energy from the solar atmosphere through the chromosphere and corona and into the distant solar wind. For the solar wind it must be determined if the Alfvén waves, for example, carry most of the energy from the photosphere and are dissipated to heat the corona and if the fluctuations in the plasma density (which lead to interplanetary scintillations of radio sources) are due to a wave-wave interaction process. Missions such as Helios and the proposed Out-of-Ecliptic are required

to determine the solar-solar wind energy balance and the significance of plasma wave processes to this balance.

6.2.9 Emission of plasma waves from cosmic plasma systems such as the solar atmosphere, flare stars, pulsars, galaxies and quasars.

Examples given in Section 5.4 suggested that some cosmic plasma wave processes may be interpreted in terms of processes that can be well studied by passive and active experiments in the magnetospheres of the Earth and, in the future, Jupiter. Once the local process is explained in terms of concrete physical principles, extensions in the theory, in laboratory experiments and in computer simulations can be made to better describe the particular cosmic plasma system.

6.3 Future Direction

I feel that future plasma wave research should emphasize the following:

6.3.1 Completion of survey for plasma and plasma wave processes throughout the solar system

Within the Earth's magnetosphere, the auroral and polar regions between 3000 km and 4 R_E have not been surveyed. Auroral particle acceleration, kilometric wave emission and plasma convection processes occur in these regions. The proposed Dynamics Explorer Mission supplemented by ISEE and GEOS would provide much-needed measurements. Pioneer 11, Pioneer Venus, MJS(U) will provide the necessary survey data for Venus, Jupiter, Saturn and possibly Uranus. However, a flyby of Neptune and Pluto, additional investigations at Mercury and orbiters

of Jupiter (as proposed), Saturn and Uranus are necessary to provide the detailed data on magnetospheric processes and their response to solar and solar wind variations. Data on plasma and plasma wave processes in the solar wind into 0.3 AU in the ecliptic are being collected by Helios 1 and 2. However, the solar heliosphere processes both closer to the sun and out of the ecliptic plane must be at least surveyed in order to develop theories for the dynamics of stellar magnetospheres.

6.3.2 More detailed measurements of electromagnetic and electrostatic wave characteristics using multiple receivers and correlation of wave measurements with more detailed energetic particle and plasma measurements in the Earth's magnetosphere.

A new level of plasma wave instrumentation and analysis technique needs to be developed to obtain more definitive plasma wave measurements. Storey and Lefeuvre[103] are developing the mathematical and data processing techniques to derive wave distribution functions $f(\underline{k}, \underline{S}, \omega)$ from measurements of the six electromagnetic wave components. Appropriate plasma wave instrumentation is to be flown on GEOS but these same measurements are also required elsewhere in the magnetosphere. Similar instrumentation and analysis methods need to be developed for determining the complete wave characteristics of electrostatic waves. GEOS will also carry an on-board correlator to produce spectrograms of a selectable frequency-time segment. This processor provides broadband data at a much lower data rate than would be required to transmit the analog data directly -- thus conserving telemetry as is required for

obtaining detailed spectral information from the solar wind and from planetary magnetospheres. Satellite-to-satellite interferometry on plasma waves will be attempted for the first time between ISEE-A and ISEE-B. This technique may be particularly valuable for determining the spatial correlations of trapped plasma waves and the source locations (and source sizes) for escaping plasma waves such as the kilometric radiation. This more sophisticated instrumentation must be accompanied by similarly sophisticated instruments for measuring the plasma parameters and particle distribution functions -- generally on pairs of spacecraft to be able to separate spatial from temporal variations.

6.3.3 Development of active plasma experiments for further probing the magnetosphere under controlled conditions and for carrying out basic wave-particle experiments in the extensive magnetospheric plasma.

Helliwell[48] reviews the techniques and results of VLF wave injection into the magnetosphere in order to study wave stimulation and particle precipitation processes (see also Figure 7). The results to date suggest that more sophisticated experiments should be carried out in conjunction with in situ spacecraft that are capable of measuring the local wavefields -- original and stimulated -- and the detailed evolution of the particle distribution functions. Such experiments are planned in conjunction with ISEE and with the proposed DE (which is more suitable) but better experiments would include the VLF transmitter also on a spacecraft. The injected electron and ion beam experiments with the Echo rockets are examples of using the magnetosphere as a

laboratory plasma in which to produce controlled wave-particle processes (see Cartwright and Kellogg[20]) as well as to study the magnetosphere itself. A concentrated effort should be made to take advantage of the Shuttle launch capability to expand these laboratory experiments along the lines reviewed by Scarf[86] for the Plasma Physics and Environmental Perturbation Laboratory (PPEPL) and the proposals for a series of Atmospheric, Magnetospheric and Plasma Physics (AMPS) payloads. Already it seems that Spacelab I will carry an electron gun with diagnostic instrumentation for waves, particles and optical emissions. However, it must be remembered that a laboratory carried by Shuttle can examine a limited plasma parameter range so that active experiments to the outer magnetosphere are also required. These can be launched from Shuttle or by conventional rocketry from Earth.

6.3.4 Performance of laboratory plasma experiments, computer simulations, and theoretical analyses that treat specific processes identified in the magnetosphere.

More emphasis must be placed on research efforts that complement the direct space measurements especially since the space experiments are being designed to obtain more quantitative results on some well identified processes. Fälthammar[26] emphasizes the increasing need for laboratory plasma experiments. He suggests that these fall into two categories: (1) configuration simulation in which the geometrical properties of a large system such as the magnetosphere or moon (terrella and lunella experiments) are modeled and (2) process simulation in which the local behavior of a real plasma exhibiting a

particular process of interest is studied. Some experiments relating to current flow and plasma convection in a dipole field (Birkeland currents); energy dissipation at a magnetic neutral line or point (neutral sheets and solar flares); anomalous resistivity, electric potentials in magnetic mirrors and electric double layers (particle acceleration in neutral sheet and on auroral field lines); and the penetration of plasma into a magnetic field region (bow shock) have been carried out.

A few experiments related specifically to space plasma wave processes have been performed: An experiment by Bernstein, et al., with a large scale electron beam produced emissions at 3/2, 5/2, ... harmonics of the electron gyrofrequency and at the electron plasma frequency[7]. Other experiments, for example, designed to study large magnetic-field-aligned electric potentials and particle acceleration, produce ion acoustic turbulence and emissions below the electron gyrofrequency and above the electron plasma frequency which are similar to the plasma waves associated with the auroral acceleration regions. Further research is required to ascertain, quantitatively, the role that these waves play [S. Torvén, personal communication].

Computer plasma simulations provide a powerful technique to study specific plasma phenomena in comparison with space measurements, laboratory results and theoretical predictions. Such computer plasmas can develop smoothly into non-linear states. Since one has information on the position of each particle and on the applied microscopic fields, diagnostic quantities can be computed to give the evoluation of the

plasma parameters, wave modes and particle distribution functions. Very few applications of this developing technique have been made specifically to space plasma problems. Goertz and Joyce[35] did investigate the formation of an electrostatic double layer in one dimension, for example, which confirmed a number of the theoretical precictions but the diagnostics to identify the associated electrostatic plasma waves has not yet been carried out. The computer simulation technique is an important key to interpreting observed processes and to extrapolating these processes into other plasma regimes.

The significant research problems identified in Section 6.2 provide a working list of problems to be attacked by theorists. "Solutions" to these problems require that the theorist be provided with suitable results from space measurements, laboratory experiments and computer simulations. In order to obtain suitable data, the theorist needs to be involved in defining the space missions, in identifying the specific problems to be attacked, and in specifying the parameters to be measured. They also need to identify the computer and laboratory experiments to be carried out and the required parameters to be determined.

6.3.5 Identification and interpretation of plasma wave processes from cosmic plasma systems seeking analogies and contrasts to solar system plasma processes.

Plasma and plasma wave processes in other cosmic systems may operate from different sources of free energy and under different plasma conditions. Examination of these processes (from the escaping plasma

wave information) is important in order to test the interpretation of solar system processes by extrapolation and to identify new processes which occur within the Universe. Such research provides the necessary cross-fertilization between solar system plasma physicists and astrophysicists.

7. CONCLUSION

It is only in the neighborhood of the Earth that one can hope to ever thoroughly investigate a large number of plasma phenomena, within reasonable financial resources, to the point where the physical mechanisms are really understood. Once that is done other planetary and cosmic plasmas can be understood much more easily (by analogy with the terrestrial magnetosphere) with limited information by remote sensing and perhaps a few in situ probe measurements.

Table 1. ELECTROMAGNETIC PLASMA WAVES

PHENOMENON	LOCATION	OBSERVED FREQUENCY	CHARACTERISTIC FREQUENCY	REFERENCES	WAVE PROPERTIES	SOURCE MECHANISMS
Geomagnetic Micropulsations and Magneto-hydrodynamic Waves	Magnetosheath Polar Cusp Plasmapause	.001 Hz-10 Hz	$< f_g^+$	9, 51, 66, 70, 78, 108, 112	Transverse and compressional Alfvén waves transmitted along geomagnetic field lines. Excited at times of geomagnetic storms. Indicators of changes in magnetosphere configuration.	Kelvin-Helmholtz or drift instability in magnetosheath and cusp. Energetic ion loss cone and gyroresonance instability at plasmapause. May also result from coupling of magnetohydrodynamic waves from solar wind.
Ion Cyclotron Waves	Plasmapause Polar Cusp	.5 Hz-100 Hz	$\leq f_g^+$	23, 24, 30, 56, 70	Spectrum of waves just below the ion cyclotron frequency. Generally L.H. polarized. Causes proton precipitation?	Growth from hot ions injected into region of low density cold plasma by drift mode or ion cyclotron resonance instabilities.
Bow Shock Plasma Waves	Bow Shock Magnetosheath	20 Hz-200 Hz	$< f_g^-$	36, 80, 81	Electromagnetic turbulence comprised of bursts of whistler mode waves.	Coupling of electrostatic turbulence in ion heating region in presence of of large temperature gradients.
Lion's Roar	Magnetosheath	90 Hz-160 Hz	$< f_g^-$	98	Packets of whistler-mode waves propagating along magnetic field. Associated with magnetic storms.	Ion cyclotron overstability from streaming 10 keV protons.
Magnetic Noise Bursts	Auroral Field Lines Neutral Sheath	10 Hz-600 Hz Peak ~ 100 Hz	$< f_g^-$	45, 46	Superposition of intense bursts of magnetic noise. Associated with tail electrostatic noise and similar to lion's roar.	Current driven instability by streaming protons producing noise with $\hat{k} \parallel \hat{B}_o$.
Ion Cyclotron Whistlers	Ionosphere Plasmasphere	10 Hz-750 Hz	$\leq f_g^+$	3, 37, 91, 92	L.H. polarized tones rising to the ion cyclotron frequency. Provides information on ion densities and temperatures.	Terrestrial lightning discharge; wave couples from R.H. to L.H. polarized in ionosphere.

Table 1. (Cont'd.)

PHENOMENON	LOCATION	OBSERVED FREQUENCY	CHARACTERISTIC FREQUENCY	REFERENCES	WAVE PROPERTIES	SOURCE MECHANISMS
ELF Hiss (Plasmaspheric)	Plasmapause Plasmasphere Detached Plasma Regions	10 Hz-5 kHz	$< f_g^-$	22, 33, 106	Whistler-mode turbulence at all local times. Fills plasmasphere. Causes electron precipitation from outer radiation zone.	Electron cyclotron instability with 10 keV electrons at equator.
Chorus	Outside Plasmasphere	10 Hz-5 kHz	$< f_g^-$	13, 82, 109	Series of overlapping rising or falling tones. Quasi-periodic. Predominantly dayside equator.	Doppler shifted cyclotron resonance with 5-150 keV electrons. Causes precipitation of > 20 keV electrons.
VLF Hiss (Auroral and V-shaped)	Auroral Field Lines	10 Hz-100 kHz	f_{LHR}-f_g^-	38, 65, 69	V-shaped noise bands on frequency-time, often superimposed. Downgoing waves < 2500 km. Associated with electrostatic turbulence.	Convective beam amplification of Cerenkov noise by 100 eV-10 keV electrons. Coupled from electrostatic noise?
Saucers	Auroral Fields Lines	500 Hz-30 kHz	f_{LHR}-f_g^-	38, 52	V-shaped noise bands in frequency-time, often superimposed. Upgoing waves > 1000 km.	Convective beam amplification of Cerenkov noise by < 5 eV electrons.
Discrete Emissions	Plasmapause	1 kHz-12 kHz	$< f_g^-$	47, 61, 73	Rising, Falling or mixed tones, sporadic or quasi-periodic. Rarely seen with spacecraft.	Electron cyclotron resonance with bunches of > 10 keV electrons.
LHR Noise	Auroral Zone Plasmapause	4 kHz-18 kHz	$\geq f_{LHR}$	21, 64, 105	Intense noise band above LHR frequency, nearly electrostatic with $\hat{k} \perp \hat{B}_o$.	Cerenkov radiation from 10 eV - 10 keV electrons. Enhanced electron whistler noise.
Upstream Whistler Mode Waves	Up-stream of Bow Shock in Solar Wind	.01 Hz-4 Hz 20 kHz-70 kHz	$< f_g^+$ $< f_g^-$	25, 28, 32 36	Two bands: below ion cyclotron and below electron cyclotron frequency.	Two-stream instability from (reflected ?) ions and electrons streaming away from bow shock.
Trapped Non-Thermal Continuum Radiation	Outside Plasmapause Inside Magnetosheath	500 Hz-20 kHz	$> f_p^-$ (local) $< f_p^-$ (solar wind)	39, 42, 54	Weak broadband noise trapped between plasmapause and magnetosheath. Generated on dayside 0400-1400 LT.	Gyrosynchrotron radiation from outer radiation belt electrons or coupled from electrostatic noise at plasmapause.

Table 1. (Cont'd.)

PHENOMENON	LOCATION	OBSERVED FREQUENCY	CHARACTERISTIC FREQUENCY	REFERENCES	WAVE PROPERTIES	SOURCE MECHANISMS
UHR Noise	Plasmapause	100-600 kHz	$\sim f_{UHR}$	74, 105	Intense noise band near upper hybrid resonance frequency.	Incoherent Cerenkov radiation from 10 eV-10 keV electrons.
Electron Whistlers	Plasmasphere Plasmapause	100 Hz-1 MHz	$< f_g^-$	3, 47, 83, 100	R.H. polarized falling tones. Travel along magnetic field lines in ionization ducts or can be refracted across field lines.	Terrestrial lightning discharge. Nose whistlers provide electron density inside and outside of plasmapause.

f_g^+ = ion gyrofrequency

f_g^- = electron gyrofrequency

f_p^+ = ion plasma frequency

f_p^- = electron plasma frequency

f_{LHR} = lower hybrid resonance frequency (LHR)

f_{UHR} = upper hybrid resonance frequency (UHR)

L.H. and R.H. = left- and right-hand polarization

$\hat{k}$ = wave normal vector direction

$\hat{B}_o$ = geomagnetic field direction

AKR = auroral kilometric radiation

Table 2. ELECTROSTATIC PLASMA WAVES

PHENOMENON	LOCATION	OBSERVED FREQUENCY	CHARACTERISTIC FREQUENCY	REFERENCES	WAVE PROPERTIES	SOURCE MECHANISMS
Tail Broadband Electrostatic Noise	Magnetotail at boundaries of Plasma Sheet	10 Hz-2 kHz	f_{LHR}-f_g^-	45, 85	Broadband emission consisting of discrete bursts with V-shaped structure, $\hat{k} \perp \hat{B}_o$. Associated with VLF hiss in same region and with AKR.	Associated with streaming protons (~ keV). Ion cyclotron or ion acoustic current driven instability.
Auroral Field Lines Turbulence	Auroral Field Lines All Local Times	10 Hz-10 kHz Peak 10-50 Hz	f_g^+-f_p^+,<f_g^-	31, 46	Broadband emission consisting of discrete bursts of V-shaped structure. Associated with VLF hiss in same region and with AKR.	Associated with regions of field-aligned currents in cusp and tail. Ion cyclotron or ion acoustic instability.
Farley Instabilities	Ionosphere E-Region	40 Hz-10 kHz	$< f_g^+$; $\leq f_{LHR}$	27, 67, 77	Narrowband ELF emission at ~ 100 Hz. Broadband VLF emission < 10 kHz at higher altitude.	Hall current instability for electrons streaming through neutrals in collisional plasma. Associated with transverse E-fields.
Electrostatic Electron Cyclotron Emissions	Outside Plasmapause; Near Plasma Sheet	200 Hz-50 kHz	~ $(n+1/2)f_g^-$ near f_p^-	4, 29, 45	Narrowband emission near $(n+1/2)f_g^-$ with several harmonics observed simultaneously.	Loss cone instability for 1-100 keV electrons in rarefied cold plasma. Causes energetic electron precipitation.
Bow Shock Turbulence	Bow Shock Transition Magnetosheath	200 Hz-30 kHz	~ f_p^+	28, 36, 80, 81	Broadband electrostatic noise.	Ion two-stream instability producing ion acoustic waves. Associated with proton heating.
Bow Shock Plasma Oscillations	Bow Shock	3 kHz-50 kHz	~ f_p^-	36, 80, 81, 84	Narrow band electron plasma oscillations associated with electron heating.	Two-stream instability from upstreaming solar wind electrons (~ 1 keV) reflected from bow shock.

Table 3. ESCAPING PLASMA WAVES

PHENOMENON	LOCATION	OBSERVED FREQUENCY	CHARACTERISTIC FREQUENCY	REFERENCES	WAVE PROPERTIES	SOURCE MECHANISMS
Nonthermal Continuum Radiation	Outside Plasmapause Magnetosphere	20 kHz-100 kHz	$> f_p^-$ of solar wind	39, 42, 54	Weak electromagnetic broadband noise escaping magnetosphere. Generated in morning and early afternoon. Associated with bands of electrostatic cyclotron noise at plasmapause.	Gyrosynchrotron radiation from outer radiation belt electrons or coupled from electrostatic waves at plasmapause.
Kilometric Radiation (Auroral or Terrestrial) (AKR or TKR)	Auroral Field Lines	20 kHz-2 MHz	$> f_g^-$ or f_p^-	1, 2, 41, 71, 79	Broadband noise bursts peaked ~ 200 kHz beamed into 2 steradian cone centered on auroral field line at ~ 2 R_E emitting ~ 10^9 W. Associated with VLF hiss, ELF hiss, auroral turbulence and tail electrostatic noise. Generation ~ 2200 LT in tail and ~ 1200 LT in cusp.	Associated with discrete auroral arcs, inverted-V electron events (field aligned currents). Proposed: doppler shifted electron cyclotron radiation $> f_g^-$ (R.H.) or electron beam instability $> f_p^-$ (L.H.) in presence of ion wave turbulence.

Table 4. MAN-MADE PLASMA WAVES

PHENOMENON	LOCATION	OBSERVED FREQUENCY	CHARACTERISTIC FREQUENCY	REFERENCES	WAVE PROPERTIES	SOURCE MECHANISMS
Power System Harmonic Radiation	Plasmapshere Boundary, L ~ 4	kHz	n x 120 Hz, n ~ 25	50	Narrowband radiation at harmonics of the rectified 60 Hz, spaced by 120 Hz and falling in kHz range. Affects other emissions at close frequencies	Harmonic radiation can be amplified by keV electrons at equator. Causes precipitation of keV electrons and may stimulate chorus emissions.
Siple Stimulated VLF Emissions	Plasmasphere Boundary L ~ 3-5	2 kHz-16 kHz	Resonant with electrons	16, 48, 49	Narrowband emissions, rising or falling tones, stimulated and modified by transmitted VLF wave.	100 kW transmitter at Siple, Antartica; growth of ~ 30 dB by cyclotron resonance with > 10 keV electrons at equator.
Electron Beam Stimulated Emissions	Ionosphere	DC-12 MHz	Various	20, 72	Waves stimulated by rocket-borne electron and argon guns: f_p^-, $2f_g^-$, and VLF and ELF whistler mode frequencies.	Electron gun 4-45 kV, 80 ma at 0-180° pitch angles, argon source of 50 V to neutralize rocket.

Table 5A. PLANETARY RADIO BURSTS:
COMPARISON OF PEAK EMISSION FREQUENCIES*

PLANET	PEAK FREQUENCY, MHz OBSERVED	PREDICTED	MAGNETIC FIELD, GAUSS OBSERVED	BODE LAW	FREQ/FIELD, MHz/GAUSS
EARTH	.25	--	0.62	--	0.40
JUPITER	8	--	14.4N 10.8S	--	0.56
SATURN	1.1	2.4	--	4.8	(0.5)
URANUS	0.5	0.8	--	1.6	(0.5)
NEPTUNE	--	0.6	--	1.3	(0.5)

*Adapted from Kennel and Maggs [60] with Uranus observations[11] added. The predicted peak frequencies are obtained by multiplying the Bode Law predicted magnetic field value by the average ratio of the observed peak frequency to polar field value for the Earth and Jupiter.

Table 5B. PLANETARY RADIO BURSTS:
COMPARISON OF PEAK POWER FLUXES*

PLANET	POWER FLUX AT EARTH, $Vm^{-2}\ Hz^{-1}$	EMITTED POWER IN 2π STR, WATTS	POWER FROM SOLAR WIND, WATTS	RATIO EMITTED TO SOLAR WIND
EARTH	--	5×10^{8}	5×10^{11}	1×10^{-3}
JUPITER	2×10^{-19}	1.5×10^{12} [10^{10}]	5×10^{13}	3×10^{-2} [2×10^{-4}]
SATURN	5×10^{-20}	1.5×10^{11}	5×10^{12}	4×10^{-2}
URANUS	1.5×10^{-20}	1.0×10^{11}	4×10^{11}	2.5×10^{-1}
NEPTUNE	(2×10^{-24} to 1×10^{-20})	(1.5×10^{11})	1.5×10^{11}	(2×10^{-4} to 1)

*Adapted from Kennel and Maggs[60] with Uranus observations[11] added.
[ ] Total emitted power estimate including beaming factor.
( ) Scaling to Neptune from other planets. The farther planets seem to have increased efficiencies if the beaming is into 2π steradians. However, the conversion efficiency for Jupiter and the Earth is much less when accounting for the measured beaming factors.

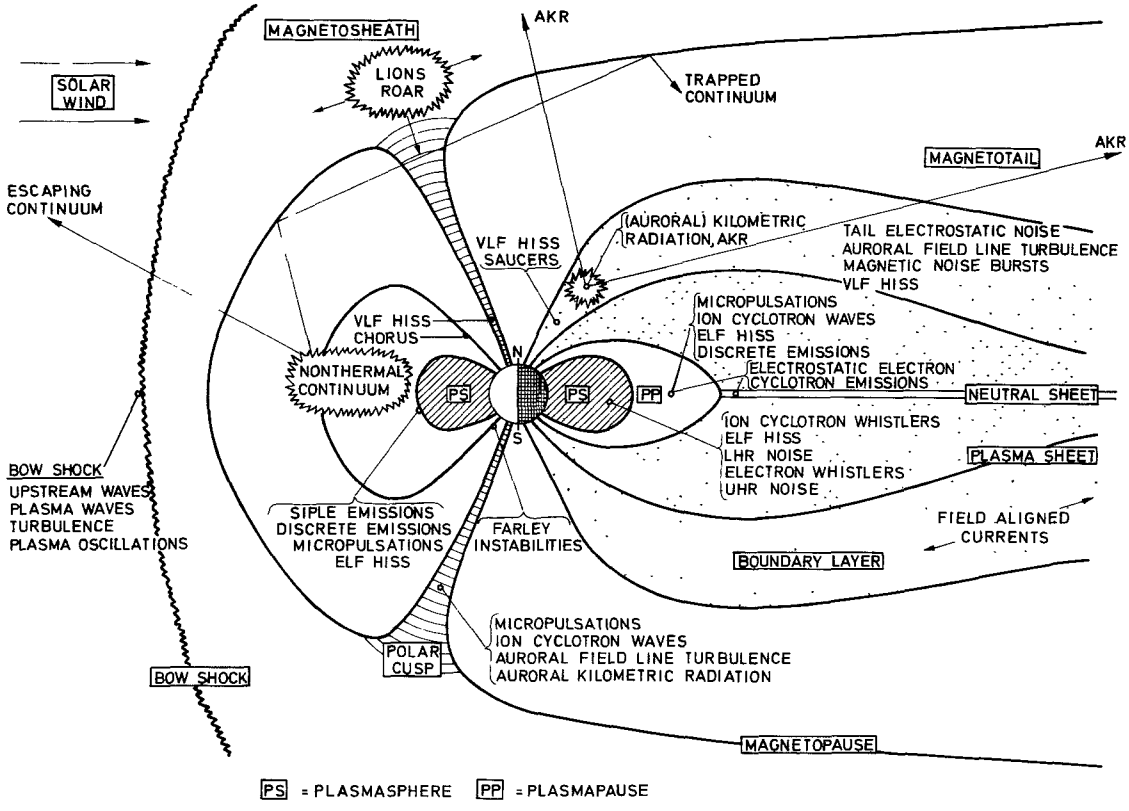


FIGURE 1 Regions of plasma wave occurrence located in a noon-midnight meridian cross-section of the Earth's magnetosphere.

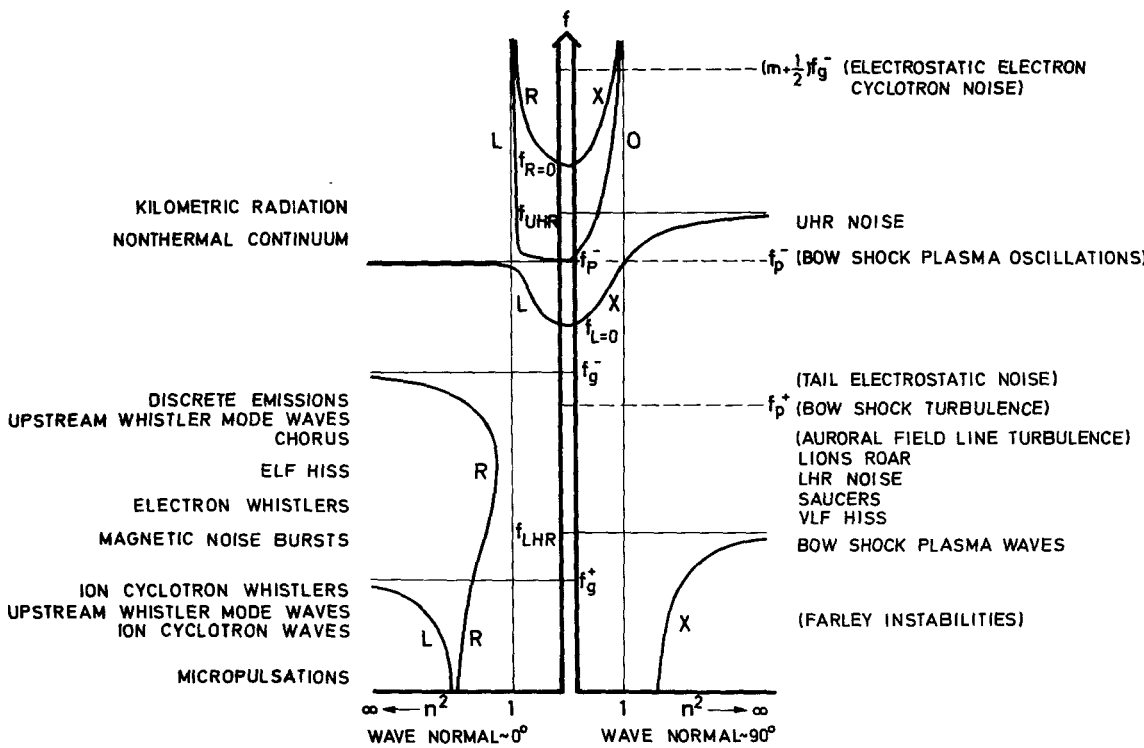


FIGURE 2 Association of plasma wave types with the characteristic frequencies of a plasma for wave normal directions nearly along the geomagnetic field direction (0°) and nearly transverse (90°). The curves represent the index of refraction, n, for the left- and right-hand (L and R) and the extraordinary and ordinary (X and O) wave modes. Wave types in parentheses ( ) are electrostatic (adapted from Jones and Grard[53]).

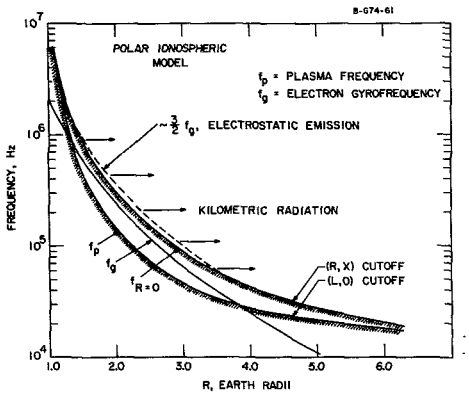


FIGURE 3a Polar magnetospheric model showing the variation of the electron plasma frequency f_p^-, gyrofrequency f_g^- and the cutoff frequency for the right-hand mode $f_{R=0}$. Also indicated is the frequency range for the kilometric radiation which is consistent with the frequency range for which the $3/2\ f_g^-$ electrostatic emission would be expected to exceed the right-hand mode cutoff (from Gurnett[41]).

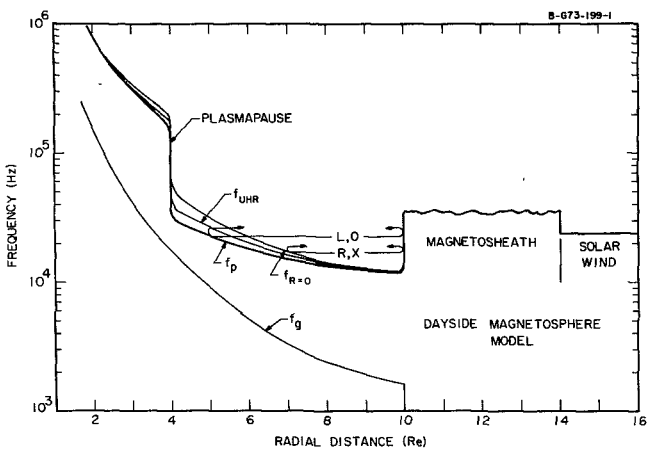


FIGURE 3b Dayside equatorial magnetospheric model showing the radial variation of the gyro, plasma, upper hybrid and R-mode cutoff frequencies. Nonthermal continuum noise is found to be trapped in the low density region between the plasmapause and the magnetosheath (after Gurnett and Shaw[39]).

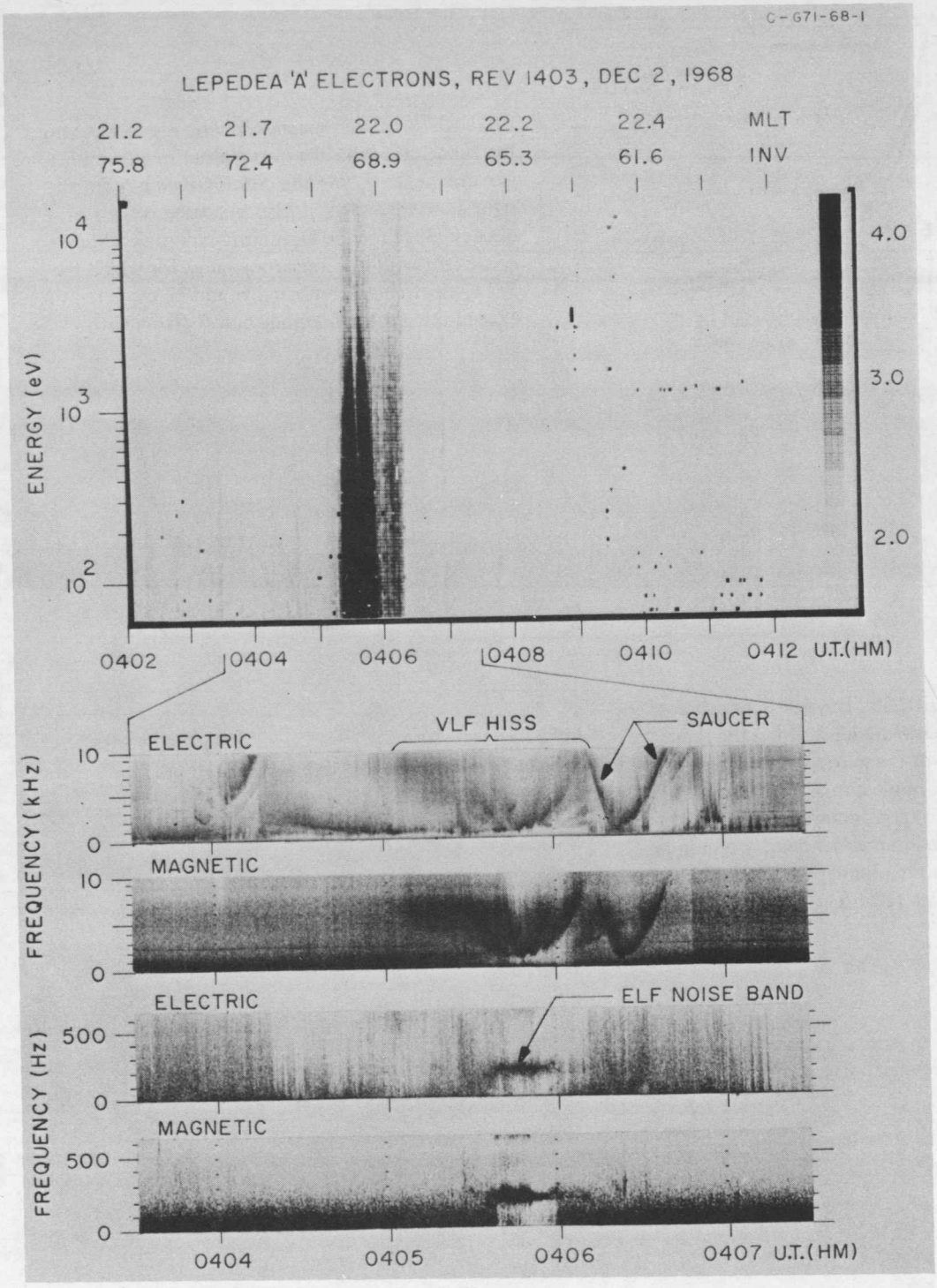


FIGURE 4 Frequency-time spectrogram of VLF hiss, saucer and ELF noise band (similar to lion's roar) associated with an intense auroral electron precipitation event evident on the energy-time spectrogram at ~ 0406 UT observed by Injun 5 at ~ 2500 km altitude (from Gurnett and Frank[38]).

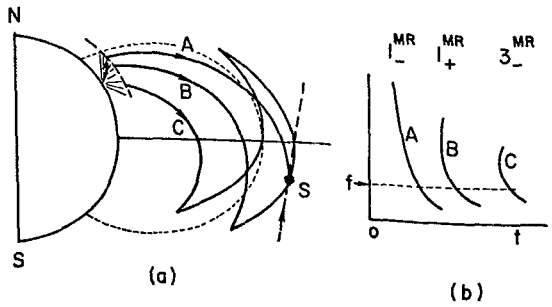


FIGURE 5a Examples of non-ducted electron whistler ray paths in the magnetosphere that make up a magnetospherically reflected (MR) whistler train. Note that the whistler energy can propagate across magnetic field lines (from Smith and Angerami[100]).

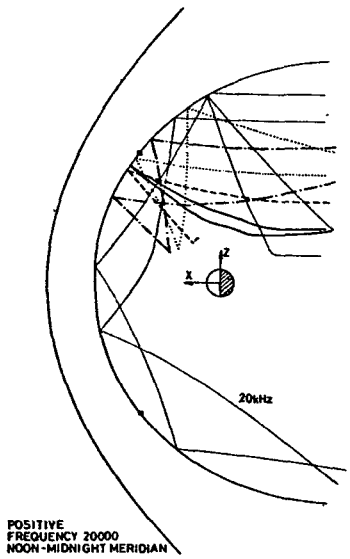


FIGURE 5b Examples of ray paths for nonthermal continuum radiation which is trapped between the magnetopause and plasmapause boundaries (from Jones and Grard[55]).

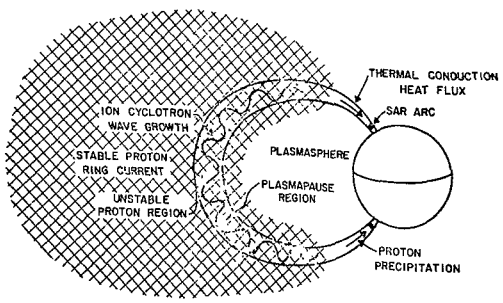


FIGURE 6 Pictorial representation of the hot ring current plasma mixing with the cold plasmasphere plasma to produce ion cyclotron waves at the plasmapause. These waves cause ion precipitation and damp as they propagate leading to a downward heat flux which produces stable auroral red (SAR) arcs in the ionosphere (from Electrodynamics Explorer Final Study Team Report, NASA/GSFC, September 1976).

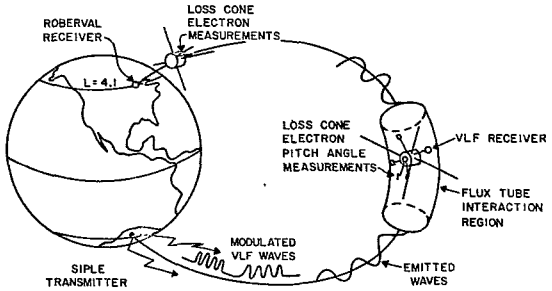


FIGURE 7 Schematic drawing of VLF wave injection from the Siple, Antarctica, transmitter along the $L \approx 4$ field line toward the Roberval, Canada, conjugate point. These injected waves organize the energetic electrons to emit and amplify waves at adjacent but variable frequencies in the equatorial region. These large amplitude waves cause electron precipitation (from Electrodynamics Explorer Final Study Team Report, NASA/GSFC, September 1976).

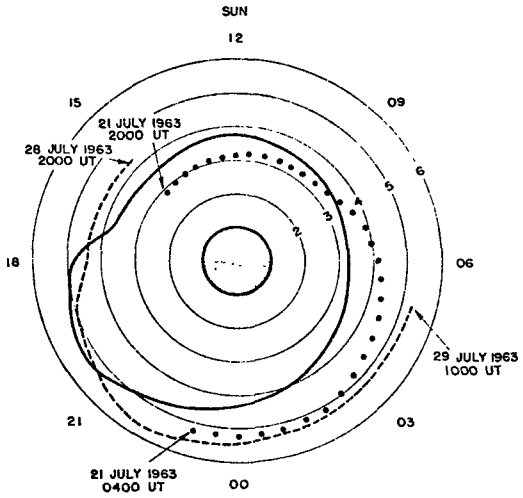


FIGURE 8 The location of the electron density knee (plasmapause boundary) as a function of local time and radial distance as deduced from the dispersion of electron whistlers for three days in July 1963 (from Carpenter[15]).

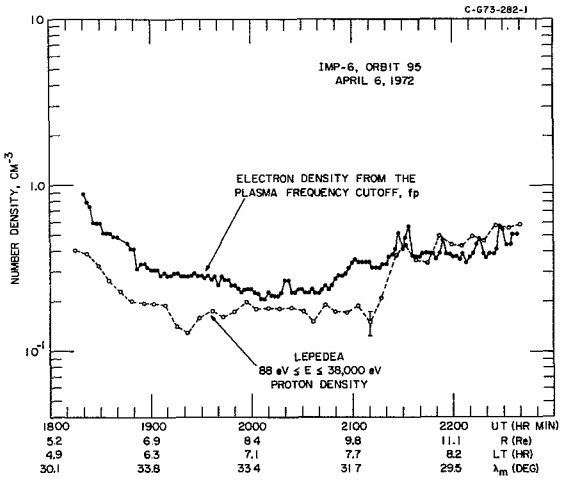


FIGURE 9 Values for the total electron density outside the plasmapause derived from the non-thermal continuum spectral cutoff at the plasma frequency as observed with IMP 6 compared to the supra-thermal proton densities determined simultaneously (from Gurnett and Frank[40]).

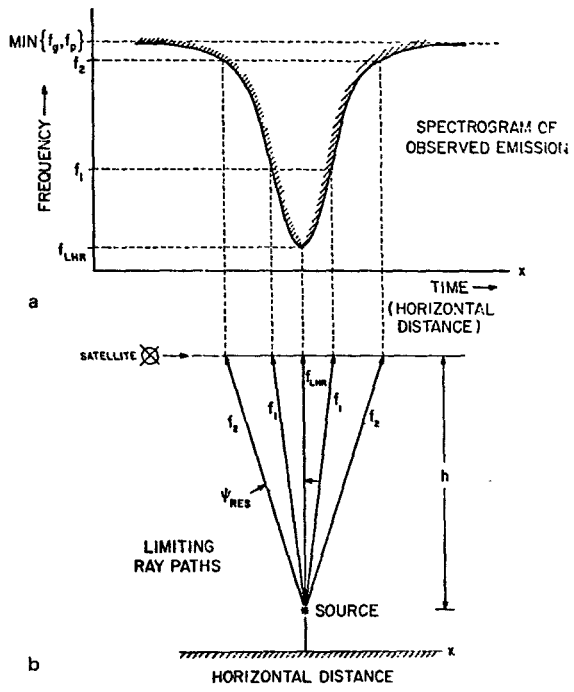


FIGURE 10a Idealized frequency-time (distance) spectrogram of saucer emission extending in frequency from the local lower hybrid resonance frequency at the satellite to the lower of the electron plasma or gyrofrequency (from Mosier and Gurnett[73]).

FIGURE 10b Representative ray paths for the saucer energy from a source below the satellite. Higher frequency components make a large angle to the source field lines (from Mosier and Gurnett[73]).

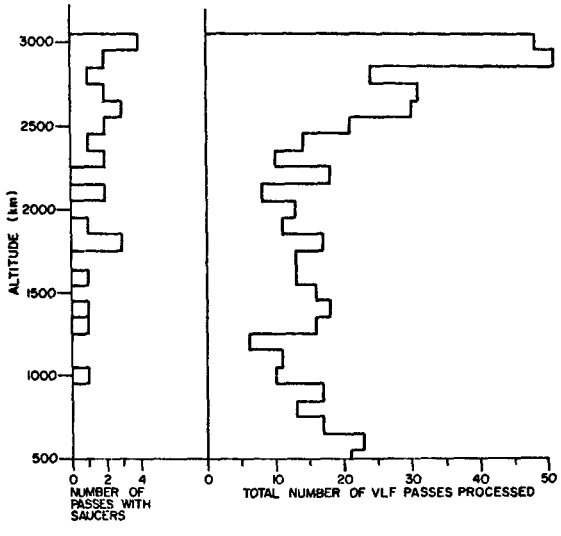


FIGURE 10c Distribution of saucer occurrence with height of the observing satellite. Generally the source region was deduced to be within 100 km of the satellite by ray tracing downward (from James[52]).

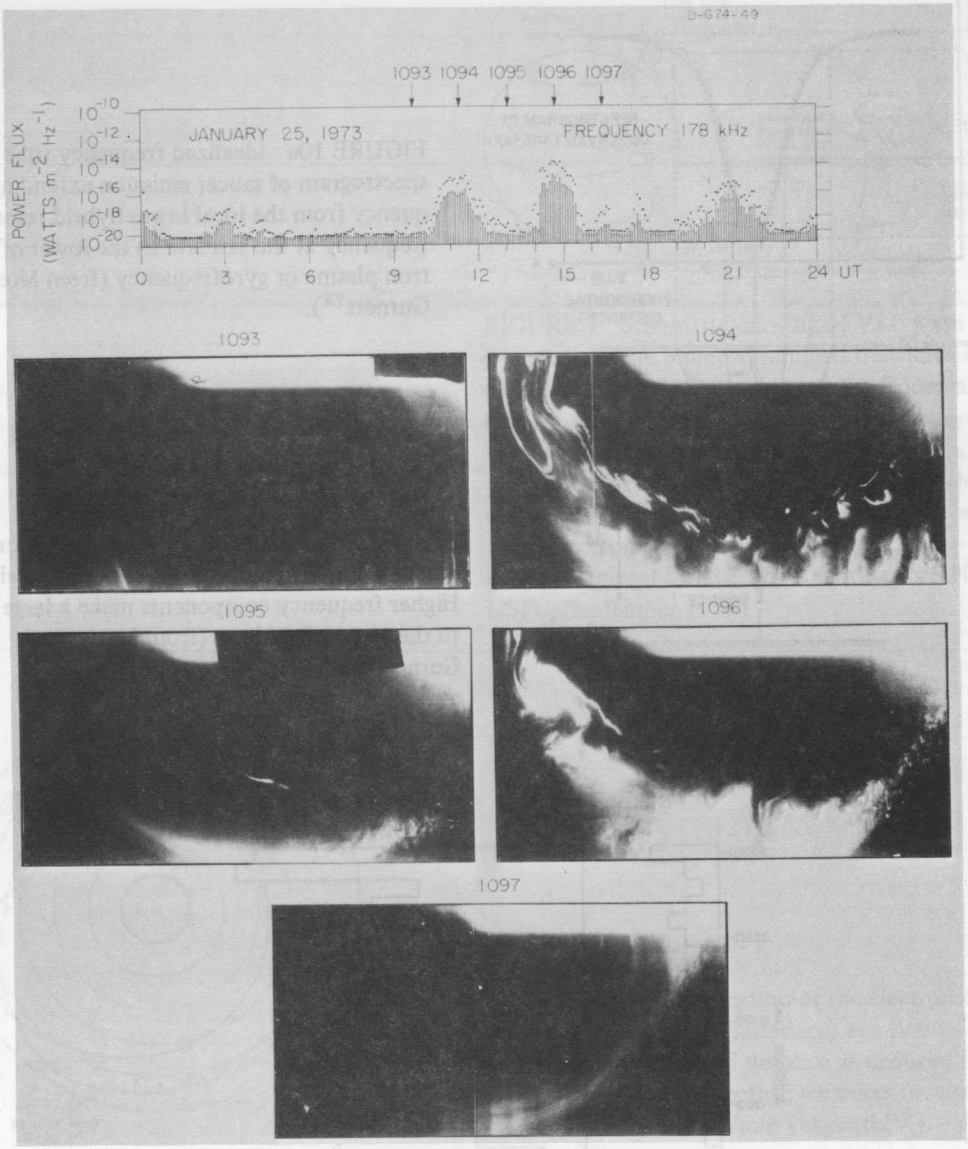


FIGURE 11 The amplitude of kilometric radiation at 178 kHz for a 24 hour period (magnetically active) on 25 January 1973 and a sequence of DAPP auroral zone photographs taken during the same period. The kilometric radiation appears to be more closely related to the discrete auroral arcs than to the diffuse aurora. On frames 1094 and 1096 both discrete and diffuse aurora are evident along with intense AKR. However on frames 1095 and 1097 diffuse aurora is occurring but the AKR is absent (from Gurnett[41]).

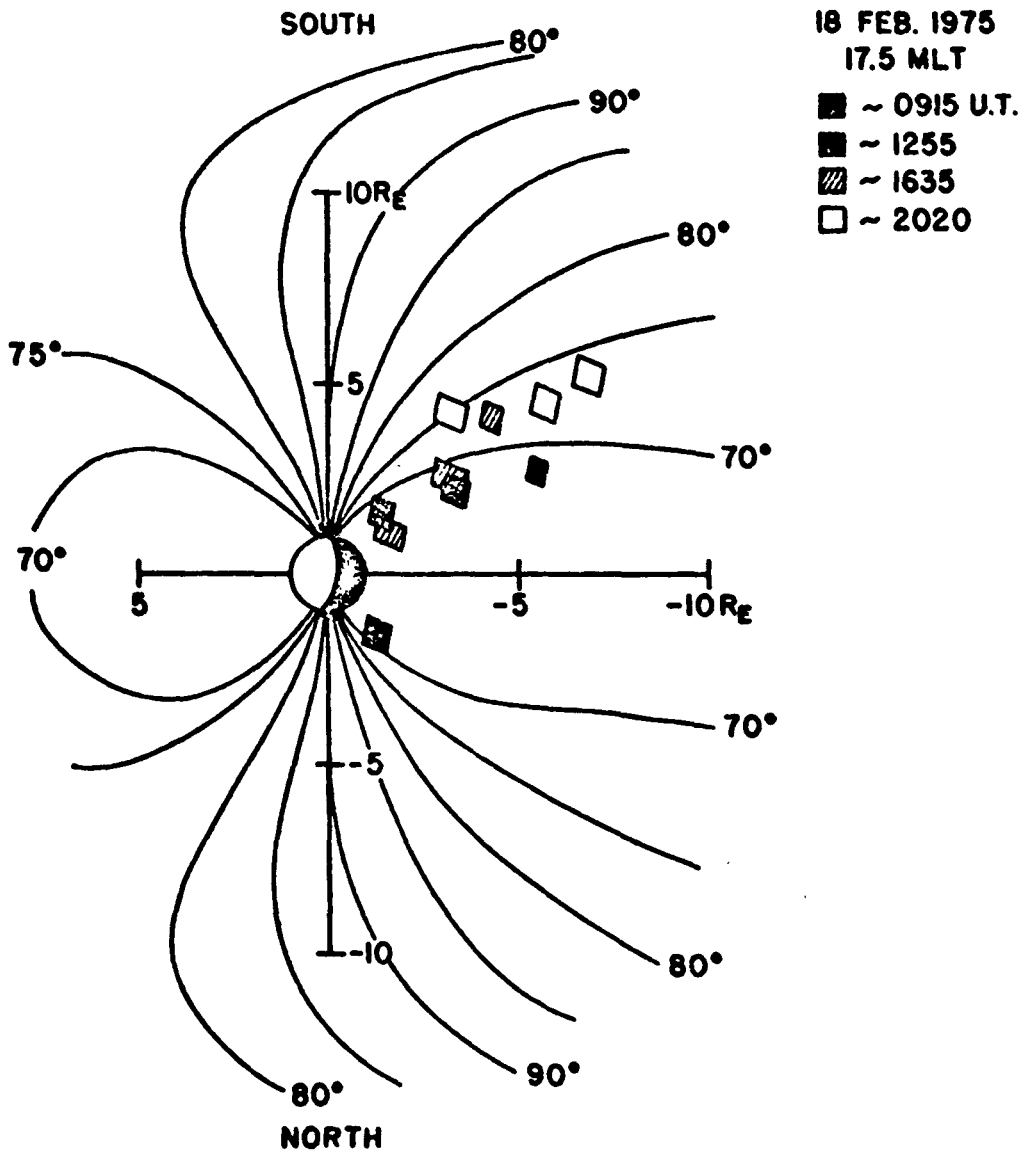


FIGURE 12 Locations of the kilometric radiation source by lunar occultations with RAE 2 at several times during 18 February 1975. Note that the source regions trace out nightside auroral field lines (from Alexander and Kaiser[1]).

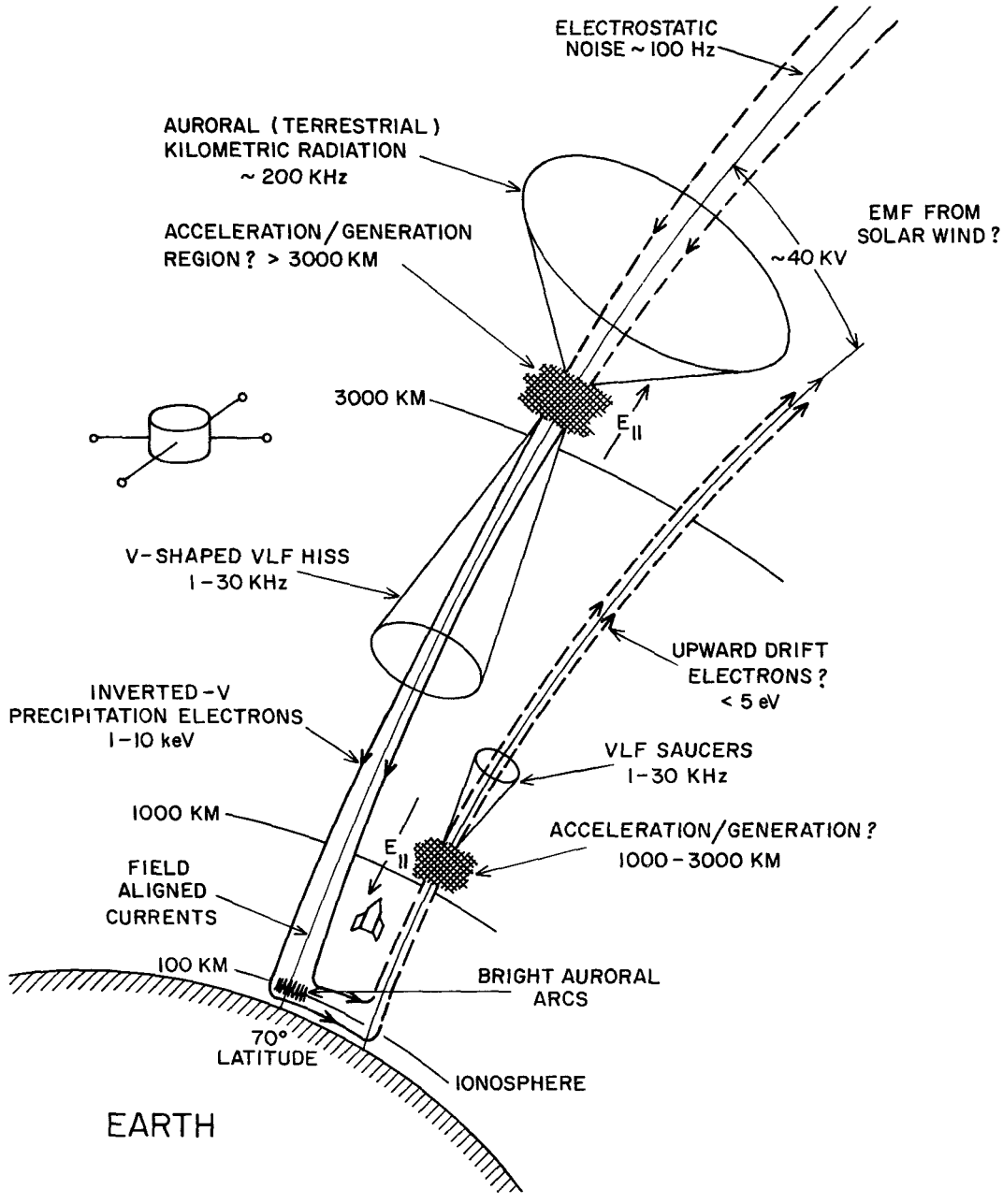


FIGURE 13 Conceptual view of auroral zone plasma wave phenomena and the associated parallel electric field regions, precipitated and accelerated electrons and bright auroral arcs. Kilometric radiation (~ 200 kHz) is an escaping plasma wave from 0.5-6 R_E in altitude whereas the VLF hiss (~ 3 kHz) is a downward propagating whistler-mode wave from nearly the same altitude range. Saucers (~ 3 kHz) are upgoing whistler-mode waves from regions adjacent to the hiss.

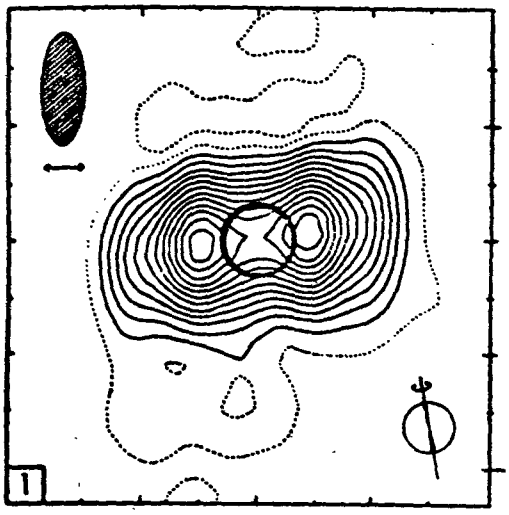
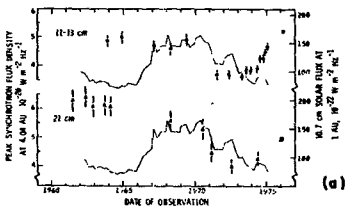
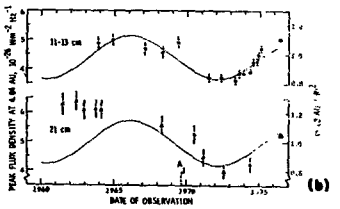


FIGURE 14 Decimetric brightness distribution of Jupiter at 21.3 cm for a central meridian longitude of 15°. Note that this synchrotron radiation peaks on either side of the dark circle representing the optical disk (from Berge and Gulkis[6]).

FIGURE 15 A comparison of Jupiter's escaping radiation at two wavelengths (~ 12 cm and 21 cm) with (a) the 10.7 cm solar radio flux (jagged curve) and (b) the square of Jupiter's distance from the sun (sine curve with adjusted phase) as a function of date. The large intensity variations are not correlated with solar or solar wind parameters directly but may be due to changes in the trapped electron distribution function and plasma parameters of the inner Jovian magnetosphere (adapted from Klein[62]).





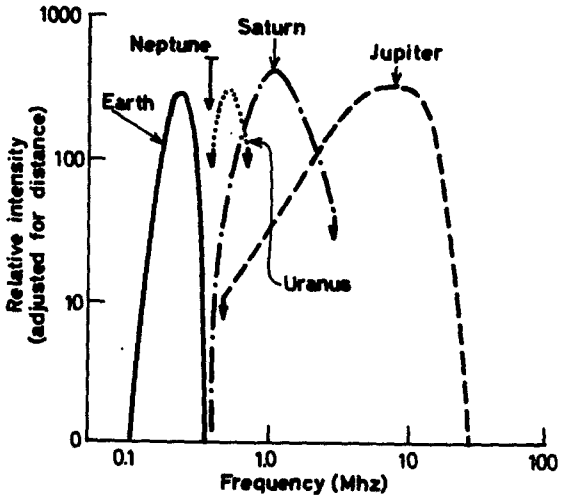


FIGURE 16 Relative intensities of the observed escaping decametric/hectometric/kilometric radiation as a function of frequency for Jupiter, Saturn and Uranus and the Earth. A predicted range for Neptune (see Table 5) is also indicated (adopted from Kaiser and Stone[57] and Brown[11]).

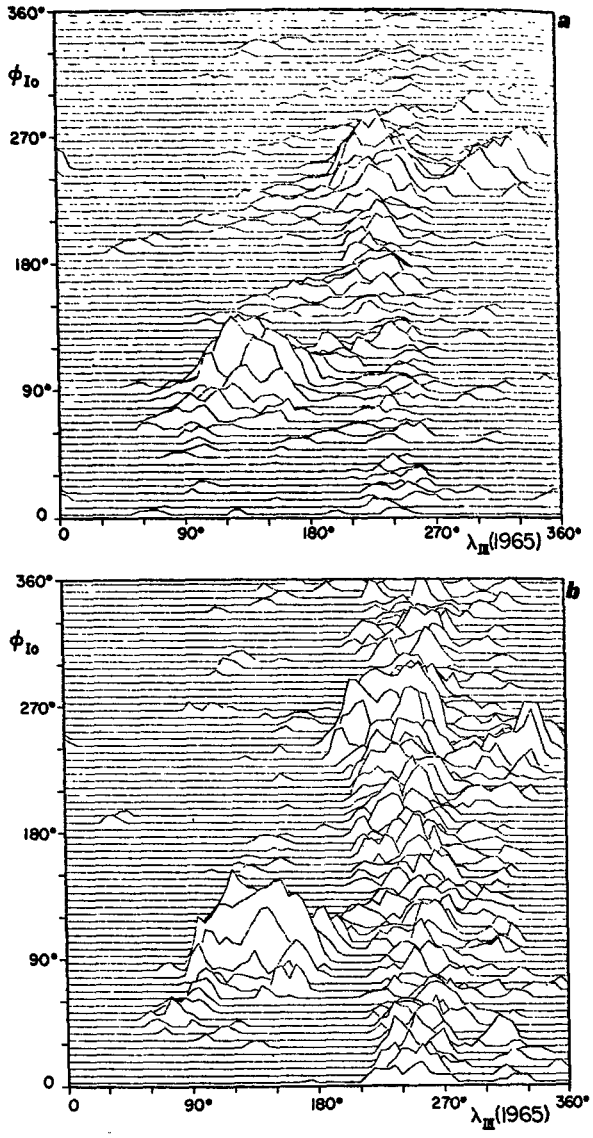


FIGURE 17 Occurrence probability for Jovian decametric emission as a function of the Io orbital position (φ_{Io}, $\varphi = 180°$ is toward Earth) and the central median longitude of Jupiter λ_{III} for two different time periods. Note particularly the Io-associated regions near $\varphi \simeq 90°$ and 240° and the Io-independent emission for $\lambda \simeq 240°$-270° (from Boyzan and Douglas[10]).

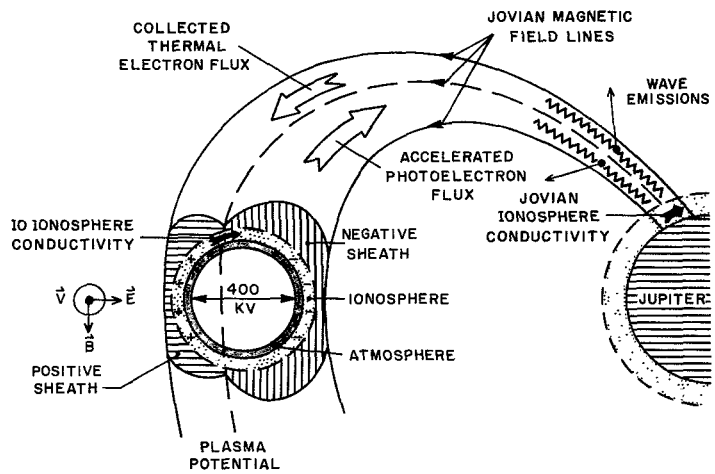


FIGURE 18 The Io-sheath-acceleration model. Io's rapid motion through the Jovian magnetic field sets up a motional potential of several hundred kilovolts. This potential is dropped across plasma sheaths in which electrons and ions can be accelerated to 100 keV energies. The potential drives a beam of these high energy electrons (carrying up to 10^{13} watts) toward the Jovian atmosphere. By a plasma instability along the field lines decametric radiation is stimulated and amplified (adapted from Shawhan *et al.*[94]).

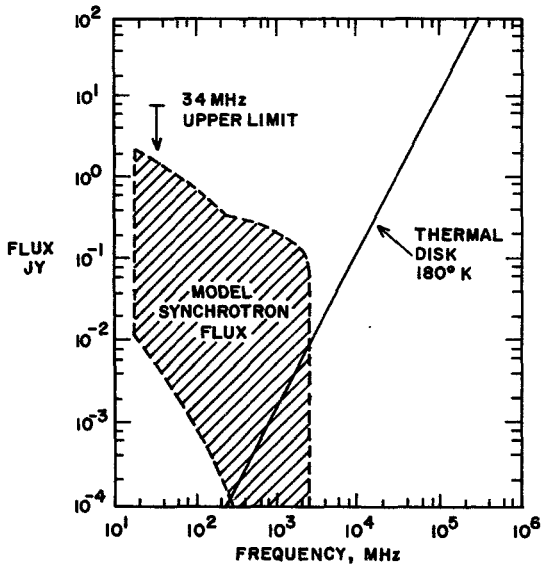


FIGURE 19 Possibility of detecting synchrotron radiation from Uranus. Most observations are above 1000 MHz and are consistent with the blackbody emission from a 180°K disk. Model calculations (shaded region) indicate that the emission should fall below 1000 MHz where radiotelescopes are less sensitive. An upper limit measurement for 34 MHz (from one of the largest low frequency radiotelescopes) is indicated. Larger radiotelescopes or measurements closer to the planet are required (from Shawhan and Cronyn[96]).

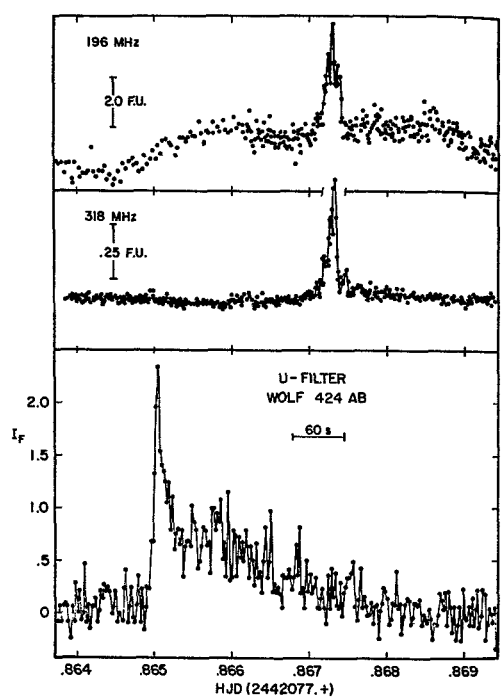


FIGURE 20 Simultaneous detection of a radio (196 MHz and 318 MHz) and optical (ultraviolet) outburst from the flarestar Wolf 424 AB on 30 January 1974. These outbursts contain orders of magnitude more energy than for a solar flare but the time scale (10's seconds) and high degree of radio polarization (60%) suggest that the basic mechanisms may be similar (from Spangler and Moffett[101]).

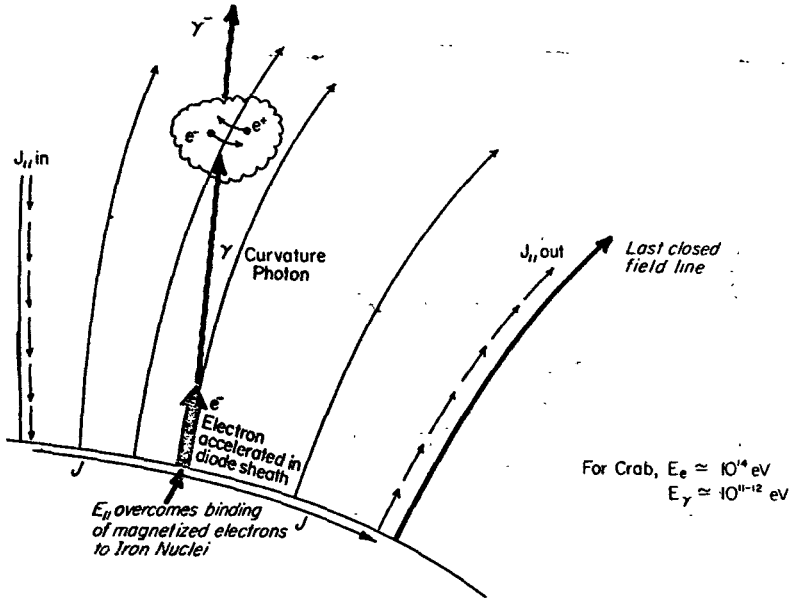


FIGURE 21 Sturrock's model of pulsars. Shown here is the pattern of currents flowing in and out of the polar cap of rotating magnetized neutron star whose dipole and spin axes are aligned. In this model the current is carried by beams of superrelativistic electrons, which emit curvature photons. When these photons propagate at a small angle to a superstrong magnetic field they can produce pairs which also radiate photons, thereby breaking down the vacuum and populating the outer magnetosphere with plasma. The electron-positron plasma is subjected to a two stream instability which produces emission in the radio spectrum (from Kennel[59]).

ACKNOWLEDGMENTS

I thank Drs. E. Ungstrup, C.-G. Fälthammar, L. Lanzerotti, R. Fredricks, D. Gurnett, C. Goertz, N. D'Angelo, and J. A. Van Allen for suggestions on the content of this review and I appreciate the detailed comments of Drs. T. Bell, R. Thorne and L. Block on the manuscript. Work on this review was initiated while I was a visiting scientist at the Danish Space Research Institute, Lyngby, Denmark, and at the Department of Plasma Physics, the Royal Institute of Technology, Stockholm, Sweden. Support at The University of Iowa is provided in part by NASA Grant NGL 16-001-002 and NSF Grant ATM76-82739.

REFERENCES

1. Alexander, J. K. and Kaiser, M. L., Terrestrial kilometric radiation: 1.-Spatial structure studies, J. Geophys. Res., 81, 5948-5956, 1976.

2. Alexander, J. K. and Kaiser, M. L., Terrestrial kilometric radiation: 2-Emission from the magnetospheric cusp and dayside magnetosheath, J. Geophys. Res., 82, 98-104, 1976.

3. Al'pert, Ya. L., Waves and Satellites in the Near-Earth Plasma, Consultants Bureau: New York, 1974.

4. Ashour-Abdalla, M. and Kennel, C. F., VLF electrostatic waves in the magnetospheres of Earth and Jupiter, The Scientific Satellite Programme During the IMS, K. Knott and B. Battrick, Eds., 305-325, D. Reidel: Dordrecht-Holland, 1976.

5. Barbosa, D. D. and Coroniti, F. V., Relativistic electrons and whistlers and Jupiter's magnetosphere, J. Geophys. Res., 81, 4531-4536, 1976.

6. Berge, G. L. and Gulkis, S., Earth-based observations of Jupiter: millimeter to meter wavelengths, Jupiter, the Giant Planet, T. Gehrels, Ed., U. of Arizona Press: Tucson, 1976.

7. Bernstein, W., Leinbach, H., Cohen, H., Wilson, P. S., Davis, T. N., Hallinan, T., Baker, B., Martz, J., Zeimke, R., and Huber, W., Laboratory observations of RF emissions at ω_{pe} and $(n + 1/2)\ \omega_{ce}$ in electron beam-plasma and beam-beam interactions, J. Geophys. Res., 80, 4375-4379, 1975.

8. Block, L. P., and Fälthammar, C.-G., Mechanisms that may support magnetic-field-aligned electric fields in the magnetosphere, Ann. Geophys., 32, 161-174, 1976.

9. Bossen, M., McPherron, R. L., and Russell, C. T., A statistical study of Pcl magnetic pulsations at synchronous orbit, J. Geophys. Res., 81, 6083-6091, 1976.

10. Boyzan, F. A., and Douglas, J. N., Directivity and stimulation in Jovian decametric radiation, J. Geophys.Res., 81, 3387-3392, 1976.

11. Brown, L. W., Possible radio emission from Uranus at 0.5 MHz, Astrophys. J., 207, L202-L212, 1976.

12. Budden, K. G., Radio Waves in the Ionosphere, Cambridge University Press: London, 1961.

13. Burton, R. K. and Holzer, R. E., The origin and propagation of chorus in the outer magnetosphere, J. Geophys. Res., 79, 1014-1023, 1974.

14. Carlqvist, P., Current limitation and solar flares, Solar Physics, 7, 377-392, 1969.

15. 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, 693-709, 1966.

16. Carpenter, D. L., and Miller, T. R., Ducted magnetospheric propagation of signals from the Siple, Antarctica, VLF transmitter, J. Gecphys. Res., 81, 2692-2700, 1976.

17. Carpenter, D. L., and Seely, N. T., Cross-L plasma drifts in the outer plasmasphere: quiet time patterns and some substorm effects, J. Geophys. Res., 81, 2728-2736, 1976.

18. Carr, T. D. and Gulkis, S., The magnetosphere of Jupiter, in Ann. Rev. Astron. Astrophys., L. Goldberg, Ed., 7, 577-618, 1969.

19. Carr, T. D. and Desch, M. D., A review of recent decametric and hectometric observations of Jupiter, Jupiter, the Giant Planet, T. Gehrels, Ed., U. of Arizona Press: Tucson, 1976.

20. Cartwright, D. G. and Kellogg, P. J., Observations of radiation from an electron beam artificially injected into the ionosphere, J. Geophys. Res., 79, 1439-1458, 1974.

21. Ceresier, J. C., Propagation perpendiculaire au voisinage de la fréquence de la résonance hybride basse, Plasma Waves in Space and in the Laboratory, Edinburgh University Press, 1970.

22. Chan, K.-W. and Holzer, R. E., ELF hiss associated with plasma density enhancements in the outer magnetosphere, J. Geophys. Res., 81, 2267-2274, 1976.

23. Coroniti, F. V., Fredricks, R. W., and White, R., Instability of ring current protons beyond the plasmapause during injection events, J. Geophys. Res., 77, 6243-6248, 1972.

24. D'Angelo, N., Ultralow frequency fluctuations at the polar cusp boundaries, J. Geophys. Res., 78, 1206-1209, 1973.

25. Fairfield, D. H., Whistler waves observed upstream from collisionless shocks, J. Geophys. Res., 79, 1368-1378, 1974.

26. Fälthammar, C.-G., Laboratory experiments of magnetospheric interest, Space Sci. Rev., 15, 803-825, 1974.

27. Farley, D. T., A plasma instability resulting in field-aligned irregularities in the ionosphere, J. Geophys. Res., 68, 6083-6097, 1963.

28. Fredricks, R. W., Scarf, F. L., and Frank, L. A., Nonthermal electrons and high-frequency waves in the upstream solar wind, J. Geophys. Res., 76, 6691-6699, 1971.

29. Fredricks, R. W. and Scarf, F. L., Recent studies of magnetospheric electric field emissions above the electron gyrofrequency, J. Geophys. Res., 78, 310-314, 1973.

30. Fredricks, R. W. and Russell, C. T., Ion cyclotron waves observed in the polar cusp, J. Geophys. Res., 78, 2917-2925, 1973.

31. Fredricks, R. W., Scarf, F. L., and Russell, C. T., Field-aligned currents, plasma waves and anomalous resistivity in the disturbed polar cusp, J. Geophys. Res., 78, 2133-2141, 1973.

32. Fredricks, R. W., A model for generation of bow-shock-associated up-stream waves, J. Geophys. Res., 80, 7-17, 1975.

33. Fredricks, R. W., Wave-particle interactions in the outer magnetosphere: A review, Space Science Reviews, 17, 741-780, 1975.

34. Gendrin, R., Waves and wave-particle interactions in the magnetosphere: a review, Space Science Reviews, 18, 145-200, 1975.

35. Goertz, C. K. and Joyce, G., Numerical simulation of the plasma double layer, Astrophys. and Space Sci., 32, 165-173, 1975.

36. Greenstadt, E. W. and Fredricks, R. W., Plasma instability modes related to the earth's bow shock, in Magnetospheric Physics, B. M. McCormac, Ed., D. Reidel: Dordrecht-Holland, 1974.

37. Gurnett, D. A., Shawhan, S. D., Brice, N. M., and Smith, R. L., Ion cyclotron whistlers, J. Geophys. Res., 70, 1665-1688, 1965.

38. Gurnett, D. A. and Frank, L. A., VLF hiss and related plasma observations in the polar magnetosphere, J. Geophys. Res., 77, 172-190, 1972.

39. Gurnett, D. A. and Shaw, R. R., Electromagnetic radiation trapped in the magnetosphere above the plasma frequency, J. Geophys. Res., 78, 8136-8149, 1973.

40. Gurnett, D. A. and Frank, L. A., Thermal and suprathermal plasma densities in the outer magnetosphere, J. Geophys. Res., 79, 2355-2361, 1974.

41. Gurnett, D. A., The earth as a radio source: terrestrial kilometric radiation, J. Geophys. Res., 79, 4227-4238, 1974.

42. Gurnett, D. A., The earth as a radio source: the non-thermal continuum, J. Geophys. Res., 80, 2751-2763, 1975.

43. Gurnett, D. A., Electrostatic turbulence in the magnetosphere, Physics of Solar Planetary Environments, D. J. Williams, Ed., American Geophysical Union: Washington, DC, Volume II, 760-771, 1976.

44. Gurnett, D. A., The earth as a radio source, Magnetospheric Particles and Fields, B. M. McCormac, Ed., D. Reidel: Dordrecht-Holland, 1976.

45. Gurnett, D. A., Frank, L. A., and Lepping, R. P., Plasma waves in the distant magnetotail, J. Geophys. Res., 81, 6059, 1976.

46. Gurnett, D. A. and Frank, L. A., A region of intense plasma wave turbulence on auroral field lines, J. Geophys. Res., 82, 1031-1050, 1977.

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

48. Helliwell, R. A., Controlled VLF wave injection experiments in the magnetosphere, Space Sci. Rev., 15, 781-802, 1974.

49. Helliwell, R. A. and Katsufrakis, J. P., VLF wave injection into the magnetosphere from Siple station, Antartica, J. Geophys. Res., 79, 2511-2518, 1974.

50. Helliwell, R. A., Katsufrakis, J. P., Bell, T. F., and Raghuram, R., VLF line radiation in the Earth's magnetosphere and its association with power system radiation, J. Geophys. Res., 80, 4249-4258, 1975.

51. Jacobs, J. A., Geomagnetic Micropulsations, Springer-Verlag: Berlin, 1970.

52. James, H. G., VLF saucers, J. Geophys. Res., 81, 501-514, 1976.

53. Jones, D. and Grard, R. J. L., Electromagnetic wave propagation in the outer magnetosphere and near-Earth magnetotail, Eldo-Cecles/ESRO-Cers Scient. and Tech. Rev., 6, 97-108, 1974.

54. Jones, D., Mode coupling of Cerenkov radiation as a source of noise above the plasma frequency, in K. Knott and B. Battrick, Eds., The Scientific Programme During the International Magnetospheric Study, D. Reidel: Dordrecht-Holland, 281-291, 1976.

55. Jones, D. and Grard, R. J. L., Propagation characteristics of electromagnetic waves in the magnetosphere, in The Scientific Satellite Programme During the IMS, K. Knott and B. Battrick, Eds., 293-302, D. Reidel: Dordrecht-Holland, 1976.

56. Joselyn, J. A. and Lyons, L. R., Ion cyclotron growth calculated from satellite observations of the proton ring current during storm recovery, J. Geophys. Res., 81, 2275-2282, 1976.

57. Kaiser, M. J. and Stone, R. G., Earth as an intense planetary radio source: similarities to Jupiter and Saturn, Science, 189, 285-287, 1975.

58. Kavanagh, L. D., Jr., Synchrotron radio emission from Uranus and Neptune, Icarus, 25, 166-170, 1975.

59. Kennel, C. F., What we have learned from the magnetosphere, Comments on Astrophys. and Space Sci., 6, 71-80, 1975.

60. Kennel, C. F. and Maggs, J. E., Possibility of detecting magnetospheric radio bursts from Uranus and Neptune, Nature, 261, 299-301, 1976.

61. Kimura, I., Interrelation between VLF and ULF emissions, Space Sci. Rev., 16, 389-411, 1974.

62. Klein, M. J., The variability of the total flux density and polarization of Jupiter's decimeter radio emission, J. Geophys. Res., 81, 3380-3382, 1976.

63. Kundu, M., Solar Radio Astronomy, Interscience: New York, 1965.

64. Laaspere, T. and Taylor, J. A., Jr., Comparison of certain VLF noise phenomena with the lower hybrid resonance frequency calculated from simultaneous ion composition measurements, J. Geophys. Res., 75, 97-106, 1970.

65. Laaspere, T. and Hoffman, R. A., New results on the correlation between low-energy electrons and auroral hiss, J. Geophys. Res., 81, 524-530, 1976.

66. Lanzerotti, L. J. and H. Fukunishi, Modes of magnetohydrodynamic waves in the magnetosphere, Rev. Geophys. Sp. Sci., 12, 724-729, 1972.

67. Lee, K., Kennel, C. F., and Kindel, J. M., High-frequency hall current instability, Radio Science, 6, 209-214, 1971.

68. Luthey, J. L., Possibility of Saturnian synchrotron radiation, Icarus, 20, 125-135, 1973.

69. Maggs, J. E., Coherent generation of VLF hiss, J. Geophys. Res., 81, 1707-1724, 1976.

70. McPherron, R. L., Russell, C. T., and Coleman, P. J., Jr., Fluctuating magnetic fields in the magnetosphere, II: VLF waves, Space Sci. Rev., 13, 411-454, 1972.

71. Melrose, D. B., An interpretation of Jupiter's decametric radiation and the terrestrial kilometric radiation as direct amplified gyromechanism, Astrophys. J., 207, 651-662, 1976.

72. Monson, S. J., Kellogg, P. J., and Cartwright, D. G., Whistler mode plasma waves observed on Electron Echo 2, J. Geophys. Res., 81, 2193-2199, 1976.

73. Mosier, S. R. and Gurnett, D. A., VLF measurements of the Poynting flux along the geomagnetic field line with the Injun 5 satellite, J. Geophys. Res., 74, 5675-5687, 1969.

74. Mosier, S. R., Kaiser, M. L., and Brown, L. W., Observation of noise bands associated with the upper hybrid resonance by the IMP 6 radio astronomy experiment, J. Geophys. Res., 78, 1673-1680, 1973.

75. Mozer, F. S., Analyses of techniques for measuring DC and AC electric fields in the magnetosphere, Space Sci. Rev., 272-313, 1973.

76. Northtrop, T. G. and Birmingham, T. J., The magnetosphere of Jupiter as observed with Pioneer 10. 3. Jovian synchrotron radiation at 10.4 cm as deduced from observed electron fluxes, J. Geophys. Res., 79, 3583-3587, 1974.

77. Olesen, J. K., Primdahl, F., Spanglev, F., Ungstrup, E., Bahnsen, A., Fahleson, U., Fälthammar, C.-G., and Pedersen, A., Rocket-borne wave, field and plasma observations in unstable polar cap E-region, Geophys. Res. Lett., 3, 399-402, 1976.

78. Orr, D., Probing the plasmapause by geomagnetic pulsations, Ann. Geophys., 31, 77-92, 1975.

79. Palmadesso, P., Coffey, T. P., Ossakow, S. L., and Papadopoulos, K., Generation of terrestrial kilometric radiation by a beam-driven electromagnetic instability, J. Geophys. Res., 81, 1762-1770, 1976.

80. Rodriguez, P. and Gurnett, D. A., Electrostatic and electromagnetic turbulence associated with the Earth's bow shock, J. Geophys. Res., 80, 19-31, 1975.

81. Rodriguez, P. and Gurnett, D. A., Correlation of bow shock plasma wave turbulence with solar wind parameters, J. Geophys. Res., 81, 2871-2882, 1976.

82. Rosenberg, T. J., Helliwell, R. A., and Katsufrakis, J. P., Electron precipitation associated with discrete very-low-frequency emissions, J. Geophys. Res., 76, 8445-8452, 1971.

83. Russell, C. T., McPherron, R. L., and Coleman, P. J., Jr., Fluctuating magnetic fields in the magnetosphere, 1. ELF and VLF, Space Sci. Rev., 12, 810-856, 1972.

84. Scarf, F. L., Fredricks, R. W., Frank, L. A., and Neugebauer, M., Nonthermal electrons and high-frequency waves in the upstream solar wind, J. Geophys. Res., 76, 5162-5171, 1971.

85. Scarf, F. L., Fredricks, R. W., Russell, C. T., Kivelson, M., Neugebauer, M., and Chappell, C. R., Observations of a current-driven plasma instability at the outer zone-plasma sheet boundary, J. Geophys. Res., 78, 2150-2165, 1973.

86. Scarf, F. L., Controlled experiments from the space shuttle, Space Sci. Rev., 15, 933-950, 1974.

87. Scarf, F. L., A new model for the high-frequency decametric radiation from Jupiter, J. Geophys. Res., 79, 3835-3839, 1974.

88. Scarf, F. L., Characteristics of instabilities in the magnetosphere deduced from wave observations, Physics of the Hot Plasma in the Magnetosphere, B. Hultqvist and L. Stenflo, Eds., 371-290, Plenum Press: New York, 1975.

89. Scarf, F. L. and Russell, C. T., Magnetospheric dynamics and wave-particle interactions, The Scientific Satellite Programme during the International Magnetospheric Study, K. Knott and B. Battrick, Eds., 261-280, D. Reidel: Dordrecht-Holland, 1976.

90. Sentman, D. D., Van Allen, J. A., and Goertz, C. K., Recirculation of energetic particles in Jupiter's magnetosphere, Geophys. Res. Lett., 2, 465-468, 1975.

91. Shawhan, S. D., Experimental observations of proton whistlers from the Injun 3 VLF data, J. Geophys. Res., 71, 29-45, 1966.

92. Shawhan, S. D., Whistlers - Use for determination of composition and temperature, Low Frequency Waves and Irregularities in the Ionosphere, N. D'Angelo, Ed., 94-102, D. Reidel: Dordrecht-Holland, 1969.

93. Shawhan, S. D., The use of multiple receivers to measure the wave characteristics of VLF noise in space, Space Sci. Rev., 10, 689-736, 1970.

94. Shawhan, S. D., Goertz, C. K., Hubbard, R. F., Gurnett, D. A. and Joyce, G., Io-accelerated electrons and ions, The Magnetospheres of the Earth and Jupiter, V. Formissano, Ed., D. Reidel Pub. Co.: Dordrecht-Holland, 375-389, 1975.

95. Shawhan, S. D., Io sheath-accelerated electrons and ions, J. Geophys. Res., 81, 3373-3379, 1976.

96. Shawhan, S. D., and Cronyn, W. M., An upper limit for the 34.3 MHz synchrotron flux from Uranus, J. Geophys. Res., 82, 1947-1950, 1977.

97. Smith, E. J., Davis, L., Jones, D. E., Coleman, P. J., Colburn, D. S., Dyal, P., and Sonnett, C. P., Jupiter's magnetic field, magnetosphere and interaction with the solar wind: Pioneer 11, Science, 188, 451-455, 1975.

98. Smith, E. J. and Tsurutani, B. T., Magnetosheath lions roars, J. Geophys. Res., 81, 2261-2266, 1976.

99. Smith, R. A., Models of Jovian decametric radiation, Jupiter, the Giant Planet, T. Gehrels, Ed., U. of Arizona Press: Tucson, 1976.

100. Smith, R. L. and Angerami, J. J., Magnetospheric properties deduced from OGO 1 observations of ducted and non-ducted whistlers, J. Geophys. Res., 73, 1-20, 1968.

101. Spangler, S. R. and Moffett, T. J., Simultaneous radio and optical observations of UV-Ceti-type flare stars, Astrophys. J., 203, 497-508, 1976.

102. Stix, T. H., The Theory of Plasma Waves, McGraw-Hill: New York, 1962.

103. Storey, L. R. O., and Lefeuvre, F., Theory for the interpretation of measurements of a random electromagnetic wave field in space, in Space Research XIV, M. Rycroft and D. Reasenberg, Eds., 381 - 386, Adademie-Verlag: Berlin, 1974.

104. Sturrock, P. A., A model of pulsars, Astrophys. J., 164, 529-556, 1971.

105. Taylor, W. W. L. and Shawhan, S. D., A test of incoherent Cerenkov radiation for VLF hiss and other magnetospheric emissions, J. Geophys. Res., 79, 105-117, 1974.

106. Thorne, R. M., Smith, E. J., Burton, R. K., and Holzer, R. E., Plasmaspheric hiss, J. Geophys. Res., 78, 1581-1596, 1973.

107. Thorne, R. M., Wave-particle interactions in the magnetosphere and ionosphere, Rev. Geophys. Sp. Phys., 13, 291-302, 1975.

108. Troitskaya, V. A., Micropulsations and the state of the magnetosphere, in Solar Terrestrial Physics, King and Newman, Eds., 213-274, Academic Press: London, 1967.

109. Tsurutani, B. T. and Smith, E. J., Postmidnight chorus: a substorm phenomenon, J. Geophys. Res., 79, 118-127, 1974.

110. Van Allen, J. A., On the magnetosphere of Jupiter, Saturn and Uranus, Highlights of Astronomy, E. A. Müller, ed., D. Reidel: Dordrecht-Holland, vol. 4, 1977.

111. Warwick, J. W., Radiophysics of Jupiter, Space Sci. Rev., 6, 841-4891, 1967.

112. Webb, D. C., Lanzerotti, L. J., and Park, C. G., A comparison of ULF and VLF measurements of magnetospheric cold plasma densities, J. Geophys. Res., accepted for publication, 1977.

THE IONOSPHERIC PLASMA

Donald T. Farley
School of Electrical Engineering
Cornell University
Ithaca NY 14853

ABSTRACT - We consider the ionosphere here as a plasma, and only incidentally as part of the Earth's upper atmosphere. The long term stability and lack of walls in the ionosphere provide advantages over laboratory plasmas for some investigations. We shall concentrate here on three areas of research: incoherent scatter, instabilities in the equatorial and auroral E regions, and instabilities in the equatorial F region. Several other topics will be mentioned briefly. Incoherent scatter research is based upon and strongly confirms linear kinetic plasma theory. The E- and F-region instabilities lead to plasma turbulence which can be conveniently studied in the ionosphere. The linear theories of these instabilities are now at least partly understood, particularly in the long wavelength regime. The nonlinear processes by which the turbulence develops are being actively investigated via a variety of experimental, theoretical, and computer simulation programs. Once a good correspondence between the observations in space and the simulations has been achieved, the latter should guide the development of nonlinear theories, which should have applications beyond those of space physics.

1. INTRODUCTION

The ionosphere is by definition ionized and hence a plasma. It offers both advantages and disadvantages for the study of plasma processes. It has no walls, is usually in a reasonably steady state (sometimes even complete thermal equilibrium), and covers a wide region of parameter space: the temperature varies over at least an order of magnitude, the ionization density varies over 3 or 4 orders of magnitude, and the magnetic field and/or collisional processes are important for the ions and/or electrons in some regions but not in others. On the minus side, it is difficult (but not impossible) to do controlled experiments in the ionosphere. In most situations the parameters cannot be varied at will; one must settle for what nature provides. This will probably be less true in the future, however, with the advent of the Space Shuttle.

The motivation of most ionospheric research is of course to gain an understanding of the important processes controlling the behavior of the upper atmosphere. The existence of the ionosphere and its importance for radio communications was recognized shortly after the turn of the century, and even now in the age of satellite communication systems, the ionosphere continues to have a considerable impact upon communications. We wish to understand the ionosphere, however, not only because of its effect upon radio propagation, but because it is a part of our environment. It is becoming increasingly clear that the whole atmosphere, including the uppermost parts, is more tightly coupled than was realized in the past. Processes in the lower atmosphere affect the upper and vice versa. The effect of tropospheric storms can be detected at altitudes of several hundred kilometers; the effect of magnetospheric and ionospheric processes on the lower atmosphere can apparently also be appreciable. The possible relationship between weather and the solar cycle is currently attracting considerable attention.

Most of us who study the ionosphere are interested in plasma physics because the ionosphere is a plasma, and not vice versa. Here, however, we will adopt the opposite point of view and consider the ionosphere

to be a plasma which just happens to be part of the upper atmosphere. We will not concern ourselves with such things as photochemistry and the production and loss of ionization, interactions with winds and gravity waves in the neutral atmosphere, general questions of morphology, etc. Instead we will examine three particular areas in which the fields of plasma physics and ionospheric physics have closely impinged upon each other. These are (1) incoherent (sometimes called Thomson) scatter, which was first developed as an electromagnetic probing technique for the ionosphere but which has also been successfully used to study high temperature plasmas in the laboratory, (2) E region ($\sim$ 100 km altitude) instabilities observed in the equatorial and auroral ionospheres, and (3) instabilities observed in the equatorial F region ($\gtrsim$ 300 km altitude). A few other topics will also be briefly mentioned.

Incoherent scatter, in addition to its usefullness as a probing technique, has provided perhaps the best detailed quantitative verification to date of the conventional linear kinetic theory of a "warm" plasma. Agreement between theory and observation is invariably within the experimental errors, which are sometimes less than one percent. Theoretical explanations of the E- and F-region instabilities are now emerging and are based on theories originally developed to explain laboratory plasma phenomena. In the ionosphere one can study these instabilities in their fully developed "turbulent" state at leisure, so to speak, since the main characteristics of the plasma change only slowly, usually on time scales of minutes or hours.

2. INCOHERENT SCATTER

Incoherent scatter in its simplest form is simply the scattering from individual free electrons in a plasma. The fact that an electron will scatter electromagnetic radiation has of course been known for many decades, but the radar cross section of an electron is so small ($\sim 10^{-28} m^2$) that incoherent scatter (in which the signals from individual electrons are randomly distributed in phase so

that one adds powers rather than voltages to get the total signal) from the ionosphere was ignored for many years. After all, the total cross section of even a 10 km cube of ionosphere with a density of the order of 10^{12} electrons/m^3, which is a typical maximum density, is still only about 1 cm^2, which is a pretty small target at an altitude of 300 km. Furthermore, the electron velocities should apparently introduce large Doppler shifts in the returned signal, thereby broadening the bandwidth and making detection even more difficult. In 1958, however, Gordon [31] pointed out that radars of that time had reached the state in which they could rather easily detect such weak scatter, provided that one used a really large antenna, an idea which led him to the construction of the Arecibo radar in Puerto Rico. Stimulated by Gordon's idea, Bowles [9] recorded the first measurement of incoherent scatter from the ionosphere using a 5MW, 41MHz transmitter coupled to a hastily constructed 1024 dipole antenna array.

From these beginnings incoherent scatter has been developed into an exceedingly powerful tool for probing the ionosphere from the ground and, perhaps to a somewhat lesser extent, has proved useful for studying laboratory plasma phenomena. Once the first ionospheric measurements were made, it was rapidly realized that one could not ignore the effect of the ions, as Gordon had done in his first preliminary calculations. For wavelengths longer than the Debye length (the usual case in ionospheric measurements), the ions play the main role in determining the shape of the spectrum of the scattered signal. A number of theorists [16, 29, 38, 58] soon had independently developed the basic theory, using a variety of approaches to the calculation (but all arriving at the same result, happily). The calculations involve finding $\langle|\Delta N(\underline{k},\omega)|^2\rangle$, the expected mean square amplitude of density fluctuation waves with wave vector $\underline{k}$ and frequency ω. The radar frequency and geometry determine $\underline{k}$, and ω is the Doppler shift introduced by the scattering. At some point all the

theories invoke the Vlasov or Boltzmann-Vlasov equation, which most plasma physics textbooks [12] point out is an approximation, since it involves a certain smoothing out of the microscopic electric fields. The equation is then linearized in the perturbations and one is soon faced with integrals of the form

$$\int \frac{v\partial F_o(v)/\partial v}{\omega - kv}\, dv$$

where F_o is the unperturbed distribution function of the ions or electrons, and v is the component of velocity in the direction of $\underline{k}$.

For any reasonable distribution function, including a Maxwellian for a plasma in thermal equilibrium, one must deal with the pole in this integral at $v=\omega/k$. The now universally accepted way to handle this pole was first described by Landau [47] and leads to the familiar Landau damping of plasma waves. In the late 1950's and early 1960's, however, there was still some discussion as to whether or not Landau damping was a real phenomenon, and at least one theory of incoherent scatter [15] was developed on the assumption that it was not (by integrating through the pole and taking the principal part of the integral rather than using an indented contour). The results of the latter theory, while similar to the others, were sufficienty different that experimental measurements of the shape of the spectrum of the scattered signal were soon able to show that the "Landau damped" theories were correct. The relevance of incoherent scatter research to some other aspects of the development of plasma physics theory is mentioned in a review by Bauer[5].

The theory leads to an expression for the radar scattering cross section σ of the plasma which is given below, in the notation of Farley [23].

$$\frac{\sigma(\omega_o+\omega)d\omega}{Nr_e^2} = \frac{|y_e|^2 R_e(y_i) + |(T_e/T_i)y_i + ik^2\lambda_D^2|^2 R_e(y_e)}{|y_e+(T_e/T_i)y_i+ik^2\lambda_D^2|^2\,\pi\omega}\, d\omega \qquad (2\text{-}1)$$

Here T_e and T_i are the electron and ion temperatures, N is the electron density, r_e is the classical electron radius (2.82×10^{-5}m), k is $4\pi/\lambda_{radar}$ for backscatter,

λ_D is the Debye length, and y_e and y_i are electron and ion "admittance" functions that can be expressed in terms of the plasma dispersion function for Maxwellian velocity distributions when the magnetic field is unimportant. For ion mixtures y_i must be replaced by an appropriate summation. The effect of a magnetic field, collisions, relative drifts, and non-Maxwellian distributions can all be incorporated into the admittance terms. This theory shows, among other things, that when the Debye term is much smaller than unity (the usual case in ionospheric measurements) the ions play the dominant role in determining the spectrum of the scattered signal, even though it is of course the electrons that do the scattering. The Doppler width of the spectrum is usually determined by the ion rather than electron velocities and hence is more than two orders of magnitude (for oxygen ions) narrower than Gordon's [31] original prediction, making detection of the signal considerably simpler.

Since 1960 the theory of incoherent scatter has been extended by the above mentioned authors and several others to include the effect of many other parameters of the ionosphere. In many situations the original thermal equilibrium, collisionless, single ion species, non-magnetic theory is directly applicable to the ionospheric measurements, but in some it is not. The electron and ion temperatures may be different, collisions between ions and neutral particles may be important, the Earth's magnetic field may affect the scattering, more than one ion species may be present, the plasma as a whole may be moving, or the electrons and one or more species of ions may all be drifting at different mean velocities, the electrons may have a non-Maxwellian distribution, particularly a high energy "tail" caused by the production of photoelectrons, or the ions may even by non-Maxwellian, particularly in the auroral zone where the plasma drift velocities relative to the neutral particles can be very large. It seems inappropriate to list all the relevant references here; many of them can be found in reviews by Evans[18,20,21] and Bauer[5].

The theoretical work has been extremely successful. There is no unresolved disagreement between theory and experiment, and many of the theoretical results have been verified to a high degree of accuracy - better than one percent in some cases. For example, Figure 1, taken from Hagen and Hsu [37], shows a series of auto-correlation functions, which are the Fourier transforms of the Doppler power spectra of the scattered signal, measured at the Arecibo Observatory in Puerto Rico at altitudes up to 1854 km. The fitted theoretical curves, for O^+-H^+ ion mixtures, are plotted on the same axes, and in most cases the two cannot be distinguished. From such a fitting procedure the composition and temperature variations with time and altitude can be determined. Figure 2 (from Hagen and Behnke [36]) shows an example of recent high altitude wide bandwidth measurements at Arecibo which succeeded in measuring the electronic component of the scatter from the ionosphere for the first time. Even more recently, Behnke and Hagen[6] have reported the first observation of the lower hybrid or whistler resonance predicted by theory.

The theory of incoherent scatter is very rich in the sense that so many parameters of geophysical interest can affect the scattered signal, particularly the spectrum of the signal. This often makes the data somewhat difficult to analyze, but on the other hand when the analysis is carried out it yields a great deal of information that is obtainable in almost no other way. The ionized particles in the upper atmosphere also serve as tracers to provide information about ambient electric fields and the dynamics of the neutral atmosphere. The incoherent scatter measurements of temperature, drift velocity, and collision frequency are most important in this application. Incoherent scatter observations in the ionosphere complement rocket and satellite *in situ* measurements by providing very detailed monitoring of many parameters at a single geographical location over a range of altitudes extending from 80 km or below to several thousand km.

Incoherent scatter probing of laboratory plasmas is also carried out, but the technique does not seem to provide as powerful a tool in the laboratory as it

does in the ionosphere. The measurements are usually made with lasers with wavelengths much less than the Debye length (in contrast to the ionospheric case), and so the ions have no effect on the scatter. The data give a measure of the electron velocity distribution (i.e., temperature and density). The lab spectral data are generally of poorer quality than ionospheric data due to statistical considerations. Since the scattering is a stochastic process, one must average many samples to get accurate estimates of the statistical parameters, even if the signal-to-noise ratio is very high. This averaging is generally easily done in the ionospheric case where changes in the plasma characteristics are slow, but this is often not possible in the laboratory. In spite of all this, however, the lab measurements do yield results which are difficult to obtain in any other way.

To summarize, some of the important consequences of incoherent scatter research to data are:

1. The extensive agreement between theory and ionospheric observations has provided a detailed quantitative verification of conventional kinetic plasma theory based on the linearized Vlasov equation.

2. Incoherent scatter radar observations have become by far the most powerful ground-based technique for studying ionospheric behavior. Most of the important parameters of the ionospheric plasma can be measured over a wide range of altitudes, usually with good time resolution as well. Unfortunately, because of the expense there are only five observatories in the world now making these observations (with a sixth under construction in northern Scandanavia).

3. Incoherent scatter probing of high temperature plasmas in the laboratory has been useful in measuring the electron temperatures.

Looking towards the future, we can make the following predictions and recommendations:

1. Theoretical research on the theory of incoherent scatter will continue,

but at a moderate level. The theory currently seems to be in very good shape, but new questions will undoubtedly arise, particularly in connection with observations in the auroral zone and polar cap, regions in which we can expect at times to find highly non-equilibrium, non-Maxwellian ion and/or electron distributions.

2. Incoherent scatter will continue to make important contributions to our understanding of ionospheric physics, particularly in polar latitudes. It is important to maintain and if possible upgrade the few existing observatories. It would also be highly desirable to build new facilities, such as the one which has been proposed for a region near L=4 (see Evans [19]), probably in the Northeastern United States and Canada (where auxiliary receiving sites are planned). This project has been stalled in an advanced state of planning for several years now. Indian scientists are eager to establish an observatory in their country, but lack the nesessary funds. It should be emphasized that although incoherent scatter radars are certainly not cheap, they are not really expensive when compared to the cost of major satellite programs. We obviously learn a great deal from satellites, but some phenomena cannot be studied effectively by them. Tidal variations in the upper atmosphere, for example, are particularly suited to study by incoherent scatter, but several observatories (the more the better) distributed over the globe must be operating simultaneously.

3. Returning more to the area of plasma physics, incoherent scatter should continue to contribute to the theory of parametric instabilities in the ionosphere[54] driven by powerful HF transmitters. Measurements at Arecibo during the past few years have given interesting results, but the geometry there is not ideal for this sort of study. A heating facility is planned in conjunction with the new Scandanavian incoherent scatter observatory, and hopefully we should learn a great deal when these begin operation.

4. The advent of the Space Shuttle will provide opportunities for various active experiments in the ionosphere. For many of these, incoherent scatter from the ground will doubtless provide convenient diagnostics. The opinion of most incoherent scatter experts is that it does not make sense to try to do incoherent scatter observations from a satellite, even a very large one.

3. E-REGION ELECTROJET PLASMA INSTABILITIES

There are substantial ambient electric fields in the ionosphere. At middle and low latitudes these are produced by the dynamo action of neutral winds driving the charged particles across the Earth's magnetic field through collisional interactions. The neutral winds, in turn, are controlled by solar induced pressure gradients, tidal forces, coriolis effects, and even ion drag (representing the work performed by the neutral wind in generating the electric fields). In polar and auroral latitudes, the electric fields are generated primarily through interactions between the magnetosphere and the solar wind and are transmitted down into the lower ionosphere along the highly conducting magnetic field lines. These electric fields generate a worldwide system of currents which flow primarily in the E region at altitudes of the order of 100-120 km. These currents are particularly intense in two regions: (1) the magnetic equator where the magnetic field geometry leads to an unusually high conductivity perpendicular to the magnetic field, and (2) the auroral zone where the electric fields perpendicular to $\underline{B}$ are often very large. In both regions the currents lead to plasma instabilities which generate electron density perturbations which have been observed by radar and *in situ* probing. These observations have motivated a very substantial amount of theoretical work. The instabilities in the two regions have obvious similarities, but there seem to be some puzzling differences as well. Since the equatorial case seems to be the better understood at the moment, we will consider it first.

Equator. At the equator typical east-west electric fields are of the order of 10^{-3}V/m, which is fairly weak. However, at the equator the Hall current, which this field would normally produce, is vertical and inhibited, and a vertical polarization field builds up which is 10-20 times larger than the original driving field [1]. This in turn drives a horizontal Hall current which is due almost entirely to electron drift. The important region is characterized by $\nu_e << \Omega_e$, $\nu_i >> \Omega_i$, where $\nu_{e,i}$ and $\Omega_{e,i}$ are the electron and ion collision and Larmor frequencies. From another point of view, the effective east-west conductivity at the equator is approximately the Cowling conductivity $(\sigma_2^2+\sigma_1^2)/\sigma_1$ where σ_2 is the Hall conductivity, which is considerably larger in the electrojet region than the Pedersen conductivity σ_1. The resulting electron drifts are generally westward during the day and eastward at night, at velocities that are typically of the order of a few hundred m/s. The currents at night are very weak because of the low electron densities in the E region, but the electric fields and electron velocities are as large as during the day. These electron velocities are much smaller than the electron thermal velocities but can at times exceed the ion acoustic velocity, which is essentially the ion thermal velocity and is roughly 350-400 m/s for typical temperature values.

The close correlation between anomalous echoes from the E region seen at the equator with ionosondes (equatorial sporadic-E) and the strength of the electrojet current as measured with magnetometers was recognized many years ago[50]. During the IGY (1957-58) a comprehensive study of equatorial scattering phenomena in both the E and F regions was carried out using several VHF forward scatter links in Peru (see Cohen and Bowles, [13] and references therein). In the early 1960's the work in Peru was resumed using frequency coherent VHF radars so that Doppler spectra of the electrojet echoes could be determined [10]. The strong echoes showed a characteristic peak in the spectrum at a Doppler shift that corresponded closely to the ion-acoustic

velocity in the medium. This result led Farley [22] and Buneman [11] to independently suggest that a type of two-stream plasma instability was the source of the irregularities responsible for the scattering. Because the current in this case is flowing perpendicular to the magnetic field, the instability threshold is greatly reduced from that of the ordinary collisionless plasma two-stream instability, for which the relative ion-electron drift velocity must exceed the electron thermal velocity when $T_e \approx T_i$, as in the ionosphere. In the electrojet case the drift must only exceed the ion-acoustic velocity by a moderate amount to trigger unstable waves traveling perpendicular (or nearly so) to the magnetic field. There were also early suggestions by Maeda et al.[49] and Knox [46] that the gradient drift instability, first discussed by Simon [63] in connection with laboratory plasmas, might also be important for the electrojet case. These suggestions did not receive much acceptance at first because they did not seem to deal successfully with VHF data. Subsequent research, however, has shown that this instability is also important.

This early work reached some important conclusions but left many unanswered questions which are only now beginning to be resolved. Research continued during the 1960's and has accelerated during the 1970's, which have already seen the publication of over 50 papers dealing with the electrojet instabilities. Radar diagnostic techniques have been greatly improved, some rocket observations have been made, the linear theory of the instability (in both fluid and kinetic theory forms) has been extended and refined, numerous theoretical studies of various nonlinear effects have been carried out (although with only mixed success), some of the effects have been simulated in laboratory experiments, and work on computer simulations of the instabilities has begun. This latter work, in particular, seems to hold great promise for the future.

Even a brief discussion of all this work and/or a listing of all the references is not possible here. A short recent review of some of the material was given by Farley [24]. The most recent radar work in Peru and the current status of some of the theoretical problems are described in a series of papers by Fejer et al. [26-28]. Much of the theoretical work on nonlinear effects done in the last few years is discussed and referenced in papers by Lee et al. [48], Rogister and Jamin [56] and Sato [60].

The main observed characteristics of the equatorial electrojet instability are as follows:

1. There is a high correlation with the magnitude and direction of the electrojet current. The medium is unstable when the electrons flow westward during the day, or in either direction at night.

2. The density irregularities are highly aligned with the magnetic field.

3. The irregularities are observed over a wider altitude range at night than during the day.

4. Echoes are obtained even with vertically pointing radars, implying unstable plasma waves with vertical $\underline{k}$-vectors (not simply a vertical component of $\underline{k}$), even though the electrojet flows horizontally.

5. Two distinct types of irregularities (plasma waves), with different thresholds of excitation, different k-spectra, and different radar Doppler spectra are observed.

A few examples of the radar data are shown in Figures 3-5. Figure 3 (from Cohen and Bowles, [14]) shows a series of power spectra measured near noon in Peru, pointing the radar east and west from the zenith as well as vertical. The electrojet was strong during this period and the oblique spectra are good examples of "type 1" (so-called because they correspond to the strongest radar echoes at 50 MHz and hence were the first detected). The east and west spectra are nearly mirror images of each other and the spectra peak at essentially the same Doppler

shift for substantially different oblique zenith angles, in contrast to the behavior for "type 2" spectra for which the main Doppler shift is proportional to the sine of the zenith angle. Figure 4 (from Balsley and Farley [2]) shows data obtained at three radar frequencies, corresponding to plasma wavelengths of approximately 9, 3, and 1 meter, respectively. The 50 MHz data show most clearly the change from type 2 to type 1 spectra and back again as the electrojet streaming velocity first increases and then decreases. Figure 5 (from Fejer et al. [27]) shows a series of spectra obtained looking vertically at Jicamarca using a narrow beam antenna, short pulses, and short integrations in order to get good spatial and temportal resolution. The "turbulent" character of the scatter is quite apparent. The shape of the spectra can change rapidly with both altitude and time. These and other data not shown indicate that the waves seen looking vertically can switch from predominantly up-going to down-going (or vice-versa) in times at least as short at 2 secs and altitude intervals at least as small as 800m.

The linear theory of the instability has been developed by numerous authors. In the notation of Fejer et al. [27], the results of the fluid development can be written as

$$\omega_r = \underline{k}\cdot\underline{V}_d/(1+\psi) \tag{3-1}$$

$$\gamma \simeq (1+\psi)^{-1}[\psi\nu_i^{-1}(\omega_r^2-k^2C_s^2) + \nu_i\omega_r/kL_N\Omega_i] - 2\alpha N_o \tag{3-2}$$

Here, ω_r and γ are the real frequency and growth rate of the plasma wave, $\underline{V}_d$ is the electron drift velocity (the ions are essentially at rest), $\underline{k}$ is the wave vector, C_s is the ion-acoustic velocity, L_N is the electron density gradient length (positive for density increasing with height), α is the recombination coefficient, and ψ is $\nu_e\nu_i/\Omega_e\Omega_i$ for $\underline{k}\perp\underline{B}$ but increases very rapidly when $\underline{k}$ departs from perpendicularity by more than 1° or so. The first two terms on the RHS of (3-2) dominate at short wavelengths and set a drift velocity threshold for instability

slightly above the acoustic velocity (since ψ in the region of most interest is 0.2-0.4). This is the so-called two-stream instability which is thought to correspond to the type 1 irregularities. The last two terms in (3-2) are most important at long wavelengths and are thought to be mainly responsible for the type 2 irregularities; the third is the gradient-drift driving term and can be either stabilizing or destabilizing, depending upon the directions of the density gradient and drift velocity, and the last gives the chemical (recombination) damping rate, which is particularly small at night when the density is low.

On the basis of this theory one can account for characteristics 1-3 listed above, as well as some of the properties of the type 2 irregularities. One cannot explain the detailed shape of the Doppler spectra of the radar echoes or predict the amplitude of the unstable waves; both of these obviously depend upon nonlinear saturation mechanisms. Nor can vertically propagating unstable waves, for which the two driving terms go to zero, be accounted for. In particular, the Doppler spectra of the type 1 irregularities show a sharp peak at kC_s, even when $\underline{k}\cdot\underline{V}_d$ is apparently substantially larger than C_s. A perhaps even more fundamental problem is to account for the existence of two distinct types of irregulariites, since there is only one root of interest of the dispersion equation. Another question is that of anomalous resistivity: Is the conductivity of the electrojet region seriously affected when strong plasma waves are generated?

In attempts to deal with some of these questions, various nonlinar processes such as quasi-linear effects, orbit diffusion, and mode coupling, combined in some cases with the effects of refraction and convection, have been investigated. Some or all of these processes undoubtedly are important, but because of the inevitable simplifications which are introduced into the models to make them mathematically tractable, the results to date have not been very successful in explaining the observations. All the nonlinear models, for example, assume a

horizontal laminar background electron drift, and hence cannot account for the vertically propagating waves. To cope with this problem Farley and Balsley [24a] suggested a two (or more) stage process in which the medium first develops large scale irregularities with associated "turbulent" gradients and electron drifts; these is turn can generate small scale irregularities (which are what the radar detects) with wave vectors in all directions. This idea was shown to be quantitatively reasonable by Sudan et al.[64], and shortly thereafter Sato [59] independently put forward a similar suggestion.

This work stimulated interest in computer simulations of this turbulent coupling process. Numerical investigations of the electrojet processes were carried out as early as 1967 by Sato and Tsuda,[61] but in my opinion at least, the first important computer simulations were those of McDonald et al.[51,52]. These showed that the turbulent process described above really takes place; small scale irregularities moving in a wide variety of directions can be generated via the cascading process when the linear theory described by (3-1) and (3-2) predicts that only large scale waves moving primarily horizontally will be unstable. Figure 6 shows the simulated development of the irregularities, using a 50x50 point grid with a spacing of 1.5m and a drift velocity of 100 m/s. It can be seen that the initial perturbation develops into a highly chaotic state within 2.7 secs. This work which was begun at the Naval Resaerch Laboratory is now being extended at Cornell University. The early results of the Cornell simulations have already given good reproductions of radar observations of the type 2 irregularities. Figure 7 (from Keskinen et al.[44]) shows a comparison between a simulated radar Doppler spectrum for 9.5 meter wavelength waves and one of the spectra shown in Figure 4.

The importance of the simulations is that they will provide an opportunity to carry out "controlled experiments" on the ionospheric processes. Once we are convinced that the numerical results truly reproduce what is going on in the ionosphere, we can vary the plasma parameters at will and study the effects, something

that cannot be done in the real ionosphere. Furthermore, the simulations permit perfect "probing" of all the plasma variables. The catch to all this of course is that the simulations always involve approximations and often cannot completely reproduce the natural phenomena, particularly over a large range of wavelengths. Nevertheless, the equatorial electrojet observations and simulations seem to provide one of the "cleanest" opportunities available for studying in detail an example of fully developed two-dimensional plasma turbulence.

What are the prospects for the future in this area; where should future efforts be directed?

1. As far as observations are concerned, joint comprehensive experiments will be of most value. We need rocket measurements of both the zero order and perturbation electron densities and electric fields, together with simultaneous radar measurements at as many angles and wavelengths as possible. Campaigns of this sort are planned for Peru.

2. The numerical simulation work should be continued and expanded. The most recent work referred to above neglects ion inertia, for example, and thus is not valid for drift velocities comparable to the acoustic velocity, the condition for generating type 1 irregularities. The simulations also need to be extended to shorter wavelengths.

3. If successful, the simulations should serve as a guide to more realistic semi-analytic theories of the nonlinear processes. Merely simulating a natural phenomenon is not of too much use if it does not lead to an improved understanding of the important physics involved.

4. The simulations should be able to resolve the question of whether or not anomalous resistivity[53a] is important in the equatorial E region.

Auroral zone. Turning now briefly to the auroral zone E region, we find that our understanding of the plasma processes there is still in quite a primitive state. Measurements of the Doppler spectra of VHF radar signals scattered from auroral zone E region plasma waves were first obtained in the 1950's [8],

and recently a group from NOAA has been actively pursuing this sort of work [4,17,32-34] Sometimes the radar spectral data look very similar to those obtained at the magnetic equator; it seems certain that the instability mechanisms are related. The auroral effects are much more variable than those at the equator, however, and the geometry is more difficult to untangle. Some spectra do not resemble anything seen at the equator. Furthermore, in contrast to the equatorial situation, in the auroral zone it is often difficult to know exactly what the electrojet current is doing. The theoretical situation is complicated by the fact that the magnetic field is nearly vertical, which means that differing altitudes with substantially differing plasma parameters (particularly the important collision to Larmor frequency ratios) are strongly coupled together. At the equator the altitudes are uncoupled. On the other hand, the induced polarization field that is an important factor in enhancing the equatorial current (and which may therefore perhaps be involved in the nonlinear saturation mechanisms) is not important in the auroral case.

My feeling, which is not shared by everyone, is that a reasonable understanding of the E-region auroral zone plasma instabilities will only be achieved after the equatorial zone processes are more fully understood. In the meantime, though, the radar, rocket, and satellite observational programs should be continued, and combined wherever possible. These coordinated programs are even more essential in the auroral than in the equatorial case, since in the auoral medium we often have only a poor idea of the ambient conditions in which the instability is developing. All-sky photographs, satellite photographs, magnetometer data, satellite and rocket electric field and density measurements, together with radar data all contribute useful information to the total picture. Again, the question of what role anomalous resistivity [53a] plays in controlling the flow of current in the auroral zone is an important one that needs to be answered.

4. EQUATORIAL F-REGION INSTABILITIES

The fact that radio waves are often strongly scattered by irregularities in the nighttime equatorial ionosphere at altitudes of the order of 300 km or more was discussed as long ago as 1938 by Booker and Wells.[7] These irregularities also cause the unusually strong radio star and satellite signal scintillations observed at equatorial latitudes. Somewhat unexpectedly, strong fading of satellite transmissions sometimes occurs at frequencies as high as 4 and 6 GHz [65], a result that must be considered in the design of communications systems. During the 30 some years following Booker and Wells' work the morphology of these irregularities was extensively studied and the propagation phenomena (how the irregularities might produce the observed radio effects, given that the irregularities exist) were reasonably well sorted out, but little progress was made in finding the source of the irregularities. Farley et al.[25] reviewed most of the ideas that had been proposed up to that date and pointed out that none of them could explain the various observations, particularly those of VHF scatter from short wavelength (3m) irregularities over a very wide range of altitudes, including regions where the mean vertical electron density gradient was positive, negative, or zero. This insensitivity to the density gradient seemed at the time to rule out gravitational instabilities as a cause. Haerendel [3,35] pointed out that in any such theory one must really consider the total ionization within a magnetic flux tube, taking into account the curvature of the lines of force near the equator, and that this total content can have an upward gradient even above the region of maximum electron density in the F-region. Still, however, this idea could not explain the observations of irregularities as high as 1000 km or more.

Until very recently, this F-region problem seemed particularly baffling because the equatorial F region is very well described and its relatively simple basic chemistry and dynamics are well understood. There are no significant currents flowing and apparently no important sources of free energy available

except those involving the mean density gradients. How then to account for the instabilities in regions of positive, negative, and zero vertical gradient? The resolution of this apparent dilemma seems quite simple in retrospect. The basic idea is that a Rayleigh-Taylor instability first develops in the region of the sharp density gradient on the bottomside of the nighttime F region; this grows to large amplitudes and "bubbles" (i.e., regions of low density) eventually detach themselves and move upward into regions of increasing density. Unless the "bubbles" break up or are otherwise dissipated, they can continue to rise through the entire F region until they reach an altitude where their density is equal to the ambient background, an altitude which could be as high as 1000 km or more. Smaller scale irregularities are presumed to form on the sharp edges of the bubbles.

These ideas, which are now generally accepted, emerged at the 1975 Gordon Conference on Space Plasma Physics, as a result of talks by M.C. Kelley and R.F. Woodman, and they are described in publications by Kelley et al.[43] and Woodman and La Hoz[68]. Figure 8, taken from the paper of Woodman and La Hoz, schematically illustrates the mechanism, using a three-fluid model in place of a continuous distribution of densities. The middle fluid is assumed to be the heaviest, and the bottom fluid the lightest. Thus the lower interface is gravitationally unstable, but the upper interface initially stable. When the bubble reaches the upper interface, however, it is lighter than the upper fluid; hence the instability continues and the bubble continues to rise.

Recently the first numerical simulations of this Rayleigh-Taylor "bubble" mechanism have been carried out by Scannapieco and Ossakow[62], and their results, which are illustrated in Figure 9, seem to confirm the model, although the growth times shown are too long. Faster growth rates would have been obtained if the bottomside of the ambient F region had been placed at a higher altitude (as is observed), where the ion-neutral collision frequency is lower.

There were a number of new experimental observations in the 1970's that led to these new insights. The satellite observations reported by Hanson and his colleagues [39] showed that very substantial discontinuities in electron density do exist in the nighttime equatorial F region. The "holes" or "bubbles" are quite apparent in the data. Spectral analysis of the observations showed that the irregularities spanned a range of wavelengths extending at least from tens of kilometers to hundreds of meters, with roughly a k^{-2} spectral power law, suggesting some sort of turbulent decay process. New radar observing techniques developed by Woodman and La Hoz [68] and used for the 50 MHz studies in Peru provided dramatic pictures of the spatial and temporal development of the unstable region, at least for the short wavelength irregularities. An example of such a display is given in Figure 10. What is shown is echo strength obtained using a vertically directed radar and plotted as a function of time and altitude. Increasing darkness indicates increasing echo strength, on a logarithmic scale. Much of the detail unfortunately is difficult to reproduce; the original computer outputs and photographs cover about 50 dB of dynamic range. Nevertheless, the rising "plumes" of irregularities and the complicated structure associated with the development of the instability are well illustrated in the figure.

Another stimulus to the resolution of this equatorial problem has been the series of barium cloud release experiments carried out in the ionosphere over the last decade. We will not discuss these here in any detail, but the observed phenomena have similarities to the equatorial processes we have been considering. The barium clouds often develop striations that are highly aligned with the magnetic field and frequently break up into quite small scale sizes. Analytical work and numerical simulations have shown that the primary cause of the striations is a gradient-drift instability operating at long wavelengths, with subsequent decay to shorter scale sizes. The observations reported by Kelley et al.[43] included a barium cloud release, the fortuitous observation of a natural "bubble"

by the same rocket that released the barium, and radar observations that tracked the motion of the cloud and the natural irregularities.

As a result of these various recent developments, it appears that at long last the general nature of the important instability processes operating in the equatorial F region is now largely understood, but many of the details are not. In particular, the mechanism for generating irregularities with wavelengths shorter than the ion Larmor radius still needs to be worked out. Hudson and Kennel [40] have suggested that a collisional drift mode instability may be responsible, but more work needs to be done. As in the case of the E-region instabilities, more numerical simulation studies are planned which should help to guide future analytic theories. We have not even begun to try to sort out and interpret the complex velocity distributions within the unstable region revealed by the 50 MHz radar spectral measurements of Woodman and La Hoz [68] These Doppler spectra are sometimes simple and sometimes complex, and can vary rapidly with both altitude and time. In addition, an "explosive" instability with growth and decay times of tens of milliseconds is sometimes seen.

What are the next steps to be taken in resolving the many remaining problems? First of all, a series of coordinated equatorial measurements is being planned for Peru, where the large Jicamarca radar is located near a rocket range and satellite tracking facilities. The plan is to measure satellite scintillations on as many frequencies as possible (using especially the "wideband" satellite) to study the long wavelength modes, use rocket probes (and passing satellite probes, as available) to make *in situ* measurements of the background and perturbation densities and electric fields, and use the radar (perhaps with several beams) to map and study the regions of short wavelength instability. The radar will also be used as an on-line monitor to guide the timing of the rocket launches. Another idea currently in the planning state is to create an artificial "hole" or "bubble" in the nighttime F region over Peru by means of a rocket release

of H_2 or water. The development of this artificial bubble would then be monitored by radar and perhaps additional rockets. The most ambitious and comprehensive experimental program now in the planning stage is the proposed Equatorial Ionospheric Irregularity Study Satellite (EQUION). This satellite will be instrumented to measure plasma density, temperature, and composition, energetic particle fluxes, optical emissions, and electrostatic and electromagnetic waves. The electron density and temperature measurements will have unusually good spatial resolution ($\sim$3m) in order to permit quantitative comparison with 50 MHz radar observations in Peru. This satellite, if funded, should make very important contributions to our understanding of the F-region instabilities. On the theoretical front, more simulation work is needed, more effort is needed in the short wavelength regime, where now very little is really understood, and possible nonlinear saturation mechanisms need to be explored. Understanding the behavior of the k-spectrum of the irregularities is particularly important for understanding and predicting satellite scintillations and their dependence upon operating frequency (see Rufenach [57], for example).

There is an additional aspect of equatorial ionospheric research that is perhaps worth mentioning. The magnetic equator passes through developing countries (with the possible exception of Brazil), and hence research in these countries has the benefit of bringing local scientists into the mainstream of research in geophysics and plasma physics. Peruvian and Indian scientists, in particular, have made substantial contributions to the E-and F-region equatorial programs.

5. OTHER TOPICS

There are several additional topics that we should at least mention here. We have briefly alluded earlier to parametric instablities in the ionosphere induced by high power HF transmitters (see Perkins et al. [54] for a comprehensive discussion) and to instabilities observed following barium cloud releases in the ionosphere (see Goldman et al. [30] for recent work on simulating

the development of striations and for references to earlier work). Ott and Farley [53] pointed out that the anisotropic ion velocity distributions that can exist in the auroral ionosphere, due to ion-neutral charge exhange collisions combined with a large $\underline{E}x\underline{B}$ drift velocity, could be unstable to the Post-Rosenbluth [55] instability. Recent satellite observations [66] seem consistent with this prediction.

Another important area of plasma research is concerned with high latitude processes which affect the ionosphere. Since these are really magnetospheric processes they will be discussed in detail elsewhere. Within a few thousand kilometers of the Earth's surface there is a complex and exciting region of plasma activity. Extremely strong radio emission occurs there with a power output of the order of 10^9 watts, rivaling Jupiter as a radio source. The region is associated with auroral generation, and recent satellite observations and barium shaped charge releases indicate electric field components parallel to the magnetic field resulting in potential drops of many kilovolts. These releases, which have given graphic proof of the frozen-in field concept, have thus also revealed the existence of decoupling mechanisms. Field aligned currents associated with the aurora are certainly related to these potential drops and associated regions of anomalous resistivity. *In situ* ionospheric observations at high latitudes have shown that ion cyclotron waves are generated in regions of intense parallel currents [41]. Velocity shear instabilities have also been observed in the F-region near auroral arcs[42,45,67].

6. SUMMARY

The ionosphere, besides being an interesting part of our environment which plays an important role in worldwide communications, is a convenient place to study certain aspects of plasma phsyics. In the last 15 years or so, there have

been a number of fruitful interactions between ionospheric plasma research and plasma research originating in other areas. The technique of incoherent scatter has become an extremely powerful tool for ionospheric research because linear kinetic plasma theory has been able to successfully explain and interpret every aspect of the observations in terms of the ionospheric parameters. This success has of course also helped to put the kinetic theory itself on very firm quantitative grounds.

We have seen that there are a number of interesting plasma instabilities in the ionosphere which lead to the development of turbulent conditions which persist for minutes or even hours in a reasonably steady state. This persistence and the absence of wall effects make the phenomena in some respects easier to study than comparable laboratory effects. If we can achieve a full understanding of the ionospheric instabilities and resulting turbulence, there will be inevitable applications to other areas of plasma physics.

Radar and *in situ* studies of the ionospheric instabilities have until now, for the most part, been isolated from one another. Future work will hopefully involve more coordination, now that this is becoming feasible, due for example to the construction of rocket facilities in Peru. The future should see an increasing use of active experiments in the ionosphere. Some perturbing with high powered radio waves has been done, barium releases have been carried out, and some hydrogen releases are being tentatively planned, but with the advent of the Space Shuttle we should expect to see a great deal more activity on this front.

The linear theories of the primary instabilities in the equatorial E and F regions seem to be in reasonably good shape, but these can only begin to account for the observations. Recent efforts to simulate the instabilities with computers have given some very encouraging results and will be actively pursued in the future. Hopefully these will lead to a better understanding of the important nonlinear limiting processes which determine the final saturation properties of

the turbulence.

Other processes not considered in detail here, such as parametric instabilities and high latitude phenomena, particularly those associated with the aurora, also hold out exciting prospects for progress in plasma physics in the next decade.

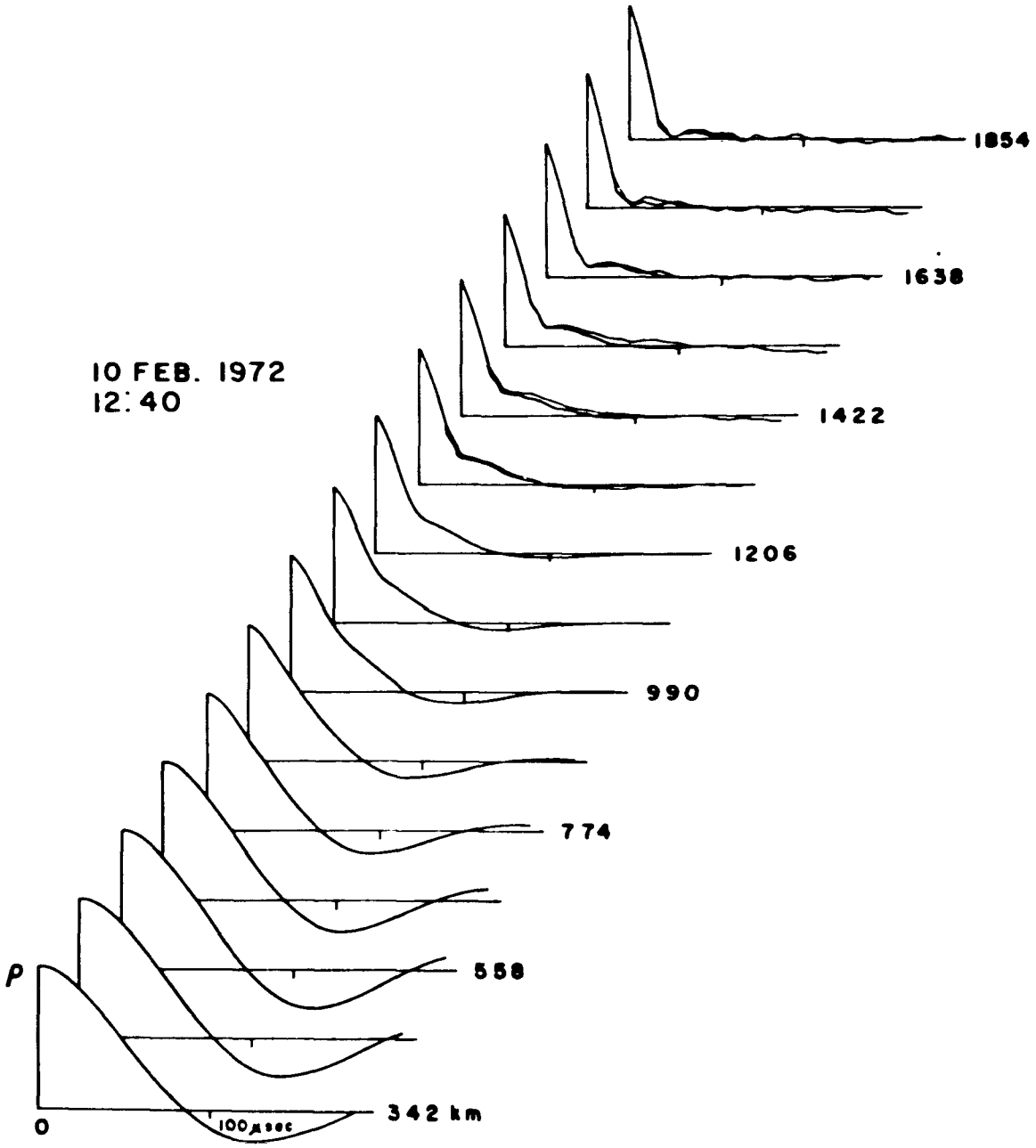


FIGURE 1 Incoherent scatter auto-correlation function measurements made at Arecibo. The pulse length used was 2 ms and the integration time was 20 min. The experimental and fitted theoretical curves are plotted together, but the agreement is so good that the two often cannot be distinguished (after Hagen and Hsu[37]).

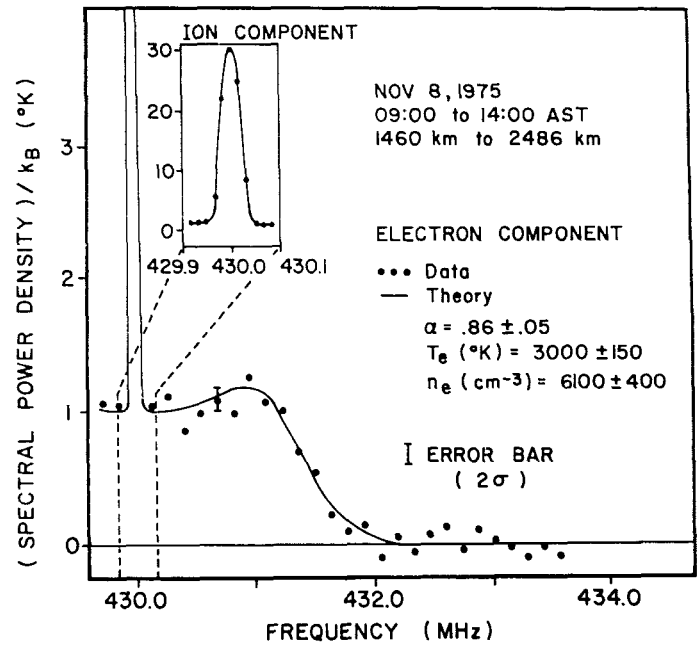


FIGURE 2 Incoherent scatter data from the ionosphere at Arecibo showing, for the first time, the electronic as well as ionic portions of the spectrum. Due to the wide bandwidth and resulting low signal-to-noise ratio, a great deal of averaging in both time and space is required (after Hagen and Behnke[36]).

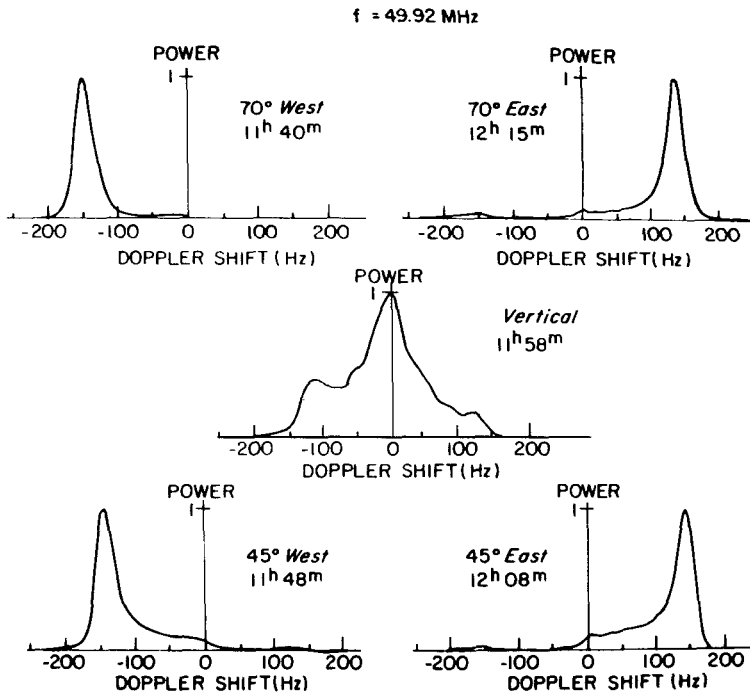


FIGURE 3 A series of power spectra of signals scattered from the equatorial electrojet near noon at Jicamarca, Peru at various angles east and west of vertical. The spectra are all normalized to unity maximum value and correspond to a relatively strong electrojet (from Cohen and Bowles[14]).

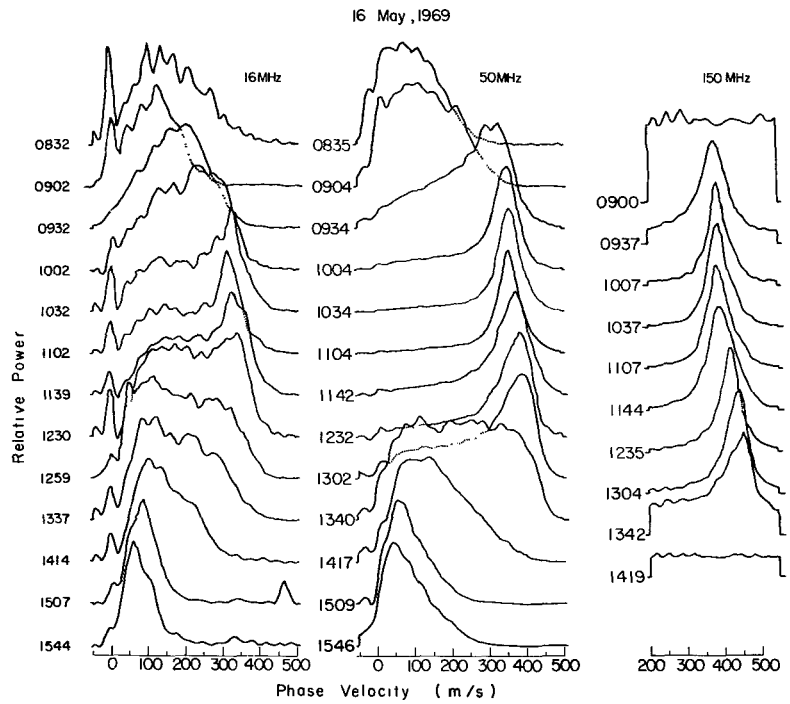


FIGURE 4 Doppler spectra from the equatorial electrojet obtained nearly simultaneously at Jicamarca at three radar frequencies. The mean local time of each 2.5 min integration is at the left of each curve. The phase velocity, rather than the Doppler shift, of the wave is plotted, and the spectra are all normalized as in Figure 3, although the maximum is not always shown. The spike that appears at the origin on several of the curves (and occasionally determines the normalization; e.g., 1416 at 146 MHz, although the spike is not shown) is due to detector bias and should be ignored (from Balsley and Farley[2]).

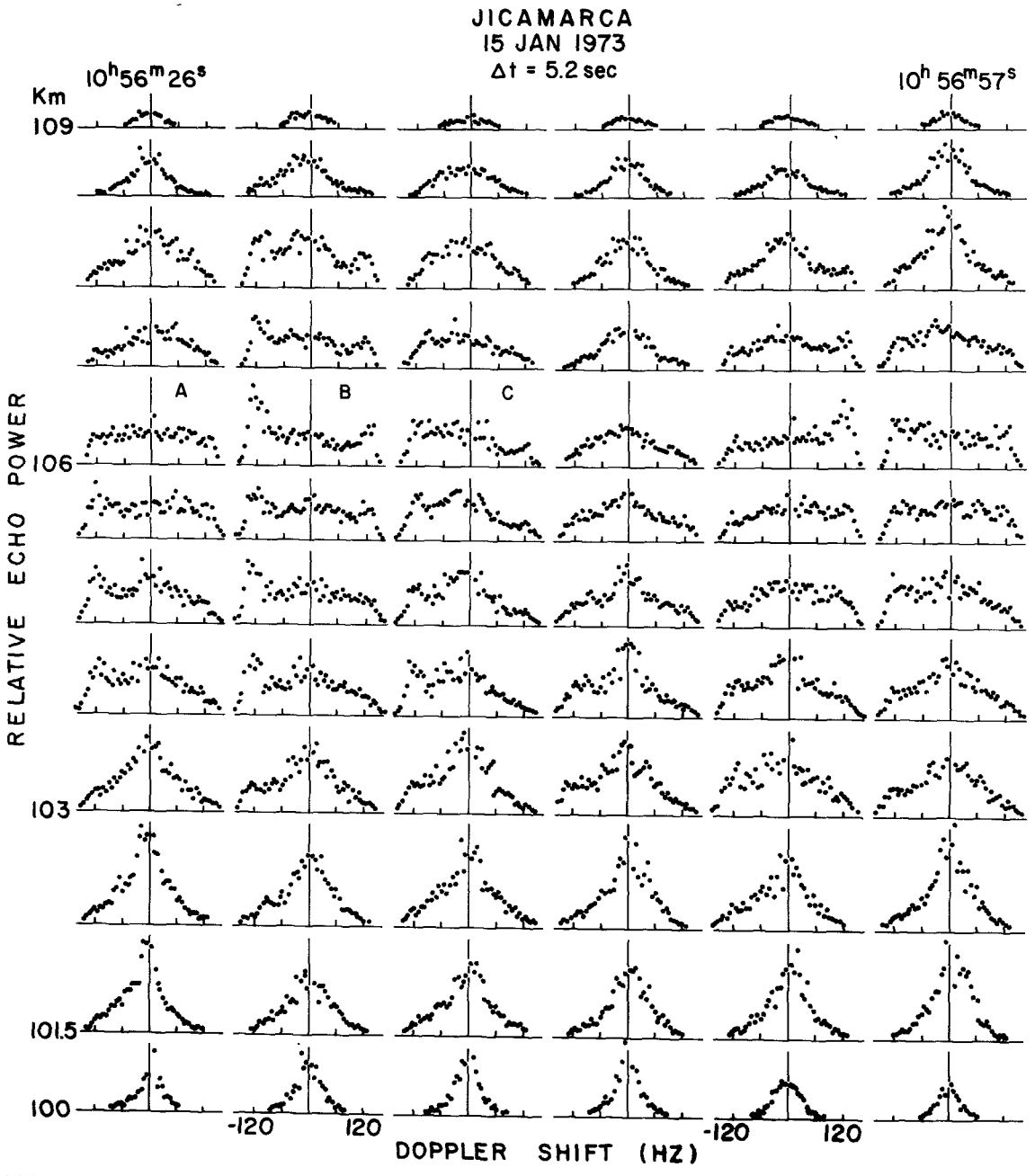


FIGURE 5 Power spectra from vertically propagating 3 m electrojet irregularities. The scattering volume extended about 1.5 km in the horizontal direction and 0.8 km in the vertical. The integration time for each spectrum was 5.2 s, corresponding to an average of 32 spectral samples (from Fejer et al.[28]).

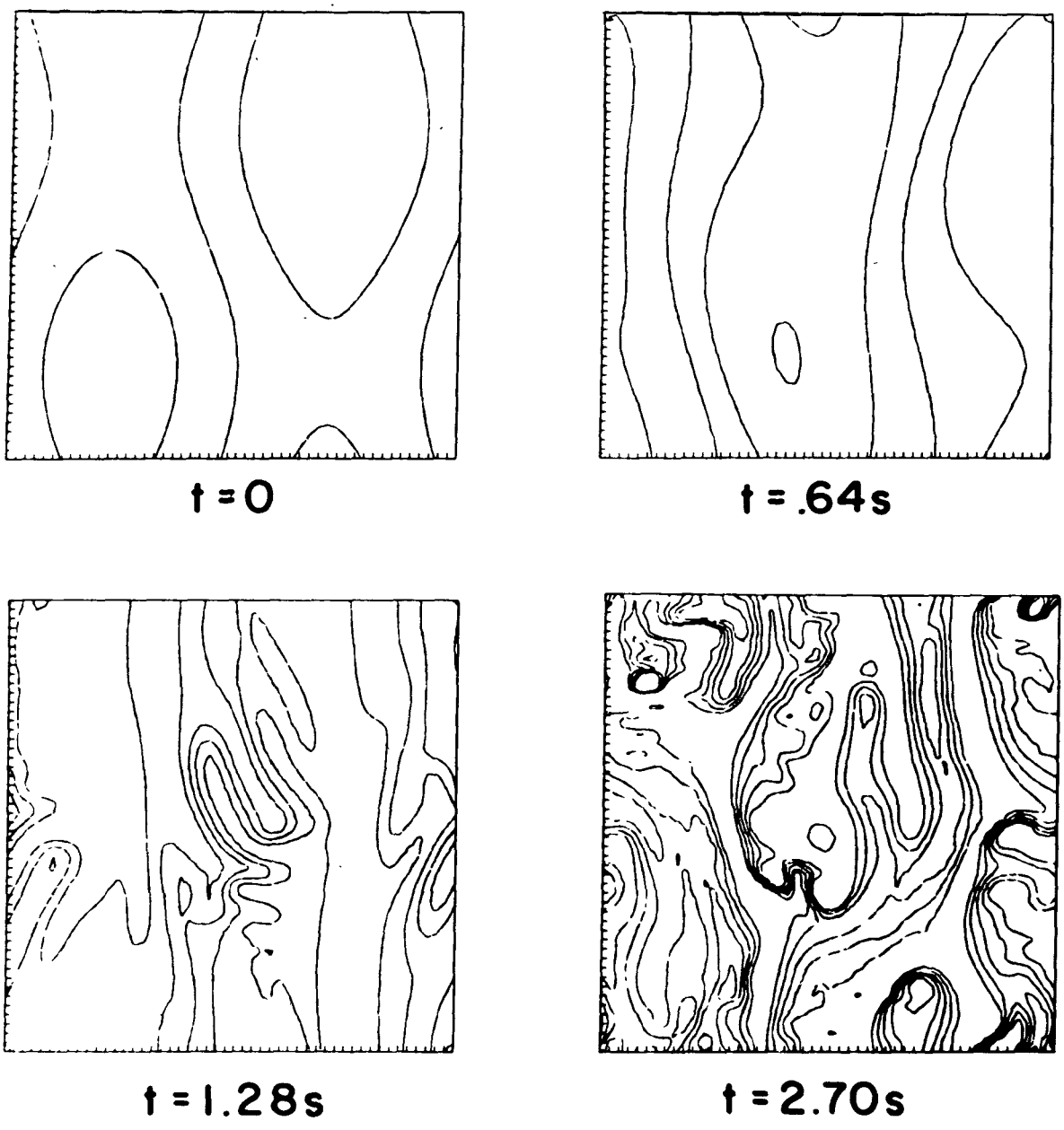


FIGURE 6 Computer simulation of the development of equatorial electrojet irregularities. The grid spacing (tic marks) is 1.5 m in both the vertical and horizontal directions and B is normal to the plane of the figure. Isodensity contours spaced by 2.5% of the ambient density are shown at four selected times (from McDonald et al.[52]).

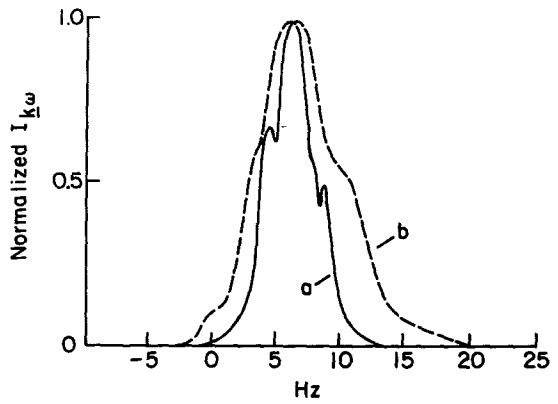


FIGURE 7 Comparison between a radar Doppler spectrum derived from a computer simulation (a) and from measurements at Jicamarca (b) similar to those shown in Figure 4 (from Keskinen et al.[44]).

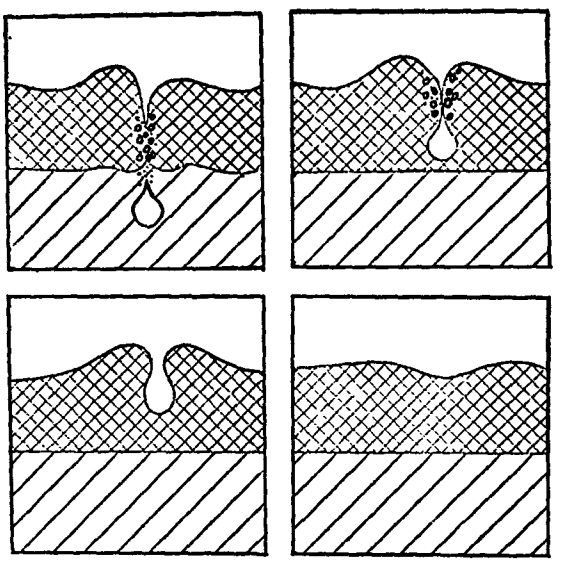


FIGURE 8 Schematic representation of a three density model of the ionosphere in which the middle layer is the heaviest and the bottom layer the lightest. The bubble forms at the lower interface and propagates through the top layers (from Woodman and La Hoz[68]).

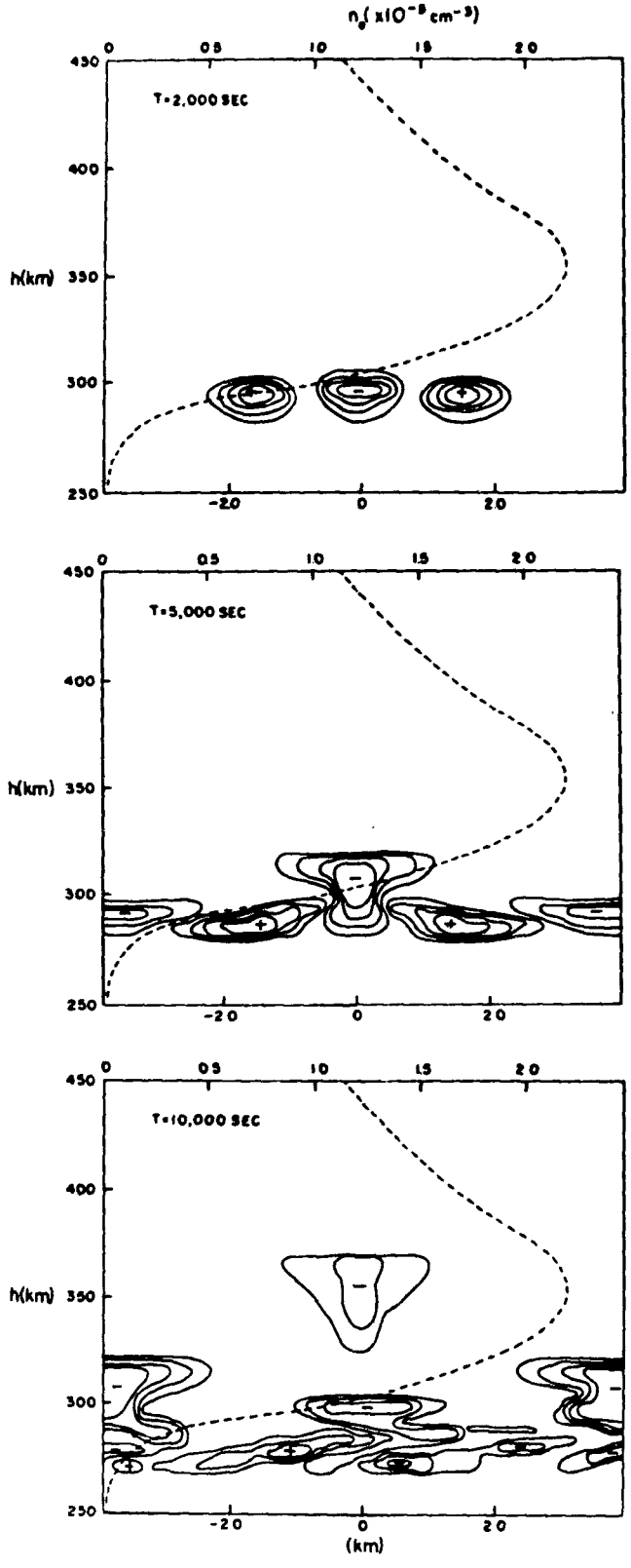


FIGURE 9 Contour plots of $\delta n/n_o$ showing the simulation of the growth of the equatorial F region Rayleigh-Taylor instability and the rising of a "bubble." The growth rate here is unrealistically small, due to the choice of initial density profile, shown by the dashed line (after Scannapieco and Ossakow[62]).

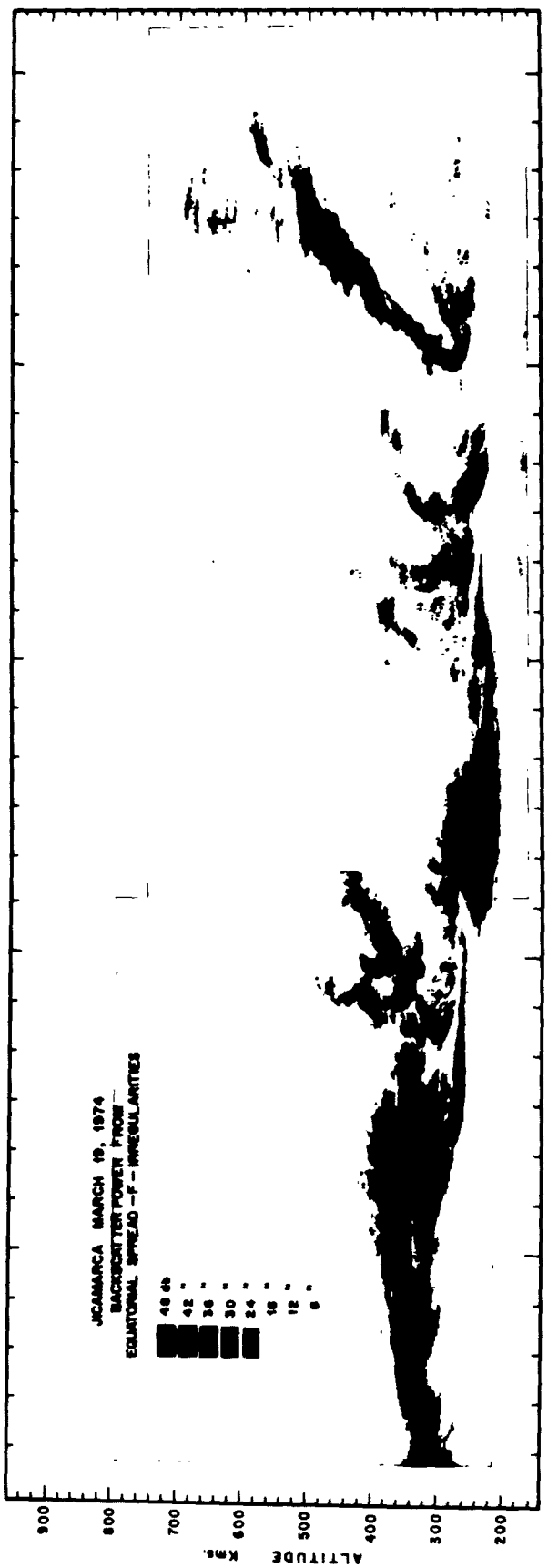


FIGURE 10 Radar echo power returned from 3-meter equatorial F region irregularities as a function of altitude and time (from Woodman and La Hoz[68]).

REFERENCES

1. Baker, W.G., and D.F. Martyn, Electric currents in the ionosphere. I. Conductivities, Phil. Trans. R. Soc. London A, 246, 281-294, 1953.

2. Balsley, B.B., and D.T. Farley, Radar studies of the equatorial electrojet at three frequencies, J. Geophys. Res., 76, 8341-8351, 1971.

3. Balsley, B.B., G. Haerendel, and R.A. Greenwald, Equatorial spread F: Recent observations and a new interpretation, J. Geophys. Res., 77, 5625-5628, 1972.

4. Balsley, B.B., W.L. Ecklund, and R.A. Greenwald, VHF Doppler spectra of radar echoes associated with a visual auroral form: Observations and implications, J. Geophys. Res., 78, 1681-1687, 1973.

5. Bauer, P., Theory of waves incoherently scattered, Phil. Trans. R. Soc. Lond. A, 280, 167-191, 1975.

6. Behnke, R.A., and J.B. Hagen, Incoherent scattering of radio waves by whistler mode oscillations in the ionosphere, submitted to Radio Sci., 1977.

7. Booker, H.G., and H.W. Wells, Scattering of radio waves by the F region of the ionosphere, J. Geophys. Res., 43, 249, 1938.

8. Bowles, K.L., Doppler shifted radio echoes from aurora, J. Geophys. Res., 59, 553-555, 1954.

9. Bowles, K.L., Observations of vertical incidence scatter from the ionosphere at 41 Mc/sec, Phys. Rev. Lett., 1, 454-455, 1958.

10. Bowles, K.L., B.B. Balsley, and R. Cohen, Field-aligned E-region irregularities identified with acoustic plasma waves, J. Geophys. Res., 68, 2485-2501, 1963.

11. Buneman, O., Excitation of field aligned sound waves by electron streams, Phys. Rev. Lett., 10, 285-287, 1963.

12. Clemmow, P.C., and J.P. Dougherty, Electrodynamics of Particles and Plasmas, Addison-Wesley, 1969.

13. Cohen, R., and K.L. Bowles, Ionospheric VHF scattering near the magnetic equator during the International Geophysical Year, J. Res. NBS, 67D, 459-480, 1963.

14. Cohen, R., and K.L. Bowles, Secondary irregularities in the equatorial electrojet, J. Geophys. Res., 72, 885-894, 1967.

15. Denisse, J.F., C.R. Acad. Sc. Paris, 253, 1539, 1961.

16. Dougherty, J.P., and D.T. Farley, A theory of incoherent scattering of radio waves by a plasma, Proc. Roy. Soc. (London), A259, 79-99, 1960.

17. Ecklund, W.L., B.B. Balsley, and R.A. Greenwald, Crossed beam measurements of the diffuse radar aurora, J. Geophys. Res., 80, 1805-1809, 1975.

18. Evans, J.V., Theory and practice of ionosphere study by Thomson scatter radar, Proc. IEEE, 57, 496-530, 1969.

19. Evans, J.V., The upper atmosphere observatory, Science, 176, 463-473, 1972.

20. Evans, J.V., Some post-war developments in ground based radiowave sounding of the ionosphere, J. Atmos. Terrest. Phys., 36, 2185-2234, 1974.

21. Evans, J.V., High power radar studies of the ionosphere, Proc. IEEE, 63, 1636-1650 1975.

22. Farley, D.T., A. plasma instability resulting in field-aligned irregularities in the ionosphere, J. Geophys. Res., 68, 6083-6097, 1963.

23. Farley, D.T., Radio wave scattering from the ionosphere, Ch. 14 in Methods of Experimental Physics (R.H. Lovberg and H.R. Griem, eds.) 9B, Academic Press, 1971.

24. Farley, D.T., Irregularities in the equatorial ionosphere: The Berkner Symposium, Rev. Geophys. Space Phys., 12, 285-289, 1974.

24a. Farley, D.T., and B.B. Balsley, Instabilities in the equatorial electrojet, J. Geophys. Res., 78, 227-239, 1973.

25. Farley, D.T., B.B. Balsley, R.F. Woodman, and J.P. McClure, Equatorial spread F; Implications of VHF radar observations, J. Geophys. Res., 75, 7199-7216, 1970

26. Fejer, B.G., D.T. Farley, B.B. Balsley, and R.F. Woodman, Oblique VHF radar spectral studies of the equatorial electrojet, J. Geophys. Res., 80, 1307-1312, 1975.

27. Fejer, B.G., D.T. Farley, B.B. Balsley, and R.F. Woodman, Vertical structure of the VHF backscattering region in the equatorial electrojet and the gradient-drift instability, J. Geophys. Res., 80, 1313-1324, 1975.

28. Fejer, B.G., D.T. Farley, B.B. Balsley, and R.F. Woodman, Radar observations of two-dimensional turbulence in the equatorial electrojet, 2, J. Geophys. Res., 81, 130-134, 1976.

29. Fejer, J., Scattering of radio waves by an ionized gas in thermal equilibrium, Can. J. Phys., 38, 1114-1133, 1960.

30. Goldman, S.R., L. Baker, S.L. Ossakow, and A.J. Scannapieco, Striation formation associated with barium clouds in an inhomogeneous ionosphere, J. Geophys. Res., 81, 5097-5113, 1976.

31. Gordon, W.E., Incoherent scattering of radio waves by free electrons with application to space exploration by radar, Proc. IRE, 46, 1824-1829, 1958.

32. Greenwald, R.A., W.L. Ecklund, and B.B. Balsley, Auroral currents, irregularities, and luminosity, J. Geophys. Res., 78, 8193-8203, 1973.

33. Greenwald, R.A., W.L. Ecklund, and B.B. Balsley, Diffuse radar aurora: Spectral observations of non-two-stream irregularities, J. Geophys. Res., 80, 131-139, 1975.

34. Greenwald, R.A., W.L. Ecklund, and B.B. Balsley, Radar observations of auroral electrojet currents, J. Geophys. Res., 80, 3635-3641, 1975.

35. Haerendel, G., Theory of equatorial spread F, preprint, Max-Planck-Institute fur Physik und Astrophysik, 1974.

36. Hagen, J.B., and R.A. Behnke, Detection of the electron component of the spectrum in incoherent scatter of radio waves by the ionosphere, J. Geophys. Res., 81, 3441-3443, 1976.

37. Hagen, J.B., and P.Y. Hsu, The structure of the protonosphere above Arecibo, J. Geophys., Res., 79, 4269-4275, 1974.

38. Hagfors, T., Density fluctuation in a plasma in a magnetic field with applications to the ionosphere, J. Geophys. Res., 66, 1699-1712, 1961.

39. Hanson, W.B., J.P. McClure, and D.L. Sterling, On the cause of equatorial spread F, J. Geophys. Res., 78, 2353-2356, 1973.

40. Hudson, M.K., and C.F. Kennel, Linear theory of equatorial spread F, J. Geophys. Res., 80, 4581-4590, 1975.

41. Kelley, M.C., E.A. Bering, and F.S. Mozer, Evidence that the electrostatic ion cyclotron instability is saturated by ion heating, Phys. Fluids, 18, 1590, 1975.

42. Kelley, M.C. and C.W. Carlson, Observations of intense velocity shear and associated electrostatic waves near an auroral arc, J. Geophys. Res., 82, 2343-2348, 1977.

43. Kelley, M.C., G. Haerendel, H. Kappler, A. Valenzuela, B.B. Balsley, D.A. Carter, W.L. Ecklund, C.W. Carlson, B. Hausler, and R. Torbert, Evidence for a Rayleigh-Taylor type instability and upwelling of depleted density regions during equatorial spread F, Geophys. Res. Lett., 3, 448-450, 1976.

44. Keskinen, M., R.L. Ferch, and R.N. Sudan, Power spectrum studies of numerical simulation of ionospheric gradient drift turbulence, paper presented at meeting of American Geophysical Union, June 1977.

45. Kinter, P.M., Observations of velocity shear driven plasma turbulence, J. Geophys. Res., 81, 5114-5122, 1976.

46. Knox, F.B., A contribution to the theory of the production of field-aligned ionization irregularities in the equatorial electrojet, J. Atmos. Terrest. Phys., 26, 239-249, 1964.

47. Landau, L.D., On the vibrations of the electronic plasma, J. Phys. USSR, 10, 25-34, 1946.

48. Lee, K., C.F. Kennel, and F.V. Coroniti, On the marginally stable saturation spectrum of unstable type 1 equatorial electrojet irregularities, J. Geophys. Res., 79, 249-266, 1974.

49. Maeda, K., T. Tsuda, and H. Maeda, Theoretical interpretation of the equatorial sporadic E layers, Rpt. Ionos. Sp. Res. Japan, 17, 147-159, 1963.

50. Matsushita, S., Intense E_s ionization near the magnetic equator, J. Geomag. Geo., 3, 44-46, 1951.

51. McDonald, B.E., T.P. Coffey, S. Ossakow, and R.N. Sudan, Preliminary report of numerical simulation of type 2 irregularities in the equatorial electrojet, J. Geophys. Res., 79, 2551-2554, 1974.

52. McDonald, B.E., T.P. Coffey, S.L. Ossakow, and R.N. Sudan, Numerical studies of type 2 equatorial electrojet irregularity development, Radio Sci., 10, 247-254, 1975.

53. Ott, E., and D.T. Farley, Microinstabilities and the production of short-wavelength irregularities in the auroral F region, J. Geophys. Res., 80, 4599-4602, 1975.

53a. Papadopoulis, K., A review of anomalous resistivity for the ionosphere, Rev. Geophys. Space Phys., 15, 113-127, 1977.

54. Perkins, F.W., C. Oberman, and E.J. Valeo, Parametric instabilities and ionospheric modification, J. Geophys. Res., 79, 1478-1496, 1974.

55. Post, R.F., and M.N. Rosenbluth, Electrostatic instabilities in finite mirror confined plasmas, Phys. Fluids, 9, 730-749, 1966.

56. Rogister, A., and E. Jamin, Two-dimensional nonlinear processes associated with "type 1" irregularities in the equatorial electrojet, J. Geophys. Res., 80, 1820-1828, 1975.

57. Rufenach, C.L., Ionospheric scintillation by a random phase screen: Spectral approach, Radio Sci., 10, 155-165, 1975.

58. Salpeter, E.E., Electron density fluctuation in a plasma, Phys. Rev., 120, 1528-1535, 1960.

59. Sato, T., Unified theory of type 1 and 2 irregularities in the equatorial electrojet, J. Geophys. Res., 78, 2232-2243, 1973.

60. Sato, T., On mechanisms governing the electrojet plasma instabilities, J. Geophys. Res., 81, 539-546, 1976.

61. Sato, T., and T. Tsuda, Computer study on nonlinear cross-field instability, Phys. Fluids, 10, 1262-1268, 1967.

62. Scannapeico, A.J., and S.L. Ossakow, Nonlinear equatorial spread F, Geophys. Res. Letts., 8, 451-454, 1976.

63. Simon, A., Instability of a partially ionized plasma in crossed electric and magnetic fields, Phys. Fluids, 6, 382-388, 1963.

64. Sudan, R.N., J. Akinrimisi, and D.T. Farley, Generation of small-scale irregularities in the equatorial electrojet, J. Geophys. Res., 78, 240-248, 1973.

65. Taur, R.R., Ionospheric scintillations at 4 and 6 GHz, Comsat Tech Rev., 3, 145-163, 1973.

66. Ungstrup, E., Observation of the Post-Rosenbluth instability in the polar cap F region, paper presented at meeting of the American Geophysical Union, December 1976.

67. Webster, H.F., and T.J. Hallinan, Instabilities in charge sheets and their possible occurrence in aurora, Radio Sci., 8, 475, 1973.

68. Woodman, R.F. and C. La Hoz, Radar observations of F region equatorial irregularities, J. Geophys. Res., 81, 5447-5466, 1976.

CONTRIBUTION TO NATIONAL ACADEMY OF SCIENCES
SPACE PLASMA PHYSICS STUDY.

UNDERSTANDING PLASMA INSTABILITIES IN SPACE: IONOSPHERIC RESEARCH AND COMMUNICATIONS APPLICATIONS

F. W. Perkins
National Center for Atmospheric Research*

Instabilities are ubiquitous in space plasmas because they ofttimes provide the dissipation and transport which are demanded by the general large-scale features of plasma flow in the ionosphere and magnetosphere. Indeed, dissipation and transport by instabilities is a feature which space plasmas share with other large, nonlinear systems; turbulent hydrodynamic flows, geophysical fluid dynamics, astrophysical hydrodynamics and controlled fusion research. One of the major scientific challenges of the present era is to develop techniques to understand these large, nonlinear systems, especially the role of instabilities. The goal of this chapter is to set forth an approach to space physics which will not only provide an overall understanding of space plasmas, but also, in this process, take up problems of true intellectual significance in just how instabilities provide transport and dissipation in large scale flows. We shall point out how two-dimensional ionospheric plasma instabilities have both important similarities and differences with turbulence in neutral fluids. The interplay between the similarities and differences serves to increase our understanding of both the plasma and neutral fluid instabilities.

*The National Center for Atmospheric Research is sponsored by the National Science Foundation.

It is also evident that a good understanding of the ionosphere is required if a model for the climatological effects of solar activity is ever to be devised, for it is in the ionosphere that the energy of solar disturbances affects neutral atmosphere flows and chemistry. Lastly, we shall discuss how ionospheric instabilities affect various communications schemes.

1. A Computational Framework for Understanding Space Plasmas

The first issue to be addressed is what are the goals of Space Plasma Physics? There is almost general agreement, I believe, that the era of space exploration has past, except for planetary studies. Already, an impressive empirical knowledge[2] has been obtained which is adequate to describe the environment in which future space missions will fly, although the complicated interaction between a given environment and spacecraft electronics, etc. is still poorly understood. A new goal is required, and in my view, the new goal of Space Plasma Physics should be understanding.

What form does one expect that understanding will take? I take the point of view that true understanding is expressed by theory which allows us to predict the consequences of various applied forces. But in nonlinear systems, one should not expect understanding to take the form of simple equations, like the Schrodinger Equation. A more likely prospect is that understanding of nonlinear, dynamical systems will be embodied in a sequence of sophisticated computer codes, each of which deals with phenomena on particular spatial and temporal scales, with the results of small-scale codes parameterized in the large-scale codes. Hence

the nonlinear effects of instabilities will appear as viscosities, resistivities, etc. in large-scale codes. By turning on and off the parameterization of an instability in a large-scale code, one can study how the sources of free energy develop which drive the instability in the first place.

This brings us to another advantage of the computational approach to understanding space plasmas--identifying the driving mechanisms for plasma instabilities. In general, the growth time for instabilities is very short compared to the time-scales for evolution of the large-scale flow. Hence the instabilities are always observed in a nonlinear steady state, and linear stability calculations based on measured data, if done correctly, should always yield zero growth rate. Most of the stability calculations found in the space physics literature are based on measured data, and hence should not show strong destabilizing terms. But a computational program, which embodies only limited physics, can evolve into regimes where small-scale processes, not described in the program, go unstable. Thus one can identify what aspect of the large-scale motion drives small-scale instabilities, and how strong these driving terms are. An excellent example of this type of approach is the recent work of Liewer and Krall on collisionless shocks.[35]

Experiments, of course, play a major role in this approach to understanding space plasmas. First, they provide the initial data with which to start computational programs. The limited data combined with the limited capability of computers guarantee that one can not "put the magnetosphere on a computer" and expect an accurate detailed description

of the magnetospheric response. But one can hope for a much more quantitative cause-and-effect understanding. The second role of experiments is to normalize the computational programs to actual observations. In nonlinear systems, we can expect many sub-grid scale processes to be going on which must be parameterized, and it is only through comparison with observations that the success of the parameterization process can be determined. Conversely, attempts to normalize computational programs to observations will channel the observational efforts on the key issues where increased experimental sophistication will directly help achieve our goal of understanding. Lastly, experiments should continue to have the ability to discover unexpected process, because we can certainly expect that in building a computational understanding there will be important processes overlooked in the first attempts. The recent discovery of electrostatic shocks[40] is a case in point.

Calling for more theory and computational models has almost reached a motherhood status in space physics recently. But, all too often, theoretical programs can devote much time to relatively straightforward research and bypass the difficult problems of true intellectual significance. The bulk of this chapter is devoted to describing how ionospheric research has in the past, and can in the future, progress towards a true understanding, taking up the difficult problems as they become germane, and developing the detailed ionospheric models which will be required in future climatological and magnetospheric studies.

To summarize the above paragraphs: Space Plasma Physics should set a goal of understanding and the most likely form this understanding will take is a set of nonlinear dynamic computational programs. Experiments provide three essential contributions: identifying the basic physical processes, providing initial data, and normalizing computational results. Experimental programs which entail the broadest set of measure-

ment techniques should be emphasized, each geared towards a key plasma physics problem.

2. The Impact of Space Plasma Physics

The scientific challenge provided by the complicated nonlinear space plasma system is one of the major reasons for undertaking studies with ambitious goals. Methods developed for treating plasma turbulence, nonlinear viscosities and resistivities, magnetic field merging, etc. will find application in many other fields of research and engineering. Achieving this impact, however, implies an aggressive program in space plasma theory with cross-ties to the other fields such as hydrodynamics, meteorology, oceanography, controlled fusion, etc. Currently, these other fields are providing the leadership in developing new theoretical methods and concepts.

Communications systems which propagate through space plasmas are used increasingly for civilian and military purposes.[63] Let us take a broad view of communications to include all forms of electromagnetic transmission of information: standard messages relayed via satellite, TV, satellite telemetry and other data, navigational signals, space track and ABM radars, and even raw energy from a satellite power station.[9] Since most of these systems are already in use, the influence of space plasmas is felt on a day-to-day basis. We shall amplify this discussion later.

A recent and discouraging development is that space might become an arena of "conventional" warfare because of the large number of space-based military systems.[58] Hence high-energy systems may be found in space, and if their performance depends on interaction with the space environment, a good theoretical understanding of space plasmas is called for to assess their chances of success. At the least, some sort of environmental impact will occur in systems tests.

The role of the ionosphere as the site of the interaction between solar activity energy and the neutral atmosphere meteorology has already been stressed.

3. Understanding of the Ionosphere

Let us turn to the question: How far has our understanding of the ionosphere and ionospheric plasma instabilities in particular advanced? The answer to this question will be given within the framework set forth in the preceding paragraphs. Evidently there is a strong theoretical bias in this approach and it does the disservice of minimizing the importance of experimental results. Hence, it should be clearly stated that the theoretical papers referred below all give careful attention to the experimental data which, in many cases, provided the theoretical puzzles in the first place. The references cited in the theoretical papers provide a good guide to the experimental literature.

The ionosphere is perhaps the best subject for the computational approach to understanding space plasmas, principally because the magnetic field is very strong and magnetic perturbations can be safely ignored in ionospheric dynamics. Our framework for understanding the ionosphere contains four parts: (1) studies of ionospheric equilibria, (2) the stability at these equilibria, (3) nonlinear calculations of the saturation of instabilities, and (4) parameterization of the results of instability calculations in the large-scale equilibria studies. The reader should appreciate that all four parts are required in fields such as meteorology, communications channel modeling, and controlled fusion which demand a predictive capability for theory, not merely explanations. Let us examine how far understanding has progressed in each of these four areas.

3a. Ionospheric Equilibria. Of all the topics within space plasma physics, the "equilibria" of the ionosphere has perhaps the most advanced understanding. In this context, equilibria means states that change only on global dimensions in space and diurnal periods in time. One-dimensional computational models[5,28,37,46,50,59,60] which calculate the evolution of the ionosphere as a function of height for a particular geographical coordinate have become quite sophisticated. Photoelectron transport, slowing down, airglow excitation, energetic electron precipitation,[59] and molecular and atomic reaction rates are all included. Nonetheless, certain external forces must be prescribed to have a well-posed model: neutral winds which can support (or push down) the F-region as well as create sporadic E-layers,[37] the solar EUV flux which provides the ionization, magnetospheric fluxes of plasma, heat, precipitating electrons, and electric current and the electric fields which these fluxes generate. In high latitude ionospheres, joule heating by electric currents plays an important role.[8] Resonantly scattered solar Lyman-α radiation is an essential element in nighttime ionosphere models.[60]

Within reasonable adjustments of these external parameters, the calculated ionospheric electron density profiles agree well with observed ones.[19,50] But often achieving good agreement forces one into adopting solar EUV fluxes that are twice the measured values,[28,50] perhaps an indication of our uncertainty of their variation during the course of the solar cycle. These models are quite good at examining the response of the ionosphere to variations in the global neutral wind patterns and electric fields.[28] Three-dimensional models, which are essential to achieving an understanding of the polar ionosphere where horizontal convection is important, have only reached a state of dimensional analysis.[33] But one-dimensional models which allow the possibility of plasma outflow, called the polar wind, are sophisticated.[6]

3b. Stability of the Ionosphere. Although a careful reader can readily identify many areas in which improvements of the equilibria models is desirable, the thrust of this chapter is to ask: to what extent have these models been examined for stability with respect to small-scale perturbations? Let us begin our answer by enumerating what types of equilibria are employed in the current models. We find (1) mechanical equilibrium, (2) thermal equilibrium, (3) chemical equilibrium, and (4) electrical equilibrium (i.e. the divergence of the current is zero).

It appears as though there has been very little systematic effort to investigate the stability of the many ionospheric equilibria models in spite of experimental evidence for small-scale structures. Instead, most instability work has concentrated on a few very short-scale (1-10 meter) instabilities discovered by radar backscatter measurements. Structure on the kilometer scale length which is studied by satellites,[1,4,16,17,30,54,55] auroral photographs and radiowave propagation studies (e.g., radio star scintillation) has received less theoretical attention, except for the equatorial ionosphere.

Let us set forth some more specific suggestions on what types of instabilities might occur in the ionosphere and indicate whether or not any stability analyses have been performed. Turning first to the question of mechanical equilibrium of the ionosphere, electric fields and neutral winds play an important role in controlling the height of the F-layer and in producing sporadic-E layers. Near the equator, the geometry is appropriate for a classic $\underline{E} \times \underline{B}$/Rayleigh-Taylor instability on the bottomside ionosphere and stability analysis clearly predict instability. This instability is invoked as the cause of equatorial spread-F.[30,56] Mid-latitude nighttime F-regions are also unstable to support by neutral

winds although the growth rates are small unless the layer is high.[42] So far no stability analysis of sporadic-E layers has been performed in spite of history of experimental data which indicate small-scale structures.[27] Sporadic-E layers are interesting because they involve a substantial variation in the chemical equilibrium from the background ionosphere. The Farley-Buneman[20] acoustic instability has long been recognized as occurring in both the equatorial and auroral electrojet; its cause is the high-current required by the electrical equilibrium. Ion-cyclotron instabilities which can be generated by parallel currents[31] or ion beams[44] have recently been detected by rocket experiments.[29] The source of driving free energy is not yet clear, however. Auroral arc models[5], which have very fast chemistry time scales as well as important electric currents, have yet to be examined for instabilities other than the Farley[20,22] instability. The gradient-drift instability has been proposed[26] to explain diffuse radio aurora but a stability analysis which deals satisfactorily with the high-latitude geometry remains to be carried out. The supersonic flow of plasma through a neutral gas which occurs in high-latitude ionospheres can generate microinstabilities if ion-neutral collisions are sufficiently frequent.[41] The auroral zone F-region might well suffer this instability, but no experimental tests have yet been carried out. To my knowledge, there are no attempts to assess the thermal stability of a plasma where joule heating by either perpendicular ion currents or parallel electron currents plays an important role. For example, one can speculate that the very small-scale instabilities[4,21] observed in the equatorial F-region could be caused by a thermal instability produced by field-aligned currents flowing between hemispheres.

Overall, most of the instability calculations have been motivated

by observations, and not by the question of whether equilibria models are stable. Even so, there remains the recently discovered shocks[40] and an impressive amount of ionospheric structure, especially at high-latitudes and short scales,[39] that are still unexplained even in a linear stability analysis.

For completeness one should also mention that man-made perturbations of the ionosphere such as barium clouds, satellite electron beam generators, radiowave modification, nuclear explosions, etc. have initiated many instabilities in the ionosphere. We shall take these up in a special discussion below.

3c. Nonlinear Saturation of Instabilities: If the stability analysis are lagging equilibrium models and observations, then nonlinear studies of the identified instabilities is still more seriously behind. Essentially only equatorial Farley and gradient-drift instabilities have been studied. Perhaps the most successful nonlinear studies to date have been the application of the computational simulation method first developed for barium clouds,[66] to gradient-drift and Rayleigh-Taylor instabilities in the equatorial E- and F-regions.[38,56] In the E-region, the computational result[38] showed how daughter instabilities with vertically-oriented wave vectors were created by the nonlinear state of the primary instability which had an approximately horizontal wave vector. The results provide an understanding of radar and rocket data. The F-region studies[56] show how large-scale density irregularities can occur and rise rapidly into the initially stable region in the post-sunset equatorial F-region. Again good agreement with experiment is reported.

The Farley acoustic instability has been studied only by analytic methods even though its two-dimensional nature suggests that computational simulations would be straightforward. The most advanced analytic theories[34,52] combine a number of previous approaches employing quasi-

linear effects[51] or strong turbulence models,[64] and nonuniform geometry. While good agreement with experiment is claimed, it seems to me that a uniform medium instability should have a uniform medium saturation mechanism and that computational simulation is a straightforward way to determine it.

From the point of view of fundamental studies, the ionosphere provides a unique opportunity to study strong two-dimensional turbulence. The equations governing the evolution of two-dimensional structures bear a close resemblence to hydrodynamic equations.[66] In these analogies, field-line-integrated plasma density or conductivity takes on the role of vorticity while the electric potential serves as the stream function. The ionospheric turbulence shares with two-dimensional fluid turbulence the property that all moments of the vorticity are conserved, but the elliptic equation for the stream function is sufficiently different that steady-state vortices can not occur in ionospheric turbulence. The ionosphere has the advantage over neutral fluids in that turbulence remains two-dimensional so that the interplay between observations, computer studies,[38,56] and analytic closure theories[52,53,64] can be quite close.

3d. Influence of Instabilities on Large-Scale Phenomena. Here there are only a few scattered results. A paper by Fedder[23] surveys what effects anomalous resistivity parallel to a magnetic field might have on auroral arc models. Several resistivity models were required, however, since the form of this resistivity is not yet clear. A nonlinear resistivity arising from the Farley-acoustic instability was calculated by Rogister and Jamin[52] but has not yet been utilized in global studies of the electrojet current

distribution. By and large, nonlinear studies of various instabilities have concentrated on properties of the instability itself, such as wavenumber spectrum, density fluctuation levels, etc. The studies generally do not produce effective resistivity formulas, turbulent diffusivities, enhanced joule heating rates, tec. which can be fed into equilibria models. Any understanding of how ionospheric instabilities effect the overall structure of the ionosphere must await future work. For example, auroral zone heating exerts a major influence on the meridional circulation of the neutral thermosphere.[14] If enhanced joule heating resulting from plasma instabilities occurs, then our model of the overall auroral zone heating could be in error.

3e. Increasing Our Understanding of Ionospheric Instabilities. Clearly a good theoretical representation of the equilibria which has withstood comparisons with experiment is the starting point for stability calculations. Both auroral arcs, where the electrical equilibrium is not clearly understoood, and, polar ionospheres where horizontal convection is a key element require improved equilibrium models. Satellite measurements experiments most likely will continue to provide the best description of magnetospheric current sources in the ionosphere, but one must be careful in deciding whether the observations uniquely define high-latitude current systems, or are merely consistent with a proposed model. Indeed, both ionospheric and magnetospheric physics could greatly benefit from a measurement program which did uniquely define the spatial pattern of electric field and field-aligned current systems. Experiments to date have chiefly been concerned about whether field-aligned current exists.[3,32] Irregularities in the polar and auroral F-regions are investigated principally by in situ satellite measurements. The Space Shuttle will give us the first opportunity to perform satellite-to-ground or satellite-to-satellite radar scattering experiments. Such observations could find out whether hitherto unobservable short-scale field-aligned instabilities exist in the auroral F-region as they do at the equator. Radar incoherent scatter studies of the auroral ionosphere will receive an important and needed boost with the operation of the European incoherent scatter facility, which will complement the very productive Chatanika facility in Alaska.

Further, even radio wave scattering studies have concentrated on density fluctuations which vary rather slowly with time. The possibility of scattering with frequency shifts close to the lower hybrid frequency as the recent Ott and Farley[41] instability predicts has been overlooked.

Evidently, experiments will continue to be the essential element in providing magnetospheric boundary conditions for ionosphere studies. But I hope that the discussion above had made clear that our ability to understand ionospheric response to changing these boundary conditions is just in its infancy. At this stage, an increased effort in nonlinear computational models is the single most important step required to increase our understanding of ionospheric plasmas. And, as our increase in understanding of ionospheric physics is promulgated through the community of magnetospheric scientists, perhaps the concept of the ionosphere being just a passive linear resistor will be replaced by better models of its nonlinear response to magnetospheric current sources.

4. The Equatorial Ionosphere - A Major Success

A careful reader of the preceding section will have noted that the equatorial ionosphere enjoys the greatest understanding. The purpose of this paragraph is to bring this material together for special emphasis. The unique geometry of the equatorial ionosphere means that it must carry strong currents in the E-region,[20] and have the density stratified in a direction orthogonal to the magnetic field.[38,56] Currently developed methods of plasma stability analysis are particularly suitable to this geometry[20,56] and the two-dimensional nature of the ensuing plasma turbulence permits accurate computational studies of Rayleigh-Taylor type

instabilities.[30,56] Hence a rather detailed understanding of equatorial spread-F has emerged from the computational simulations and has been verified by measurements.[30] These results are significant not only in their contribution to ionospheric physics but also in that they represent a unique comparison between nonlinear Rayleigh-Taylor instability theory and experiment. Nonetheless, there is still room for improvement: the analytic nonlinear theories[34,51,52,64] of the current-driven Farley-instability should be compared with truly definitive computational studies, comparable to those used for the Rayleigh-Taylor instability.

5. Active Ionospheric Experiments

A number of active experiments have been carried out in the ionosphere with unique and valuable scientific ramifications. Large barium releases have generated major perturbations in the ionospheric conductivity. The computational models[15,24,57,66] developed to describe the evolution and striation of these barium clouds have lead directly to the computational models which have been so successful in modelling equatorial spread-F.

The plasma instabilities which ensued when the ionosphere was illuminated by high power radio waves provided one of the first experimental demonstrations of plasma heating via nonlinear processes as well as motivating important theoretical work on the anomalous absorption caused by these instabilities.[45] The self-focusing instability of high-power radio waves was also observed experimentally[62] and theory[43] suggests that ionospheric striation via self-focusing may be an important environmental impact of a satellite power station.[9]

Electron beams launched by satellite borne electron guns provide interesting puzzles in the plasma physics of space-craft neutralization.[10,65]

So far there has been insufficient analysis to turn these interesting observations into new concepts in plasma physics. The latest theoretical work on neutralization is apparently that of Linson.[36]

Altogether, a variety of interesting and new scientific concepts have emerged from active experiments in the ionosphere. On the other hand, the contributions which particular active experiments have made, while valuable and unique, appear to be one-shot affairs, and not a basis for continuing research. But the general technique of a sequence of new and well-conceived active experiments will teach us more about the ionosphere and magnetosphere than most passive observation programs. The reason is that a deeper and more thorough understanding of space plasmas is required to predict the outcome of active experiments, rather than explain observations. Hence, if active ionospheric experiments are to continue to be of general intellectual significance, qualitatively new experiments must be invented. For example, the Space Shuttle could be used for hypersonic gas releases.[18]

6. Communications Applications of Ionospheric Instabilities

The unexpected influence of ionospheric structures on transionospheric satellite communications[63] links has lead to much research activity, a recent conference,[25] and an even more recent review article.[13] Both measurements[7,12,16,61] and nonlinear plasma instability calculations[38,56] show that the wavenumber spectrum of ionospheric irregularities has a power-law form. From such irregularity models, the effects on various properties of communications channels have been calculated.[11,47-49] The key role for plasma instability theory is to calculate the wavenumber spectrum of the irregularities so that one can obtain an understanding of how ionosphere effects on communications scale as a function of frequency as well as why and where instabilities occur.

7. Summary

Although observations have shown that the ionosphere is subject to several plasma instabilities, I hope that this chapter makes clear that a comprehensive study of ionospheric instabilities, their reaction back on the sources of free energy, and their effects on communication systems is just beginning. Even the equilibria states of the ionosphere have been modelled only by one-dimensional codes, and these require neutral winds, solar UV flux, and global electric fields as input parameters. Three-dimensional equilibria models, required for polar regions, have not yet been created. A theoretical program to systematically study what instabilities can occur in the various equilibria states has not really commenced, and the nonlinear study of ionospheric instabilities is even further behind. Hence we can conclude that a truly profound understanding of the ionosphere as a dynamical system with turbulent transport processes awaits substantially more research. Much of this research will be of "true intellectual significance" in that it must necessarily take up basic questions of turbulence. The complementarity between ionospheric and hydrodynamic turbulence holds the promise of being able to achieve a deeper understanding of both of these phenomena.

Acnowledgement: Discussions with R. G. Roble made substantive contributions to this chapter. I am also indebted to M. C. Kelly and S. L. Ossakow whose thoughtful reviews resulted in many improvements in the manuscript.

References

1. J. R. Aarons, R. Allen, and T. Elkins, Frequency dependence of radio-star scintillation, J. Geophys. Res. 72, 2891 (1967).

2. S. I. Akasofu and S. Chapman, Solar-Terrestrial Physics (Oxford, 1972).

3. H. R. Anderson and R. R. Vondrak, Observations of Birkeland current at auroral latitudes, Rev. Geophys. and Space Phys. 13, 143 (1975).

4. B. B. Balsley, G. Haerendel, and R. A. Greenwald, Equatorial spread-F: recent observations and a new interpretation, J. Geophys. Res. 77, 5625 (1972).

5. P. M. Banks, C. R. Chappel, A. F. Nagy, A new model for the interaction of auroral electrons with the atmosphere: spectral degradation, backscatter, optical emissions and ionization, J. Geophys. Res. 79, 1459 (1974).

6. P. M. Banks, R. W. Schunk and W. J. Raitt, Temperature and density structure of thermal proton flows, J. Geophys. Res. 79, 4691 (1974).

7. S. Basu and S. Basu, Correlated measurements of scintillations and in situ F-region irregularities from OGO-6, Geophys. Res. Lett. 3, 68 (1976).

8. H. F. Bates, Atmospheric expansion through joule heating by horizontal electric fields, Planet. Space Science 21, 2073 (1973).

9. W. C. Brown, Satellite power stations: a new source of energy?, IEEE Spectrum 10, 38 (1973).

10. D. G. Cartwright and P. J. Kellogg, Observations of radiation from an electron beam artifically injected into the ionosphere, J. Geophys. Res. 79, 1439 (1974).

11. E. Costa and M. C. Kelley, Calculations of equatorial scintillations at VHF and Gigahertz frequencies based on a new model of the disturbed ionosphere, Geophys. Res. Lett. 3, 677 (1976).

12. R. K. Crane, Spectra of ionospheric scintillation, J. Geophys. Res. 81, 2041 (1976).

13. R. K. Crane, Ionospheric scintillation, Proc. IEEE 65, 180 (1977).

14. R. E. Dickinson, E. C. Ridley, and R. G. Roble, Meridional circulation in the thermosphere I. Equinox condition, J. Atmos. Sci. 32, 1737 (1975).

15. J. H. Doles, III, N. J. Zabusky and F. W. Perkins, Deformation and striation of plasma clouds in the ionosphere 3, numerical circulations of a multi-level model with recombination chemistry. J. Geophys. Res. 81, 5987 (1976).

16. P. L. Dyson, J. P. McClure and W. B. Hanson, In situ measurements of the spectral characteristics of F-region ionospheric irregularities, J. Geophys. Res. 79, 1497 (1974).

17. P. L. Dyson and J. Winningham, Topside ionospheric spread-F and particle precipitation in the day-side magnetospheric clefts, J. Geophys. Res. 79, 5219 (1974).

18. D. S. Evans, High-velocity gas releases as a method for perturbing the upper atmosphere, J. Geophys. Res. 79, 3882 (1974).

19. J. V. Evans, A review of F-region dynamics, Rev. Geophys. Space Phys. 13, 887 (1975).

20. D. T. Farley, Two-stream plasma instability as a source of irregularities in the ionosphere, Phys. Rev. Lett. 10, 279 (1963).

21. D. T. Farley, B. B. Balsley, R. F. Woodman and J. P. McClure, Equatorial spread-F: implications of VHF radar observations. J. Geophys. Res. 75, 7199 (1970).

22. D. T. Farley and B. G. Fejer, The effect of the gradient drift term on type 1 electrojet irregularities, J. Geophys. Res. 80, 3087 (1975).

23. J. A. Fedder, Effects of anomalous resistivity on auroral Birkeland current systems, Ann. Geophys. 32, 175 (1976).

24. S. R. Goldman, S. L. Ossakow and D. L. Book, On the nonlinear motion of a small barium cloud in the ionosphere, J. Geophys. Res. 79, 1471 (1974).

25. J. M. Goodman, Ed., Effect of the ionosphere on space systems and communications, U.S. Government Printing Office, 1975.

26. R. A. Greenwald, Diffuse radar aurora and the gradient drift instability, J. Geophys. Res. 79, 4807 (1974).

27. A. DasGupta and L. Kersley, Summer daytime scintillation and sporadic-E, J. Atmos. Terr. Phys. 38, 615 (1976).

28. K. J. Jones, Wind, electric field and composition perturbations of the mid-latitude F-region during magnetic storms, J. Atmosph. and Terrest. Phys. 35, 1515 (1973).

29. M. C. Kelley, E. A. Bering, and F. S. Mozer, Evidence that the electrostatic ion-cyclotron instability is saturated by ion heating, Phys. Fluids 18, 1590 (1975).

30. M. C. Kelley, G. Haerendel, H. Kappler, A. Valenzuelo, B. B. Balsley, D. A. Canter, W. L. Ecklund, C. W. Carlson, B. Hausler, and R. Trobert, Evidence for a Rayleigh-Taylor type instability and upwelling of depleted density regions during equatorial spread-F, Geophys. Res. Lett. 3, 448 (1976).

31. J. M. Kindel and C. F. Kennel, Topside current instabilities, J. Geophys. Res. 76, 3055 (1971).

32. D. M. Klumpar, J. R. Burrows, and M. D. Wilson, Simultaneous observation of field aligned currents and particle fluxes in the post midnight sector, Geophys. Res. Lett. 3, 395 (1976).

33. W. C. Knudsen, Megnetospheric convection and the high-latitude F2 ionosphere, J. Geophys. Res. 79, 1046 (1974).

34. K. Lee, C. F. Kennel and F. V. Coroniti, On the marginally stable saturation spectrum of unstable type 1 equatorial electrojet irregularities, J. Geophys. Res. 79, 249 (1974).

35. P. C. Liewer and N. A. Krall, Anomalous penetration of a magnetic pulse into a plasma, Phys. Rev. Lett. 30, 1242 ((1973).

36. L. M. Linson, Current-voltage characteristics of an electron-emitting satellite in the ionosphere, J. Geophys. Res. 74, 2368 (1969).

37. M. A. MacLeod, T. J. Keneshea, R. S. Narcisi, Numerical modeling of a metallic sporadic-E layer, Radio Science 10, 371 (1975).

38. B. E. McDonald, T. P. Coffey, S. Ossakow, and R. N. Sudan, Preliminary report of numerical simulation of type 2 irregularities in the equatorial electrojet, J. Geophys. Res. 79, 2551 (1974).

39. H. G. Möller, Backscatter results from Lindau II - the movement of curtains of intense irregularities in the polar F-layer, J. Atmos. Terr. Phys. 36, 1487 (1974).

40. F. S. Mozer, C. W. Carlson, M. K. Hudson, R. B. Torbert, B. Parady, J. Yatteau, and M. C. Kelley, Observations of paired electrostatic shocks in the polar magnetosphere, Phys. Rev. Lett. 38, 292 (1977).

41. E. Ott and D. T. Farley, Microinstabilities and the production of short-wavelength irregularities in the auroral F-region, J. Geophys. Res. 80, 4599 (1975).

42. F. Perkins, Spread-F and ionospheric currents, J. Geophys. Res. 78, 218 (1973).

43. F. W. Perkins, and E. J. Valeo, Thermal self-focusing of electromagnetic waves in plasmas, Phys. Rev. Lett. 32, 1234 (1974).

44. F. W. Perkins, Ion-streaming instabilities: electromagnetic and electrostatic, Phys. Fluids 19, 1012 (1976).

45. A comprehensive account of radio wave heating experiments is found in the special issue of Radio Science, Nov. 1974.

46. M. H. Rees and R. G. Roble, Observations and theory of the formation of stable auroral red arcs, Rev. Geophys. Space Phys. 13, 201 (1975).

47. C. L. Rino and E. J. Fremouw, Statistics for ionospherically diffracted VHF/UHF signals, Radio Science 8, 223 (1973).

48. C. L. Rino, R. C. Livingston, and M. E. Whitney, Some new results on the statistics of radio wave scintillation 1. Empirical evidence for Gaussian statistics, J. Geophys. Res. 81, 2051 (1976).

49. C. L. Rino, Some new results on the statistics of radiowave scintillation 2, scattering from a random ensemble of locally homogeneous patches, J. Geophys. Res. 81, 2059 (1976).

50. R. G. Roble, The calculated and observed dirunal variation of the ionosphere over Millstone Hill on 23-24 March 1970, Planet. Space Sci. 23, 1017 (1975).

51. A. Rogister, Nonlinear theory of 'type 1' irregularities in the equatorial electrojet, J. Geophys. Res. 76, 7754 (1971).

52. A. Rogister and E. Jamin, Two-dimensional nonlinear processes associated with 'type 1' irregularities in the equatorial ionosphere, J. Geophys. Res. 80, 1820 (1975).

53. T. D. Roglien and J. Weinstock, Theory of the nonlinear spectrum of the gradient drift instability in the equatorial electrojet. J. Geophys. Res. 79, 4733 (1974).

54. C. L. Rufenach, Power-law wavenumber spectrum deduced from ionospheric scintillation observations, J. Geophys. Res. 77, 4761 (1972).

55. R. C. Sagalyn, M. Smiddy and M. Ahmed, High-latitude irregularities in the topside ionosphere based on Isis 1 thermal ion probe data, J. Geophys. Res. 79, 4252 (1974).

56. A. J. Scannapieco, and S. L. Ossakow, Nonlinear equatorial spread-F Geophys. Res. Lett. 3, 451 (1976).

57. A. J. Scannapieco, S. L. Ossakow, S. R. Goldman, and J. M. Pierre, Plasma cloud late time striation spectra, J. Geophys. Res. 81, 6037 (1976).

58. Soviet Killer Satellites: U.S. ponders a response, Science 193, 865 (1976).

59. D. J. Strickland, D. L. Book, T. P. Coffey, and J. A. Fedder, Transport equation techniques for deposition of auroral electrons, J. Geophys. Res. 81, 2755 (1976).

60. D. F. Strobel, T. R. Young, R. R. Meier, T. P. Coffey, and A. W. Ali, The nighttime ionosphere: E-region and lower F-region. J. Geophys. Res. 79, 3171 (1974).

61. R. R. Taur, Simultaneous 1.5 and 4-GHz ionospheric scintillation measurements, Radio Science, 11, 1029 (1976).

62. G. D. Thome and F. W. Perkins, Production of ionospheric striations by thermal self-focusing, Phys. Rev. Lett. 32, 1238 (1974).

63. W. F. Utlaut, Ionospheric scintillation: a potential limitation to satellite communications - important unknown scintillation factors, AIAA 12th Aerospace Sciences Meeting, paper 74-56 (1974).

64. J. Weinstock and A. Sleeper, Nonlinear saturation of 'type 1' irregularities in the equatorial electrojet, J. Geophys. Res. 77, 3621 (1972).

65. J. R. Winckler, R. L. Arnoldy, and R. A. Hendrikson, Echo 2: a study of electron beams injected into the high latitude ionosphere from a large sounding rocket, J. Geophys. Res. 80, 2083 (1975).

66. N. J. Zabusky, J. H. Doles, and F. W. Perkins, Deformation and striation of plasma clouds in the ionosphere, 2, numerical simulation of a nonlinear two-dimensional model. J. Geophys. Res. 78, 711 (1973).

IMPACTS OF IONOSPHERIC/MAGNETOSPHERIC PROCESSES ON TERRESTRIAL SCIENCE AND TECHNOLOGY

Edited by

L. J. Lanzerotti
Bell Laboratories
Murray Hill, New Jersey 07974

CONTENTS

1. INTRODUCTION

Naturally-occurring plasma processes in the space environment around the earth have produced visual spectacles for mankind in the form of the aurora since the polar regions were first visited and populated. With the growth of a technological civilization the auroral phenomenon assumed a more important role than that of a merely awesome display in the heavens. Its importance occurred because of the fact that many inventions and technical advances made use of the auroral-producing natural environment for their operation and/or were designed to function in an environment whose physical state is determined by little-understood processes.

Early in this century, of course, it was the existance of the ionosphere that made successful Marconi's wireless transmissions over large distances. It was whistler waves, produced by lightning and propagating in the magnetosphere, that caused interference on early communications systems in the first World War. It has long been known that drastic changes in the ionospheric conductivity produced by particle ejections from solar flares and by large magnetic storms would affect radio communications over a large portion of the globe.

As the technical needs of society have begun to require the utilization of larger portions of the space around the

planet, including the placing of sophisticated systems into this space, the impacts of the environment on the systems have become ever more important considerations. In addition, new ground-based technologies can be influenced and affected by space plasma processes. It is not unreasonable to expect, as past experience shows, that as technology changes and advances, subtle and not-so-subtle physical processes in the environment (probably unthought of at present; see chapter by Sturrock, this volume) will be of importance in determining how the technology will be adapted for implementation.

This chapter contains ten brief reports written by engineers and scientists who have been concerned with the implementation of science and technology under the constraints of the terrestrial space environment as it is presently known and understood. The subject areas were selected both because of their present technological and scientific importance as well as for their possible future interest in technological planning.

Common themes occur in several of the reports. For example, whenever society has needed long conductors to accomplish a technical task, the waves and current systems of the ionosphere and magnetosphere have been found to produce undesired effects by inducing currents in the conductors. These effects may consist of disruptions of the conductors in power lines or of the cables used in long-haul communications. The effects may take the form of unwanted background that hampers corrosion engineering survey work on pipelines.

Ionospheric and magnetospheric current systems may produce unwanted effects on long conductors, but they are needed scientifically as the "source" systems for carrying out deep induction studies of the earth's crust and upper mantle. However, the scale size of these source systems must be known, and the dynamical nature of the ionosphere and magnetosphere must be considered during such studies. These same current systems, so necessary for scientific deep-induction studies, provide an unwanted background for geophysical prospecting work in the search for petroleum and mineral resources.

The particles in the magnetosphere can affect space systems in a variety of ways. A concern since the discovery of the radiation belts has been the radiation damage of semiconductor components and systems by high energy particles of magnetospheric and solar origin. However, the low energy particles also affect spacecraft systems by producing differential electrical charging, with subsequent electrical breakdowns and discharges on satellite surfaces. Satellite logic and control systems can be drastically altered by such discharges.

The following reports do not, by any means, touch upon all of the technological and science areas that presently are impacted by magnetospheric and ionospheric processes. They do, however, provide representative examples of the wide range of topics of current interest. Each individual section is identified as to its major contributor.

2. SCINTILLATIONS OF COMMUNICATION STATELLITE SIGNALS AT GHz FREQUENCIES

(R. R. Taur, COMSAT Laboratories, Clarksburg, MD 20734)

2.1 Introduction

Rapid fluctuations (scintillations) of radio frequency signals propagated through the ionosphere have long been a concern to all designers of communication systems, civilian as well as military. The scintillations occur at all frequencies, from the HF(3-30 MHz) on up. It was believed that by going to microwave frequencies (GHz and higher), in addition to increased communication bandwidth, the ionosphere would no longer be a problem because it is essentially non-absorbing at such frequencies.

However, it was realized from the initial implementation of the INTELSAT network (using 4 and 6 GHz) that scintillations of the signal amplitudes (with peak-to-peak excursions ranging from 1 to 10 dB) are frequently reported from some earth stations located at low latitudes. Studies of the characteristics and patterns of the fluctuations led to the suggestion that they are produced by irregularities in the F-region (200-300 km) of the ionosphere (Craft and Westerlund, 1972; Taur 1973). Since that time, extensive experimental and theoretical investigations have been conducted to further understand and explain this phenomenon.

The morphology of the satellite-signal characteristics is now reasonably well mapped out as a result of the observational work. However, there is not yet an adequate theory that can satisfactorily relate ionospheric scintillations at GHz frequencies

to ionospheric scintillations at VHF, (30-350 MHz) which have been studied for many years. This situation may be attributed to the difficulty of calculating the multiple scattering effects as well as a lack of detailed knowledge concerning the structure of ionospheric irregularities, whose production mechanism remains largely unknown. Nevertheless, it appears that with the current knowledge of the morphological behavior and statistics of the ionospheric scintillations, reasonable design margin requirements can be estimated for inclusion in satellite communications systems in order to ensure satisfactory performance of the systems (Taur, 1973).

2.2 Characteristics of Ionospheric Scintillations

2.2.1 Patterns of Occurrence

Scintillations in the GHz frequency range occur principally between ±25° geomagnetic latitude near the years of maximum solar activity (Taur, 1973). Small amplitude signal fluctuations in the auroral zone have been reported by Pope and Fritz (1970). The boundary of the geomagnetic equatorial zone that produces scintillations appears to shrink with decreasing solar activity (Taur, 1974). Measurements were made in southeast Asia during 1975-1976 to investigate the equatorial scintillation "boundary" during near-minimum solar activity. Preliminary results of data analyzed to date indicate that, during periods of minimum solar activity, the boundary is located at about ±20° geomagnetic latitude. The variations of scintillations with latitude and longitude have not been established as yet due to the limited number of observing stations and satellites.

The scintillations are observed almost exclusively shortly after local ionospheric sunset during the months near equinox (Taur, 1973). As the number of sun spots decreases, it has been observed that the total number of days on which scintillation is observed decreases accordingly. There is apparently no change in the diurnal pattern as the level of solar activity changes.

2.2.2 Spectral Properties

Figure 2.1 contains a typical power spectrum of ionospheric scintillation at 4 GHz. These data were taken from a scintillation event measured at the Hong Kong earth station in September 1973. It is seen that, for frequencies > 0.125 Hz, the spectrum rolls off with an average slope of f^{-3}. This implies that under weak scattering conditions the 3-dimensional spatial power spectrum of the ionospheric electron density fluctuations should follow a power law proportional to f^{-5} (Taur, 1976). The roll-off of the scintillation spectrum starting at about 0.125Hz indicates that the relative perpendicular velocity between the ionospheric irregularities and the signal wave path is about 50 m/s. If it is assumed that the irregularities are mainly moving horizontally, the corresponding drift velocity should be about 150 m/s for a path elevation angle of about 20°.

During the weak scintillations, the scintillation index (defined as the normalized rms deviation in carrier power and which varies with wavelength λ between 1.5 and 4 GHz (Taur, 1976)) changes to a λ^2 variation between 4 and 6 GHz (Craft and Westerlund, 1972). Older theories would predict a frequency dependence of $\lambda^{1.5}$ under the assumption of a power-law size distribution of the ionospheric electron density under weak scattering conditions (Crane, 1975).

2.3 Conclusions

The effect of ionospheric scintillations on communications satellite signals at GHz frequencies is morphologically defined to a reasonable extent with the exception of a few areas, such as frequency dependence and directional variations, for which more measurements should be made to verify existing postulates. The solar activity dependence and the seasonal and diurnal patterns of GHz scintillations have suggested that the irregularities are more likely to be produced when the ionosphere is dense and unstable. These findings, together with some reported differences between weak scintillation activities on easterly and westerly communications links, have led to the hypothesis of a sharp wedge-like irregular region in the ionosphere near the sunset line. However, there is as yet no theory that can satisfactorily explain the production mechanism of the ionospheric irregularities that produce the scintillations in this frequency range.

In addition, there is little information pertaining to the thickness and spatial distribution of these irregularities. Such information is necessary for the development of a theory to correlate VHF and GHz scintillations (Warnik and Liu, 1974). The recent experimental and theoretical work on the Rayleigh-Taylor instability as related to spread-F conditions (Kelly et al., 1976; Scannapieco and Ossakow, 1976) could prove important in the ultimate understanding of GHz scintillations in the equatorial regions.

2.4 References

Crane, R. K., 1975, Spectra of Amplitude and Phase Scintillation, Proc. of the IES Symposium, Naval Research Laboratory, Arlington, Va., 1975.

Kelley, M. C., G. Haerendal, H. Kappler, A. Valenzuela, B. B. Balsley, P. A. Carter, W. Ecklund, C. W. Carlson, B. Hausler, and R. Torbert, Evidence for a Rayleigh-Taylor type instability and upwelling of depleted density regions during equatorial spread-F Geophys. Res. Letters, 3, 448, 1976

Pope, I. H. and R. B. Fritz, Observation of Simultaneous Scintillation on VHF and S-Band Satellite Transmissions at High Latitudes, NOAA TR-ERL-207-OD-6, November 1970.

Scannapieco, A. J., and Ossakow, S. L. Nonlinear equatorial spread F, Geophys. Res. Letters, 3, 448, 1976.

Taur, R. R., Ionospheric Scintillations at 4 and 6 GHz, COMSAT Technical Review, 3, 145, 1973.

Taur, R. R. Gigahertz Ionospheric Scintillation, COMSAT Technical Review, 4, 1974.

Taur, R. R., Simultaneous 1.5- and 4-GHz Ionospheric Scintillation, Radio Science, 1976.

Warnik, A. W., and C. H. Liu, Ionospheric Irregularities Causing Scintillation of GHz Frequency Radio Signals, J. Atmos. Terr. Phys., 871, 1974.

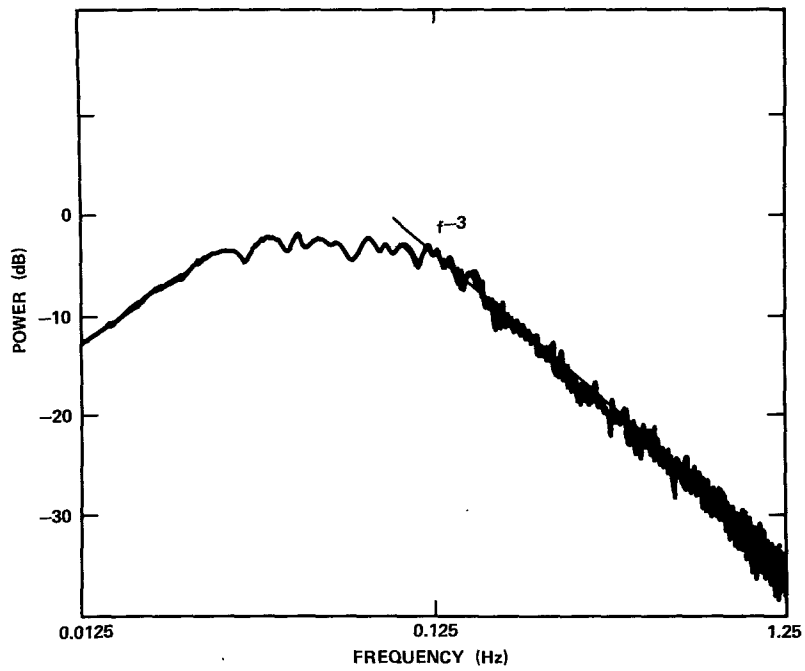


FIGURE 2.1 Power spectrum of 4 GHz ionospheric scintillation event measured at the Hong Kong earth station in September 1973.

3. MAGNETIC STORMS AND CABLE COMMUNICATIONS

(C. W. Anderson, American Telephone and Telegraph Company, Basking Ridge, N.J. 07920)

3.1 Introduction

The correspondence between geomagnetic storms and difficulties on cable communications systems has been observed for many years, both on oceanic systems and land-based systems. The "great" geomagnetic storm of March 24, 1940 (Germaine, 1940) was reported to have rendered inoperative eighty percent of all long-distance telephones out of Minneapolis, Minnesota. During another "great" geomagnetic storm on February 10, 1958, (Winckler, 1959) Western Union experienced severe interruptions on its North Atlantic telegraph cables. At the same time the Bell System transatlantic cable from Clarenville, Newfoundland, to Oban, Scotland, had induced voltages estimated to be 2,650 volts. Although the transatlantic cable system was never entirely out of service, the effect of the time-varying earth potentials was to have voices transmitted in the eastward direction at alternately loud squawks and faint whispers while the westbound signal strengths remained near normal. Another "great" geomagnetic storm in terms of communication cable considerations occured on August 4, 1972 (World Data Center Report, 1973) and provided the opportunity for a detailed study of ground induction effects on cable systems.

3.2 Geomagnetic Data Correlations

The existence of the Solar Forecasting Center maintained by the Space Environmental Laboratory, NOAA, in Boulder, Colorado, makes attractive the notion of forecasting magnetic storms which will cause difficulties on cable communications systems. Forecasts of the geomagnetic disturbance A index and its related Kp index, while useful for many activities, has not been found useful for predicting those large magnetic storms which will induce electric fields at the earth's surface greater than $\sim$10 V/Km, the fields required for severe communications impairments (Anderson et al., 1974). For example, an investigation by Anderson (1974) showed that unexplained outages of the L-4 long haul cable system occurred on only 8 of 24 days between January, 1969 and August, 1971 during which the Kp index was $\pm$7 (in a quasilogarithmic scale of 0 to 9).

The attempts to correlate the L-4 cable system outages with large Kp values have probably been frustrated because Kp is only a measure of the magnitude of the largest geomagnetic variation in a three hour interval and is not a measure of the time rate of change of the field. For example, Lanzerotti and Surkan (1974), through a statistical analysis of geomagnetic power spectra and Kp, concluded that there "exist no one-to-one relationships between power level

and K indices" and that "no simple, single-valued function can relate the power level to the K index." A geomagnetic index based on geomagnetic fluctuation levels ($\Delta B/\Delta t$) would greatly assist engineers in making estimates of the effects of geomagnetic field-induced currents on a system-wide basis.

3.3 Geomagnetic Disturbances

The geomagnetic disturbances that are normally associated with power system problems are believed to be manifestations of the intensification of the auroral electrojet giving rise to polar magnetic substorms (Albertson et al., 1973; Kisabeth and Rostoker, 1971). From a lack of sufficient geomagnetic data coverage it is not known if these types of disturbances have induced earth potentials large enough to produce the cable system problems reported in the past. However, such currents were not the cause of the problems during the August 4, 1972 storm when an extended ($\sim$30 minute) shutdown of a link in the Bell System long-haul coaxial cable (L-4) system occurred between Plano, Illinois, and Cascade, Iowa. A summary of the study by Anderson et al., (1974) is given here to demonstrate that for this occurrence, when reasonably good data were available, the induced currents in the earth were produced by magnetospheric disturbances as opposed to solely ionospheric currents.

One minute scalings of ground-based magnetometer chart-recorded data were made from 14 geomagnetic observatories in Canada and the United States. These data were used to examine the development of the geomagnetic disturbance around the time of the L-4 cable outage (∿2242UT on August 4, 1972). The rate of change of the magnetic field intensity and direction over North America for the one-minute interval 2241 to 2242UT is shown in Figure 3.1. The contours, plotted at 200γ/min levels (1 γ = 10^{-9} Tesla), were constructed on the basis of linear interpolation between magnetic observatory-measured values. A large change in the field magnitude was recorded over western Canada at this time. The change and the development of the disturbance did not resemble an ionospheric current system (Anderson et al., 1974).

At the time of the L-4 outage, two near-equatorial U. S. satellites, Explorer 45 and ATS-5, were over the Western Hemisphere and recorded magnetic field changes that signified the compression of the magnetopause to inside the spacecraft orbits at ∿4 R_e altitude (Explorer 45) and ∿5.5 R_e altitude (ATS-5). The crossings of the magnetopause occurred about two minutes apart at the two spacecraft, signifying large distortions in the boundary. A sketch of the equatorial plane view of the earth and the magnetospheric boundary at 2240UT and 2242UT is shown in Figure 3.2 together with the locations of the two spacecraft and four of the ground magnetic observatories

3.4 Induced Currents

The estimated direction of the field change at the time of the L-4 outage (∿70° with respect to the Plano-Cascade route) and its magnitude at Plano (∿700γ), both derived from Figure 3.2, were used together with a three-layer earth conductivity model to calculate the potential difference along the 242 km route. It was found that the calculated surface electric field of 7V/km along the Plano-Cascade cable was more than sufficient to cause a shutdown in the system, which was designed for earth potentials of 6.5 V/km ± 20%.

The design limit arose from the basic powering scheme for the cable. The power system is grounded at one end and the voltages of the DC-DC converters are balanced so that the voltage to ground at the "floating ground" is zero. In the presence of slowly varying voltages, the floating ground has a threshold of 370V, above which, for protection, it becomes automatically grounded. The floating ground must be restored to its normal condition manually.

3.5 Summary

It is not known as yet exactly what type of mechanism produced the large magnetic variations on August 4, 1972, but it appears they were associated with large assymetric distortions of the earth's magnetosphere, and hence large magnetopause currents, rather than the classic ionosphere currents. This area of investigation is

relatively undeveloped at present and, if pursued vigorously, should be beneficial to further understanding of solar wind-magnetospheric interaction processes.

The susceptibility of cable communication systems to geomagnetic storms has decreased over the years due to the phasing out of ground return cable systems. The major systems that would be (or might possibly be) affected are those that are DC powered from an earth ground. Efforts in the Bell System have now produced power regulation systems that can withstand earth potential variations up to 7.5 V/km without any noticeable transmission impairments. However, further work needs to be done on investigating the scale sizes of large geomagnetic disturbances and how these are related to the powering systems of transcontinental and transoceanic cables.

3.6 References

Albertson, V. D., J. M. Thorson, Jr., R. E. Clayton, and S. C. Tripathy, Solar-Induced currents in power systems: cause and effects, IEEE Trans. Power Apparatus and Systems, PAS-92, 471, 1973.

Anderson, C. W., III, Effects of the August 4, 1972, Magnetic Storm on Bell System Long-Haul Communication Routes, Corrosion '74, National Association of Corrosion Engineers, March 1974.

Anderson, C. W., L. J. Lanzerotti and C. G. MacLennan, Outage of the L-4 System and the Geomagnetic Disturbances of August 4, 1972, Bell System Technical Journal, 53, 1817, 1974.

Germaine, L. W., Magnetic Storm of March 24, 1940 - Effects in Communication Systems, EII Bulletin, May 7, 1940.

Lanzerotti, L. J., and A. J. Surkan, ULF Geomagnetic Power Near L=4,4 Relationship to the K-Index, J. Geophys. Res., 79, 2413, 1974.

Winckler, J. R., Auroral X-Rays, Cosmic Rays and Related Phenomena during the Storm of February 10, 1958, J. Geophys. Res., 64, 1959.

World Data Center A, Collected Data Reports on the August 1972 Solar-Terrestrial Events, UAG-28, 1973.

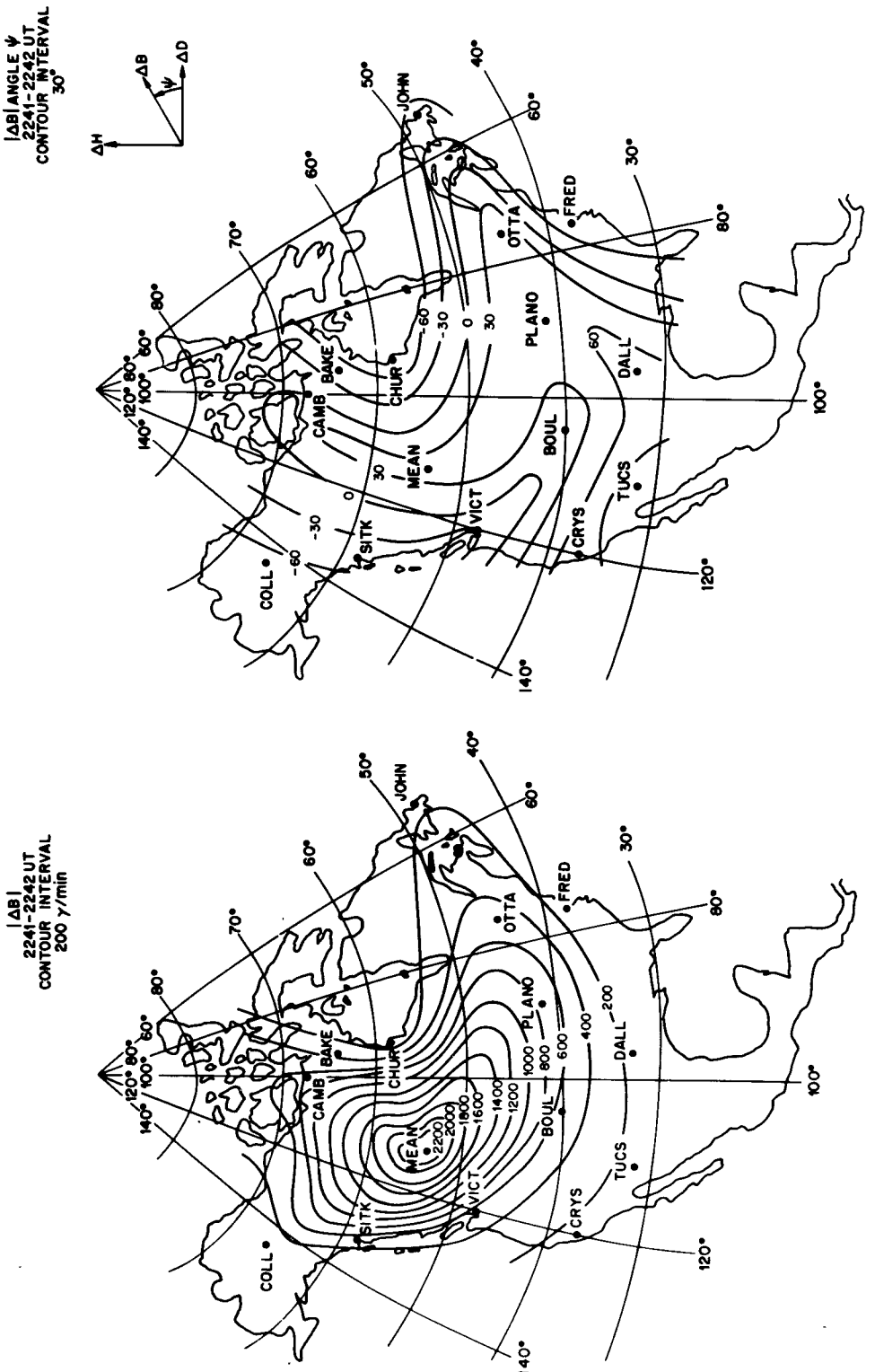


FIGURE 3.1 Rate of change of magnetic field intensity and direction over North America for the one-minute interval 2241 to 2242 UT on August 4, 1972.

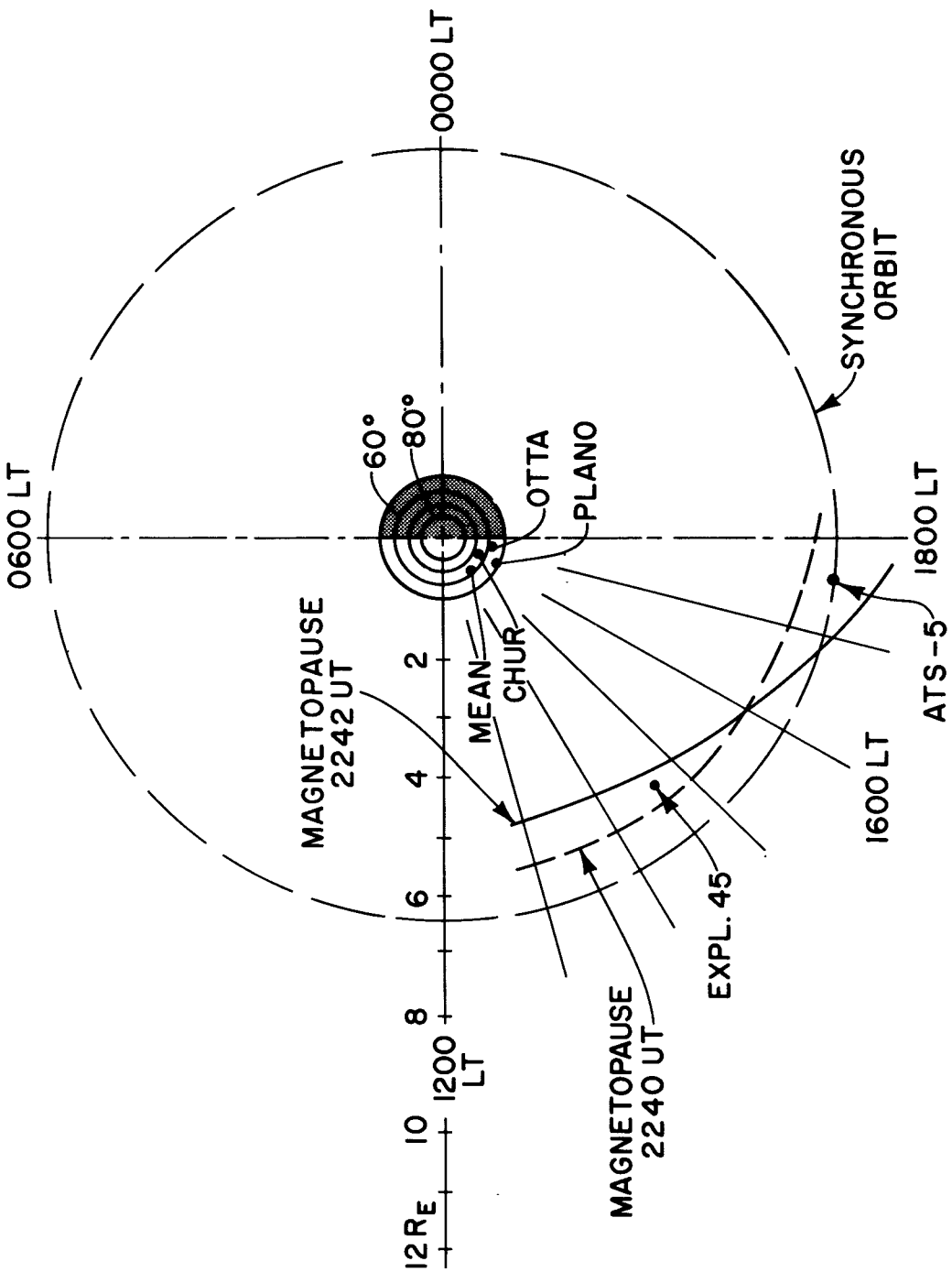


FIGURE 3.2 Equatorial plane view of earth and magnetospheric boundary in afternoon sector at 2240 UT and 2242 UT on August 4, 1972. The locations of the ATS-5 and Explorer 45 satellites are shown, as are several ground stations: Churchill (CHUR), Meanook (MEAN), Ottawa (OTTA) and Plano.

4. MAGNETOSPHERE IMPACTS ON GROUND-BASED POWER SYSTEMS

(D. J. Williams, Space Environment Laboratory, NOAA, Boulder, Colorado, 80302).

4.1 Introduction

One of the most dramatic manifestations of large scale magnetospheric plasma dynamics is the geomagnetic storm - a phenomenon identified as an intense worldwide decrease of the earth's surface magnetic field intensity produced by the sudden inward motion, and subsequent electric and magnetic field-produced drift, of the high altitude magnetospheric plasma. The impact of the geomagnetic storm on an important operational ground-based system - the ground-based power distribution system - is discussed in this section. The relevance here of magnetospheric research lies in the fact that our improved knowledge of the geomagnetic storm has allowed the power-distribution engineers to much better understand the reasons for their problems and thereby has guided future design work towards the most appropriate operational solutions.

It has been known since at least the Second World War that geomagnetic storms can cause severe disturbances in the operations of communications and power-distribution systems. The power-distribution effects can vary from reduced operating efficiency and interrupted service to users to equipment damage. Such effects have been well-documented in the literature (Davidson 1940; Brooks, 1959; Slothower and Albertson, 1967; McKinnon, 1972; Albertson, et al., 1973).

For example, during the geomagnetic storm of March 24, 1940, transformer banks in several northeast United States power stations were taken out of service due to incorrect operation of differential relays. Electric service was temporarily halted in portions of New England, New York, Eastern Pennsylvania, Minnesota, Quebec, and Ontario. During the February 9, 1958, magnetic storm Toronto, Canada, was plunged into a temporary blackout because of the tripping of circuit breakers in an Ontario transformer station. The development of the geomagnetic storm during the August 1972 solar flare activity (see also Section 3, above) was held responsible for the failure of a 230 KV power transformer at the British Columbia Hydro and Power Authority.

4.2 Cause of Problem

The cause of these disruptions of power-distribution systems is the induction of earth currents and resultant earth surface potentials during a geomagnetic storm. Calculations (Kellogg, 1966; Albertson and van Baelen 1970) have shown (and observations validate) that induced voltages up to 10 volts per mile can easily be generated during a geomagnetic storm. Induced voltages of this magnitude can disturb systems containing long conductive lines with equipments which are interconnected and grounded at large separation distances.

4.3 Effects of Induced Currents

Arbitrary differential relay operation in power-distribution systems during geomagnetic storms produces a judgmental problem in that system operators are unsure of whether or not the relay indication is of an induced transformer effect or a real transformer malfunction. Standard operating procedure is to remove transformers from the distribution networks during cases of relay tripping in order that they may be inspected. A complete inspection takes a few days and is required in order to protect a valuable and sizeable commercial investment (power transformers cost in the $0.5 - 1.0 million range). Since the northeast U.S is supplied by a complex power grid stretching throughout the northeastern and northcentral United States, as well as southern Canada, the arbitrary loss of transformers within this network during a geomagnetic storm could conceivably result in a massive power blackout. The final understanding that geomagnetic activity is the cause of these relay perturbations has led to a reasonable operational procedure to significantly reduce potential hazards. Given prior knowledge of geomagnetic disturbances, operators now do not remove transformers from the supply network in the case of a relay trip during a geomagnetic storm.

A more serious effect of geomagnetic storms is the current induced in the windings of power transformers

(Albertson, et al., 1973). The magnitude of these currents can be as high as 100 amps and can result in half-cycle saturation of the transformer core. Figure 4.1 shows examples of such induced currents. At the bottom of the figure is shown the world-wide magnetic activity index, D_{st}, in gammas ($1\gamma=10^{-9}$ tesla) for the period 13 - 16 May, 1969. A geomagnetic storm begins at the end of 14 May. In the center of Figure 1 a magnetogram trace recorded at Great Whale River in the auroral zone is shown on an expanded scale from the shaded time interval of the D_{st} plot. The rectangles placed above the magnetogram trace show the time periods of induced currents in transformer windings and the magnitudes of the currents. Actual recordings of the induced currents are shown on an expanded scale in the upper right-hand corner of Figure 4.1 for the interval 2100 UT 15 May through 0200 UT 16 May for three power stations. The Philadelphia Electric Company Power Station at Delta, Pennsylvania (on nearly the same geomagnetic meridian as Great Whale River) shows an induced current of approximately 80 amps during this interval.

Some direct consequences of currents causing half-cycle saturation of transformer cores are fluctuations in the distribution system voltages and internal localized heating in the transformers themselves. Excess internal local heating produced by these induced currents may thermally degrade either the insulation of the transformer core

laminations or any major insulation adjacent to a structural steel path. Such thermal degradation is cumulative in nature and will contribute to an overall shorter insulation life expectancy. As lifetime considerations are a major element in the costing of electric power, this particular problem becomes very real to the everyday user of electricity.

Possible means for reducing the induced current magnitudes are being studied and developed (Albertson, et al., 1973). While future systems should have better protection features, it is unclear whether or not it will prove practicable to incorporate protective devices in the huge existing power network. Thus, the problems posed to ground-based power distribution systems by geomagnetic storms are quite real. While these problems can best be solved by power systems engineers, our continuing and advancing understanding of magnetospheric processes and how they effect geomagnetic storm development should be made available to operators of these systems in order that they may understand in a much more comprehensive way the input conditions defining their problems.

4.4 References

Albertson, V. D., J. M. Thorson, Jr., R. E. Clayton, and S. C. Tripathy, Solar-induced currents in power systems: cause and effects, IEEE Transactions on Power Apparatus and Systems, PAS-92, 471, 1973.

4.4 References (cont'd.)

Albertson, V. D., and J. A. van Baelen, Electric and magnetic fields at the earth's surface due to auroral currents, IEEE Transactions on Power Apparatus and Systems, PAS-89, 578, 1970.

Brooks, J., The Subtle Storm, A Reporter At Large, New Yorker Magazine, 39, Feb. 19, 1959.

Davidson, W. F., The magnetic storm of March 24, 1940 - Effects in the power system, Edison Electric Institute Bulletin, 365, July 1940.

Kellogg, P. J., Terrestrial effects of solar activity, Proceedings of Minnesota Power Systems Conference, October, 1966.

McKinnon, J., The August 1972 solar activity and related geophysical effects, NOAA Space Environment Laboratory Report, December 1972.

Slothower, J. C., and V. D. Albertson, The effects of solar magnetic activity on electric power systems, Journal of the Minnesota Academy of Science, 34, 94, 1967.

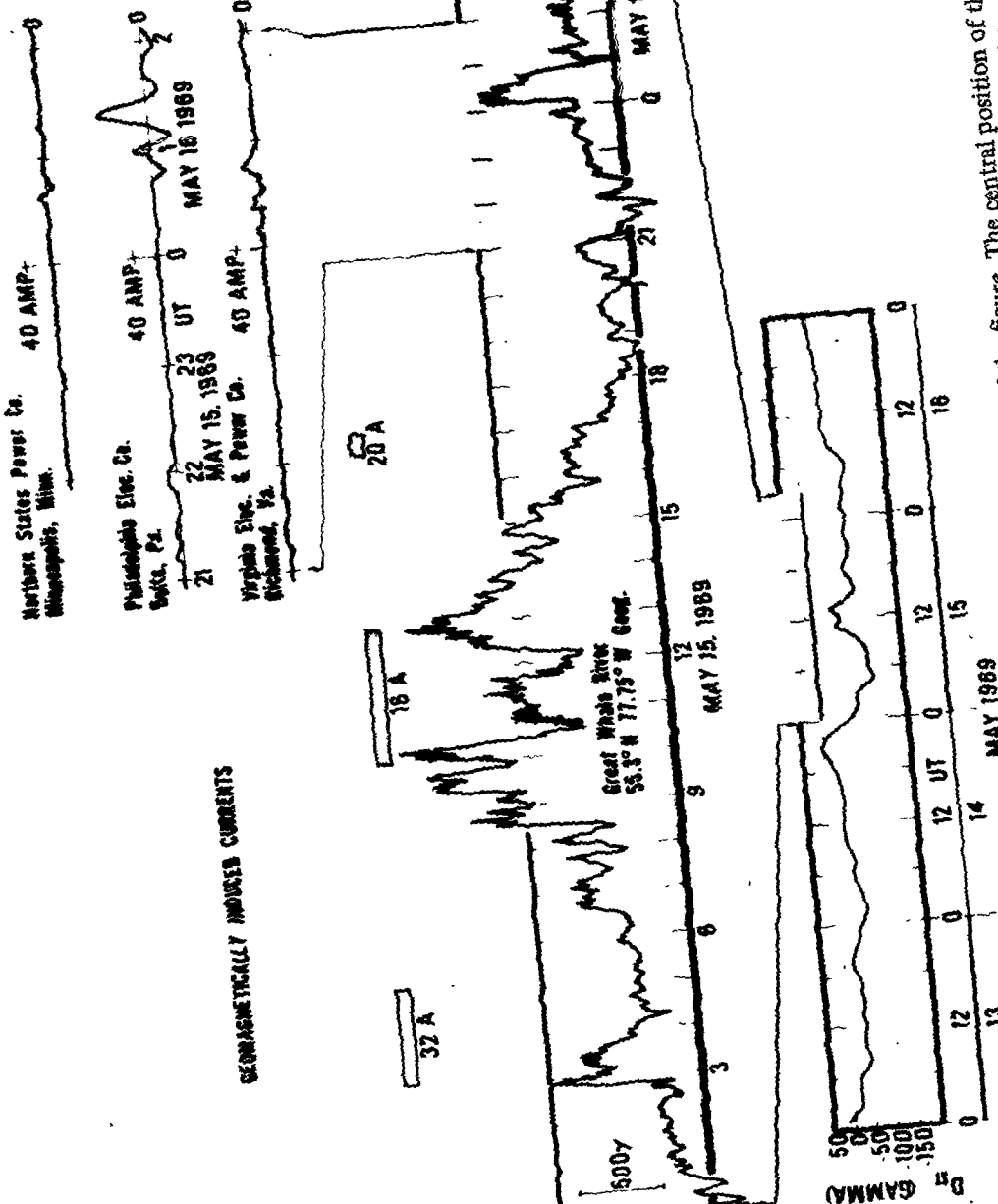


FIGURE 4.1 The world-wide geomagnetic activity index D_{st} is shown in the lower position of the figure. The central position of the figure shows, on an expanded time scale, a magnetogram trace obtained at Great Whale River during the geomagnetic storm shown on the D_{st} plot. Units are in gamma ($1\gamma = 10^{-9}$ tesla). Times and magnitudes of the currents induced in transformer windings are indicated. The top position of the figure shows three induced current traces obtained during the time period indicated in the magnetogram trace.

5. THE SPACE RADIATION ENVIRONMENT: EFFECTS ON SPACE SYSTEMS

(G. A. Paulikas, Aerospace Corp., El Segundo, Calif., 90009).

5.1 Introduction

When Arthur C. Clarke proposed the concept of synchronous-orbiting communications satellites, he could not have anticipated that energetic particle radiation, trapped in the earth's magnetic field, would be one of the factors that designers of synchronous-orbiting spacecraft would have to face. Similarly, designers of early concepts of space stations envisioned these stations to orbit the earth at an altitude of about 1000 miles - an orbit which we now know passes through the heart of the inner Van Allen radiation belt. Interestingly enough, the space station featured during the 1960's in Disneyland's "Voyage to the Moon" was in just such an orbit and an entire generation of visitors were misled regarding the very real hazards to man and machines posed by energetic trapped radiation.

Discovery of the Van Allen belts was the first major scientific result derived entirely from observations carried out on board spacecraft. Several generations of experiments, and many billions of data bits later, the scientific interest and practical applications derived from our knowledge about the trapped radiation continue

to make research on energetic particles in the magnetosphere exciting and rewarding.

The phenomenology of the radiation belts and the relevant physical processes differ for different regions of space, and a full description of the practical benefits to be gained by an improved understanding of the space environment would be unduly complex for the present purpose. Accordingly, this section presents a description of the radiation environment at synchronous orbit (altitude ∿5.6 earth radii above the surface) and the effects of this radiation environment on the cost and capabilities of present and future space systems.

5.2 Synchronous Altitude Environment

The synchronous orbit, because of its obvious utility for terrestrial applications, is undoubtedly the single most heavily populated orbit in space. More than fifty spacecraft, spread over longitude, are flying in this orbit today, and the plans of various nations (Hearth, 1976; Bekey and Mayer, 1976) emphasize an even heavier utilization of it for future communications, earth observations, meteorology, and data relay spacecraft. In addition to the practical uses of the orbit, it also happens that this altitude region is the location of a fascinating space plasma physics laboratory. The plasma pause, the extraterrestrial ring current, the boundary of the zone

of trapped energetic particles (and hence the approximate limits of access of solar particles) and the earthward terminus of the magnetotail plasma sheet, all meet and interact near this location. The study of the dynamics and interactions of the various plasmas with vastly different temperatures and densities and the development of an understanding of the processes which generate and transport energetic electrons and energetic protons is an area of research vital to economic utilization of synchronous-orbiting spacecraft.

The trapped radiation environment at the synchronous orbit consists of electrons and protons (and higher atomic number particles) with energy spectra as exemplified in Figure 5.1. These data may be considered representative although it is known that order-of-magnitude changes in intensity are very common. It is this variability that provides both interesting problems to scientists and significant headaches to space system designers. Some of the problems known to occur when a spacecraft is immersed in this environment are sketched below.

5.3 Impacts of the Environment

The fluxes of low energy particles-electrons and protons with energies less than about 20 KeV - are sufficiently intense to cause degradation of thermal

control surfaces and exterior coatings; this portion of the particle energy spectra also determines whether the phenomenon of spacecraft charging (discussed in Section 6, below) occurs. The high specific ionization of protons in the hundredsof KeV to MeV energy interval makes mandatory the complete coverage of solar cells for protection. Trapped energetic electrons are the major cause of radiation damage to solar arrays as well as to semiconductor devices contained in the deep interior of a spacecraft. For example, a typical solar cell array of a spacecraft loses a few percent of the original power output per year of exposure (Figure 5.2).

The useful lifetime of radiation sensitive devices such as computer memories, although these may be deeply buried in the interior of a spacecraft, are totally determined by their accumulated radiation dose from trapped energetic electrons during periods of low solar activity, and by the sum of the dosage from trapped electrons and solar cosmic ray protons (and alpha particles) near the maximum in the solar cycle. The recent trend toward the use of large-scale integrated circuit semiconductor devices aboard spacecraft has served to reinforce the appreciation that modern electronics, although capable of spectacularly efficient performance, remain sensitive to radiation damage. Thus, a considerable premium is placed

on accurate knowledge of the environment so that quantitative predictions of component lifetime (or necessary shielding weight) can be made. (The problem is magnified because mass market pressures, rather than considerations of radiation hardness, drive the semiconductor manufacturers.)

Quantitative predictions of the space radiation environment at the accuracy desired by spacecraft designers elude space scientists. Order-of-magnitude variations in the electron and proton fluxes are commonly observed (Figure 5.3). The environment at synchronous altitude resembles, in its dynamic behavior, nothing as much as terrestrial weather, with the important qualification that all quantities must be plotted on logarithmic scales. The spacecraft designer calls for 10% predictions; the space scientists offer him factors of two in accuracy. A factor of two uncertainty in the expected radiation dose may mean a considerable uncertainty in the expected lifetime. For a spacecraft costing $50 million, any increase in replenishment rate has formidable fiscal implications. As a result, the designer overdesigns and, in so doing, increases the costs above the minimum in order to compensate for the inadequate knowledge about the space environment.

Uncertainties of a factor of two in the environment lurk everywhere. For example, the fact that the geomagnetic equatorial plane does not coincide with the geographical equatorial plane means that some longitudes

in the synchronous orbit are a more benign location than others - by about a factor of two.

The radiation intensities also seem to fall off rather rapidly as one moves off the magnetic equator. It may thus be advantageous for some purposes to place a spacecraft into a slightly inclined synchronous orbit, purchasing in the process an improvement in the radiation environment at the expense of additional operational complexity. Solar activity, affecting the magnetosphere via the solar wind and the interplanetary magnetic field, causes both long term (11 years) and short term (days) changes in the energetic radiation environment.

Research in the physics of the trapped energetic particles has long since passed the exploratory stage. We now fully appreciate that the trapped energetic particles are but the tail end of a broad distribution (Figure 5.1) whose evolution in time and space is governed by a complex of interactions with electromagnetic waves generated by space plasmas themselves. Some of the waves are generated by the local distribution of particles; other families of waves are generated by different particle populations in other regions of the magnetosphere and reach the region of interaction via complicated paths. Studies of wave-particle interactions have reached the level of sophistication and precision that experiments - in the laboratory sense of the word - to change the energetic particlepopulations

seem feasible using man-generated beams of properly tailored electromagnetic waves. While the practical effects of such experiments, if successful, may seem on the surface desirable - for example, sweeping the synchronous "corridor" free of energetic electrons - no one has yet considered how an environmental impact statement for such a necessarily global experiment would be written.

The contributions that solar cosmic ray protons make to the total radiation dose accumulated by a synchronous orbiting spacecraft were mentioned earlier and illustrated in Fig.5.2. It turns out that solar protons have surprisingly efficient access to this region of space and the radiation damage associated with a single major solar proton event may be comparable to the damage caused in several years of exposure to the trapped energetic electron population (Figure 5.2). The discovery that even low energy solar particles can readily penetrate to the synchronous orbit was one of the surprising results returned by the first scientific experiment flown aboard a synchronous orbiting spacecraft (Lanzerotti, 1968). The earth's magnetic field is a far less effective shield than had been expected on the basis of simple calculations. It now appears that fully self-consistent calculations of solar cosmic ray trajectories, including models of the

geomagnetic field that realistically account for internal and external current systems and hydromagnetic wave populations, are necessary in order to arrive at an agreement between theory and the experimental observations of solar cosmic ray access to, and diffusion in, the inner magnetosphere.

It is clear that long-term studies of the radiation environment are necessary in order to develop a data base of sufficient spatial and temporal coverage to allow accurate specification of the radiation environment. Simultaneously, the development of an understanding of the physical processes governing the evolution of the space environment is necessary if we are ever to approach the predictive capability which we seek.

5.4 Summary

As noted in the scientific discussions, there are strong similarities and analogies between the work magnetospheric physicists must do in order to carry out analyses of the plasmas encountered in space and the research going on in other branches of plasma physics. For example, the processes governing the loss of particles from the radiation belts are far better understood than the processes which energize and transport the particles about the magnetosphere. The former, essentially involving wave-particle interactions, is concerned with resonances

between various families of electromagnetic waves and gyrating particles playing the key role in diffusing particles in pitch angle and hence out of the "containing" volume. The later processes - energization and transport - are more akin to stochastic or hydromagnetic problems often encountered in laboratory work. Using detailed wave-particle interaction calculations, space scientists have obtained quite satisfactory agreement between the theoretically calculated and the experimentally observed structure of portions of the earth's radiation belts (Coroniti and Thorne, 1973). No comparable progress is yet visible in quantitatively calculating the generation and transport rates of energetic electrons; such an accomplishment would have clear and beneficial practical consequences.

Looking toward the future, an enormous expansion is foreseen in the number of space systems designed to serve terrestrial applications and terrestrial customers. The large majority of earth-applications spacecraft are most efficient if located in geostationary orbit. A sample of the kind of synchronous-orbiting applications spacecraft which seem possible with only modest extrapolations of present technology can be found in reports edited by Hearth (1976) and in Berkey and Mayer (1976). It is clear that the investment in such space systems is likely to run into the tens of billions of dollars. For example, studies

of a Space Solar Power Station (SSPS, see Section 7, below) envisions solar arrays 50 square kilometers in area, costing $2 - $3 billion! The premium on a precise, quantitative understanding of the space environment and the impact of the environment on space systems is enormous. Even savings of fractions of a percent, derived as a result of better information regarding the energetic radiation (for example, extending the life of a spacecraft and thus decreasing the replenishment rate) translate directly into savings which run into the tens of millions of dollars.

5.5 References

Bekey, I., and H. Mayer, 1980-2000: Raising our Sights for Advanced Space Systems, Aeronautics and Astronautics, 14, Number 7/8, 34, 1976.

Coroniti, F. V., and R. M. Thorne, Magnetospheric Electrons, Ann. Rev. Earth and Planet. Sci.,1, 107, 1973.

Goldhammer, L. J. and S. W. Gelb, Synchronous Orbit Performance of Hughes Aircraft Company Solar Arrays, paper presented at the 11th Intersociety Energy Conversion Engineering Conference, September 1976.

Hearth, D. P., ed., Outlook for Space, National Aeronautics and Space Administration Report NASA SP-386, January 1976.

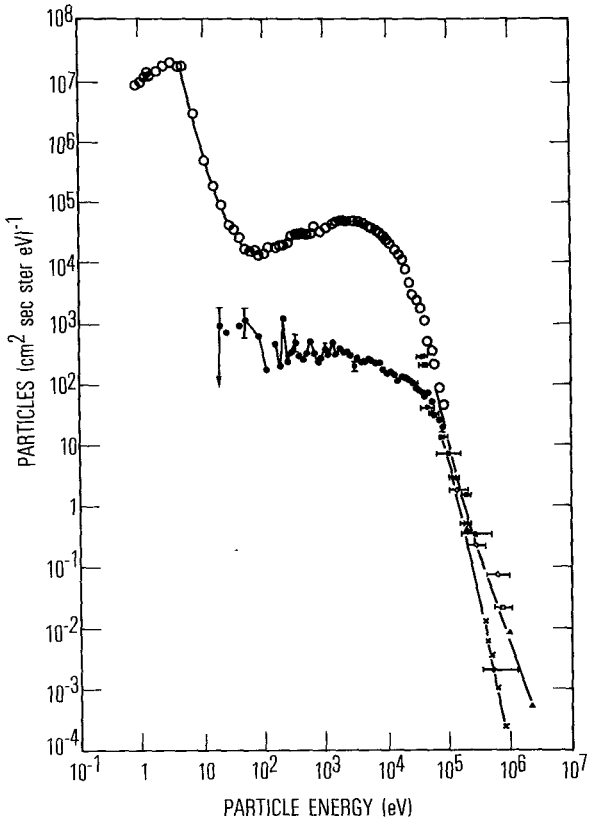


FIGURE 5.1 Energy spectrum of trapped electrons and protons as observed in the synchronous orbit by experiments on ATS-6.

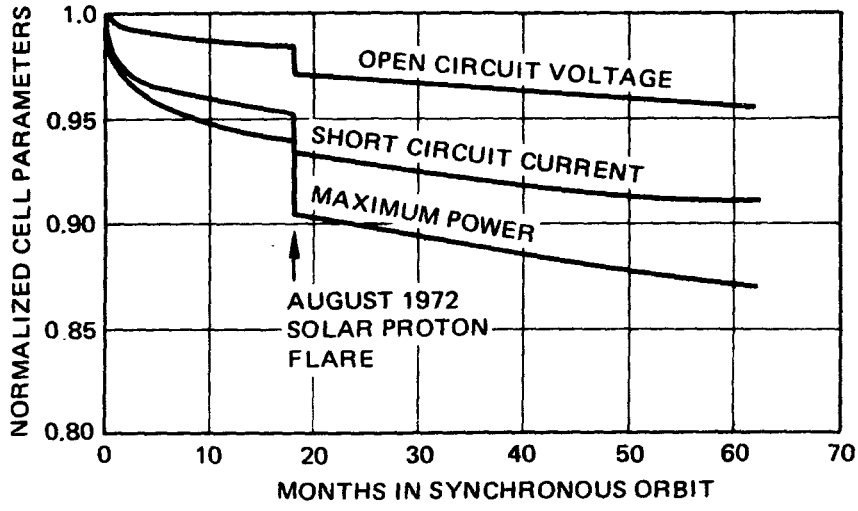


FIGURE 5.2 Performance of a solar cell array as a function of time after launch. The long-term decrease in performance is attributable to damage by energetic electrons; the step function decrease is the effect of one large solar proton event (from Goldhammer and Gelb, 1976).

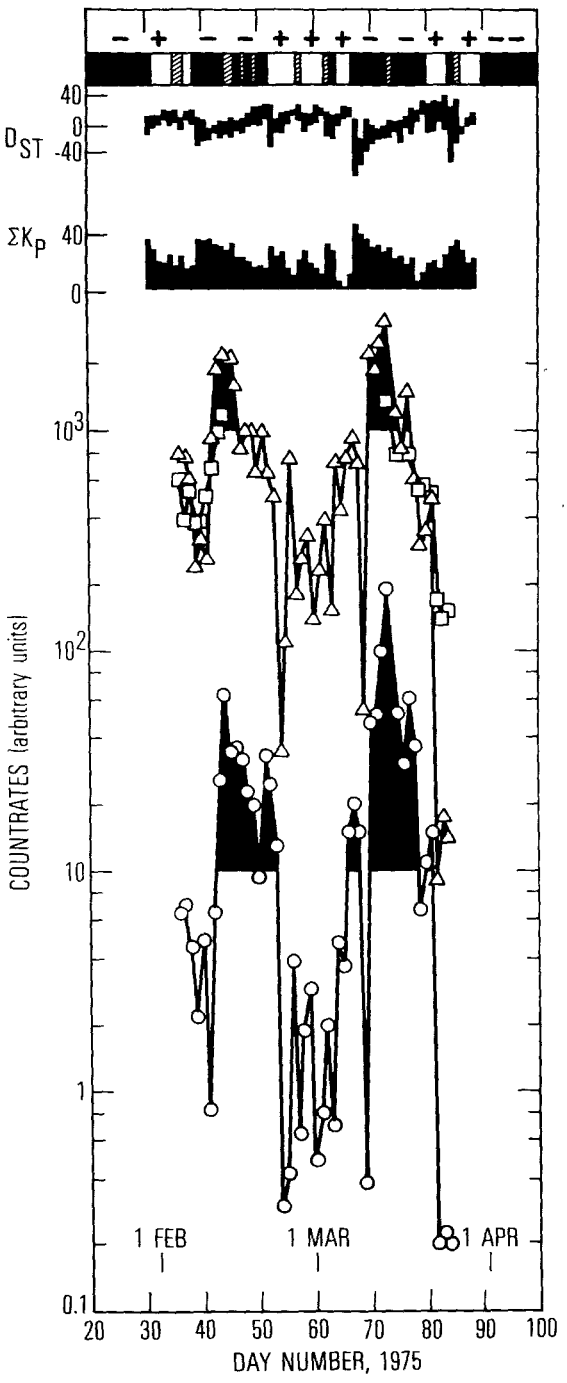


FIGURE 5.3 Variations in the energetic electron flux (□, △ = electrons > 1.6 MeV; 0 = electrons > 3.9 MeV) at the synchronous orbit (bottom traces) correlated with the direction of the interplanetary magnetic field and two geomagnetic disturbance indices (from Paulikas and Blake, 1976).

6. ELECTROSTATIC CHARGING ON SPACECRAFT

(S. DeForest, University of California, San Diego, La Jolla, Calif., 92037)

6.1 Observations and Effects

6.1.1 Scientific Problems

An electrically isolated probe inserted in a plasma will generally assume a negative electrostatic potential with respect to the undisturbed plasma. The magnitude of this potential is of the order of the electron temperature expressed in electron volts. This is true for spacecraft inserted into natural plasmas as well as for probes in laboratory plasmas. Thus, potentials of a few volts are commonly observed on spacecraft in the plasmasphere (Whipple, 1974) but potentials greater than 10,000 volts can be seen on spacecraft in the plasma sheet or in plasma cloud regions of the magnetosphere (DeForest, 1972). The record observed to date is 19,000 volts measured on the synchronous orbit ATS-6 satellite during an eclipse of the sun on the satellite by the earth.

All uncontrolled spacecraft charging compromises or ruins the measurement of low energy particles of all species. When the complication of photoelectrons and other locally-produced secondary particles are considered, a spacecraft can assume a slightly positive potential, but the plasma sheath surrounding the vehicle has a potential minimum (Whipple, 1975, Grard, 1972). Therefore, natural

plasma particles of both signs are excluded from direct measurement. This is a very serious problem since the low-energy particles that do reach a badly mounted instrument are locally-produced secondaries which reflect from the sheath potential minimum. These secondaries are only indirectly linked to the ambient natural flux. Their densities will vary, therefore, in a way that mimics the behavior of the ambient particles.

Nearly all spacecraft have exteriors consisting of various insulators. This type of outer surface is usually desired for thermal control. However, the charge stored on different parts of the spacecraft can cause differential potentials of thousands of volts between insulating surfaces. In general, the sunlit side of a spacecraft willbe held to near zero potential while the dark side charges negatively. For spacecraft at the important synchronous orbit, (see Section 5 above) this means that locally-produced electric fields of magnitude few volts per meter can distort both particle and field measurements.

6.1.2 Operational Problems

The overall charge state of a spacecraft probably has minimal impact on operations. However, if the vehicle is charged negatively for long periods some increased surface degradation due to ion bombardment will be noticed.

The major operational effect arises from differential charging. Differential charging is produced by

exposed insulators or electrically floating conductors. Discharges from external vacuum-deposited aluminum (VDA) can remove the outer surface and thus degrade the thermal properties. Similarly, the outer surfaces of optical elements can be seriously affected.

McPherson (1975) and Rosen (1975) have given convincing evidence that several different classes of spacecraft operating in synchronous orbit experience operating anomolies due to discharges produced by electrostatic charging. A frequent occurrence is false commands being given by noise on the spacecraft command lines. In at least one case, total loss of an Air Force spacecraft most probably occurred because of extreme differential charging prolonged by an intense geomagnetic substorm.

6.2 Understanding and Controlling Spacecraft Charging

6.2.1 Understanding

Theoretical and experimental studies are currently under way (McPherson, 1976) under sponsorship of several agencies to understand and control spacecraft charging. The level of understanding needed varies with the user. Scientifically, some investigators want to understand the phenomenon completely as an interesting plasma physics problem with far-reaching cosmological implications. Magnetospheric physicists would like to be able to measure uncontaminated natural plasmas. Practically, many systems

offices are trying to run operational spacecraft, and they simply want to fix the problem of extraneous commands - with as little redesign of existing vehicles as possible.

In general, an approach to understanding must begin with a reasonable model of the environment in which the vehicle is to operate. For many common orbits (synchronous in particular) insufficient data presently exist to specify the environment completely. For successful modeling and prediction, a complete distribution function from less than an electron volt to about 100 KeV is needed for all ion species and electrons. This must be known as a function of time and space. This requirement also implies a knowledge of local electric and magnetic fields. For example, McIlwain (1976) has only rather recently demonstrated the existence of strong magnetic field-aligned ion fluxes. No mass spectrometers have yet flown at synchronous orbit (although both the European Space Agency GEOS and the U. S. Air Force SCATHA spacecraft will carry them).

An understanding also involves knowledge of the response of various materials to the environment. This response includes photoelectron emission, particle backscatter coefficients, and secondary electron production. Differences in these properties can create potential differences between adjacent materials even when they are exposed to an identical environment. From the spacecraft charging

point of view, the best outer layer is a conductor, but frequently the spacecraft thermal constraints cannot be met with pure conductors. For other spacecraft parts, such as solar cells, making the outer layer conducting is prohibitively expensive, and is of dubious value if the conductor cannot withstand the radiation environment.

6.2.2 Controlling

Once the mechanism of charging is understood, it can be controlled by a careful selection of the spacecraft external surfaces and their placement. This care will insure the health and design life of a spacecraft. However, design problems with the outer surfaces of spacecraft optics may still exist.

Scientific spacecraft which have as a prime mission the measurement of low energy particles will most probably have to carry some means of active control of the spacecraft potential. This control could be achieved by an electron gun or some other type of active device. In a few cases, there is the possibility of simply enclosing the entire spacecraft within a continuous conducting surface.

6.3 Present and Future Impacts

Just as the need for spacecraft magnetic cleanliness specifications has been known for some time, so now electrostatic cleanliness specifications are being proposed for spacecraft. The need for these specifications is

evident for scientific observations as well as for the reliable operation of many classes of space vehicles. Such a specification has been proposed for the NASA ISEE program and the Dynamics Explorer.

In the future it can be expected that spacecraft construction will take into consideration the protection of electronic devices from environmentally-produced discharges. The principal way in which this will be accomplished is through careful attention to details of grounding and to the development of new materials which will have thermal and electrical properties tailored to suit the needs of both the thermal design and the electrostatic specifications. McPherson (1976) has recently reviewed the work being done in this field.

Another prospect for the future is active control of the potential by electron guns, ion accelerators, or other means. One current suggestion is to shine an ultraviolet light on the dark side of a spacecraft in order to eliminate differential charging.

6.4 Other Spinoffs

Cosmological Studies

As mentioned above, spacecraft charging has cosmological implications. Many investigators are interested in the electrostatic charging of bodies in space by particles because of the implications for planet formation. There is obviously much carry-over from the study of spacecraft to celestial objects.

For large bodies, the effects of natural charging might be readily evident. For instance, dust migration by charging has been proposed as an explanation for the difference in surface features between the front and back side of the moon (Grard, 1972).

6.5 Summary

Spacecraft can now be designed to withstand the effects of charging - after all, most of them do quite well most of the time anyway. For a scientific payload, great care must be taken to avoid spurious noise in the measured particle fluxes and the electric and magnetic field data. However, for many operational spacecraft the question becomes one of probability and economics. What construction techniques will have the highest expected return (e.g., communications, weather observation, etc.) for the lowest price? An answer to that question requires a better understanding of the environment and materials response before specific recommendations can be made to the spacecraft designer.

6.6 References

DeForest, S. E. Spacecraft Charging at Synchronous Orbit, J. Geophys. Res., 77, 751, 1972.

Grard, R. J. L., Photon and Particle Interaction with Surfaces in Space, P. Reidel Publishing Co., Dordrecht-Holland, 1973.

6.6 References (cont'd.)

McIlwain, C. E., Auroral Electron Beams Near the Magnetic Equator, in Physics of the Hot Plasma in the Magnetosphere, edited by B. Hulquist and L. Stenflo, Plenum Publishing Corp., New York, N. Y., 1976.

McPherson, D. A., D. P. Cauffman, and W. Schober, Spacecraft Charging at High Altitude -- the SCATHA Satellite Program, AIAA paper 75-92, presented at AIAA meeting in Pasadena, Ca. 1975.

McPherson, D. A., Spacecraft Charging Modeling Status, prepared for SAMSO, Los Angeles Air Force Station, Los Angeles, Ca., 1976.

Rosen, A., Spacecraft Charging Environment Induced Anomolies, AIAA paper 75-91, presented at Pasadena, Ca., January 1975.

Whipple, E. C., J. M. Wapnock, and R. H. Winkler, Effect of Satellite Potential on Direct Low Density Measurements Through the Plasmapause, J. Geophys. Res. 79, 179, 1974.

Whipple, E. C., Observation of Photoelectrons and Secondary Electrons Reflected from a Potential Barrier in the Vicinity of ATS-6, EOS 56, 408, 1975. (Also to be published by AIAA, 1975).

7. SPACE POWER SYSTEMS: IONOSPHERIC IMPACTS

(B. K. Ching, The Aerospace Corporation, El Segundo, Ca., 90245)

7.1 Introduction

Among the prospects for meeting our country's energy needs in the 21st century, perhaps the boldest, and certainly the most exciting, are the space power systems. As currently envisioned, huge power stations, each capable of serving a large city, would orbit at geosynchronous altitude and transmit space-generated microwave power to earth.

The impact of space power systems on the ionosphere, and conversely, of the ionosphere on the power systems, has not been treated with any depth in the many reviews that have been written on the power satellite concept (e.g., Glaser, 1977). Indeed, the microwave spectral region was selected for power transmission largely because of the atmosphere's transparency to such frequencies. Thus, it would seem that interactions between the beam and the medium would be of little or no consequence. Preliminary systems studies (Maynard et al., 1975), however, have identified potential ionospheric problems related to microwave transmission and vehicle exhaust products. Since such problems could have an impact on communications systems and other microwave transmission, they must be considered in the design and development of space power systems.

7.2 Microwave Transmission

Baseline designs of power satellites call for an average flux through the atmosphere of about 100 W/m^2 in a beam approximately 7 km in diameter. Only about a thousandth of one percent of the energy in the beam will be absorbed in the ionosphere, and thus there will be virtually no impact on the transmitted power level. On the other hand, 0.001% corresponds to 1 mW/m^2, which is of the same order of magnitude as the solar energy contained in wavelengths shorter than 100 Å (the extreme ultraviolet, EUV) that represents the major source of heat and ionization in the thermosphere and ionosphere. A comparison of solar EUV and microwave energy absorption rates in the upper atmosphere is shown in Figure 7.1. The microwave absorption profile does not reflect possible anomalous absorption, which might occur if the response of the medium to the additional energy input were to trigger further enhanced absorption and plasma instabilities.

Microwave and EUV radiations interact in completely different ways with the atmosphere. The microwave photons are much less energetic than the solar radiation and are unable to produce ionization or dissociation. In fact, the solar EUV photons interact initially with the atmospheric molecules and atoms, whereas the microwave radiation reacts with the ambient electrons. Despite this fundamental

difference, energy considerations alone suggest a potentially significant effect on the thermal state of the atmosphere-ionosphere system, at least in the vicinity of the microwave beam.

Concern over potential microwave-induced perturbations of the ionosphere may be attributed in large part to the results of recent ground-based radio-frequency ionospheric modification experiments (see the special issue of Radio Science, 9 (11), 1974). These experiments, conducted at input heater frequencies of a few to ten MHz, utilized power outputs of the order of a few tens of $\mu W/m^2$. Assuming that the level for interactions scales as the square of the frequency, microwave-induced ionospheric effects would be expected to occur at power levels similar to those being considered for the space power systems.

Although it is not presently known exactly what effects might be stimulated by high power microwave transmissions, some indication of what might result can be provided by the results of the recent RF ionospheric experiments. It is found that the energy absorbed from the heater beam by the ionospheric electrons raises the electron temperature. This electron heating results in a spatial redistribution of the electron density due to thermal expansion along the earth's magnetic field lines.

Since the electrons diffuse up field lines, the expansion results in a depletion in the electron density in the upper ionosphere. In contrast, in the lower ionosphere, the enhanced electron temperature lowers the rate of electron-ion recombination, resulting in an increase in the electron density. With increased power levels, plasma instabilities are excited that produce additional nonlinear absorption of the heater beam that results in the generation of field-aligned density irregularities. These irregularities give rise to various scattering phenomena, the outstanding example being spread-F.

Of course, any modification of the ionosphere by the microwave beam of the space power system will likely be localized in extent and thus will probably not have an impact on a global-scale. The absorption of energy is initially confined to the beam width (roughly 7 km), although subsequent interactions and the excitation of plasma instabilities will increase the sphere of influence beyond the limits set by the beam width. In the RF ionospheric heating experiments the recovery of the ionosphere to its normal state occurred fairly rapidly following the heater turn-off. For space power transmission systems, however, the power will be "on" for the lifetime of the satellite (perhaps 30-100 yrs.) and the beam position will be essentially fixed in geographic space (due to the use of

geosynchronous orbits). Thus, any perturbations induced by the beam will not be very effectively diluted as might be the case with a non-geostationary beam. Because of the coupling between the thermosphere and ionosphere, some changes may be induced in the neutral structure which would provide some "feedback" to the ionosphere. For example, enhancement of the neutral temperature within the beam would lead to changes in neutral composition, which, in turn, would affect the electron-ion recombination rate and hence the electron density. The heated column of air, which presents a discontinuity in the medium, could also give rise to atmospheric gravity waves. The wave features would be impressed on the ionized component.

The possible effects of microwave transmission in the thermosphere and ionosphere are summarized in Table 7.1. Whether any or all of these phenomena would occur as a result of space power systems is a question yet to be answered. A mitigating factor is that microwave frequencies are far-removed from ionospheric plasma frequencies. However, the high flux levels could be counter-active, particularly if they are sufficient to trigger instabilities in the plasma.

7.3 Vehicle Exhaust

The sheer size of a power satellite, 10^7-10^8kg, is indicative of the high level of vehicle activity that

will be needed to lift and assemble and then service and maintain a fleet of 50-100 power satellites. It is estimated that the equivalent of 10 shuttle launches per day will be necessary during the peak of the construction period (Williams, 1975).

Advanced heavy lift launch vehicles will probably expel only water vapor when at ionospheric altitudes (Beichel, 1974). However, the expulsion of large quantities of water molecules can produce significant changes in the local ionospheric structure, as was strikingly demonstrated during the launch of Skylab 1 (Mendillo et al., 1975). Water molecules interact efficiently with atomic oxygen ions, resulting in increased rates of electron-ion recombination. In the case of Skylab 1 there was a substantial depletion ($\geq$50%) of the total electron content that persisted for several hours and encompassed a region on the order of 2000 km in diameter.

Water vapor from booster engines could also affect the D-layer of the ionosphere, which is composed mainly of water cluster ions. The continual injection of water vapor in substantial amounts is likely to affect the ion formation rates and, hence, the properties of the layer.

In addition to the launch vehicles, which ascend only to low earth orbit, orbit transfer vehicles will be needed to tug the assembled satellites to geosynchronous altitude. These transfer vehicles require a high-performance propulsion system, with the likely candidate being the solar-

electric propulsion system. Such a system operates basically by the acceleration, and then expulsion, of a heavy, ionized metal. Preliminary studies have not indicated serious pollution problems in the ionsphere or magnetosphere; however, further investigation using realistic traffic models is necessary.

7.4 Conclusions

Considerations of ionospheric effects will no doubt influence decisions related to transmission power levels of space power systems as well as the propellants and propulsion systems necessary for placing the systems in orbit. Any modification of the ionosphere must be kept at levels tolerable to other uses, such as communications and radar systems. Realistic traffic models, including allowance for growth on an international basis, will be crucial for a sound assessment of the potential impacts. Since the power levels and the volume of space traffic will both be far greater than that encountered in present day operations, a vigorous program of theoretical and experimental investigation is required in order to provide the proper guidance in the design and development of space power systems.

7.5 References

Beichel, R., Propulsion Systems for Single-Stage Shuttles, Astron. and Aeron., 12 (Nov.), 32, 1974.

Glaser, P. E., Solar Power from Satellites, Phys. Today, 30, 30, Febr. 1977

Maynard, O. E., W. C. Brown, A. Edwards, J. T. Haley, G. Meltz, J. M. Howell and A. Nathan, Microwave Power Transmission System Studies, Vol. II, NASA-CR-134886, 1975.

Mendillo, M., G. S. Hawkins and J. Klobuchar, A Sudden Vanishing of the Ionospheric F Region Due to the Launch of Skylab, J. Geophys. Res., 80, 2217, 1975.

Williams, J. R., Geosynchronous Satellite Solar Power, Astron. and Aeron., 13, 46, Nov. 1975.

Table 7.1

Possible Effects of Microwave Propagation

IONOSPHERE

- Electron temperature increase
- Electron density decrease in D-region, increase in F-region
- Modification of electron energy distribution
- Anomalous absorption and heating, leading to field-aligned irregularities and radio scattering phenomena

THERMOSPHERE

- Neutral temperature increase
- Modification of relative composition
- Modification of airglow characteristics
- Excitation of atmospheric gravity waves

MUTUAL COUPLING EFFECTS

- Neutral composition affects the electron-ion recombination rate
- Ion density affects the neutral wind system
- Neutral winds and gravity wave structure affect the ion distribution

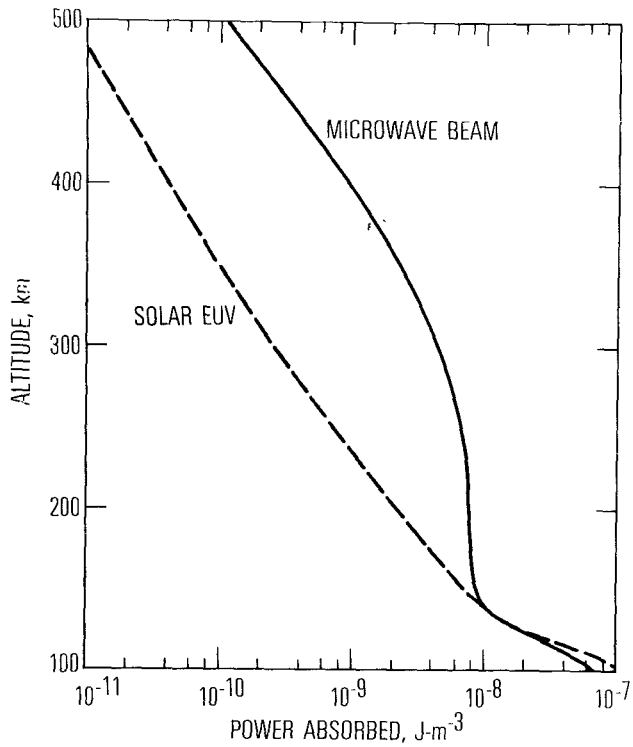


FIGURE 7.1 Daily average absorption of solar radiation at middle latitudes compared with the energy that would be absorbed from a microwave beam of uniform flux 100 W/m^2 at a frequency of 2.5 GHz.

8. DEEP EARTH INDUCTION STUDIES MAKING USE OF MAGNETOSPHERIC-IONOSPHERIC CURRENT SYSTEMS

(D. I. Gough, University of Alberta, Edmonton, Alberta, Canada)

8.1 Source Current Systems

Time-varying electric currents in the magnetosphere and ionosphere induce secondary currents in the solid earth. The fields of the internal currents can be used in the study of conductive structures within the earth. From the standpoint of such induction studies, the sources of the current systems external to the solid earth are in the magnetosphere and ionosphere. Useful current systems, in order of increasing period, include those of magnetic pulsations, magnetospheric substorms, the daily geomagnetic variation, and the ring-current decay field D_{st} in magnetic storms.

An important parameter in induction studies is the skin depth, the depth at which a field of period T falls to 1/e of its incident amplitude. In convenient units the skin depth is $30.2\times \sqrt{T\rho}$ km where T is expressed in hours and the resistivity ρ is expressed in ohm-m. Resistivities of rocks in the crust and upper mantle range from less than 1 ohm-m to 10^4 ohm-m. Source fields can be characterized: (a) by their spectra; that is by the periods T at which they apply useful energy, and the corresponding effective depths of penetration in the earth; (b) by their spatial geometries. The spatial geometries

of the magnetospheric and ionospheric sources can be complicated and are different for the individual sources.

Magnetic pulsations produce fields with periods up to a few minutes and with amplitudes of a few nanoteslas (nT). They are useful mainly in magnetotelluric study of the sedimentary rocks of the crust to depths of a few kilometers. These sources can be highly localized in the ionosphere/magnetosphere.

Magnetospheric substorms are very important in carrying out induction studies by magnetometer arrays and by the magnetotelluric method. In low and middle latitudes substorms provide transient magnetic fields of order 20 nT in the period range $15 < T < 150$ minutes. The geometry of the currents is complicated and is still undergoing active exploration (Rostoker, 1972, Kisabeth and Rostoker, 1977). Present indications are that a westward ionospheric current in the auroral oval joins magnetic field-aligned currents at its ends, which connect the ionospheric segment to a closure in the magnetosphere. This transient current system can be regarded as superimposed upon steady-state "soleonoidal" currents involving a southward ionospheric current in the morning sector auroral zone and a northward ionospheric current in the evening sector, with sheets of field-aligned current at the northern and southern limits.

In middle and low latitudes the magnetic storm source field may change only smoothly and moderately across

a magnetometer array, and anomalous fields within the array can be identified with induced currents and in some cases can be quantitatively modeled. However, at the present time, in latitudes near the auroral zone the complicated source-field geometry encourages the use of arrays for the study of the external current systems rather than for induction studies of the solid earth.

The daily variation dynamo (Matsushita, 1975) and its associated magnetic field produce a travelling wave with respect to the solid earth. At middle latitudes the S_q (quiet day) variation has spectral peaks at periods of 24, 12, 8 and 6 hours; on distributed days shorter periods are added to the spectrum. Even in a medium of resistivity 10 ohm-m, typical of the upper mantle, the skin depth at T=24h is 470 km, so the daily variation fields penetrate deeply into the upper mantle. At the dip equator the concentration of S_q currents in the equatorial electrojet provides a narrow, linear localized current source useful in induction studies at crustal and uppermost mantle depths.

The spectrum of the D_{st} field, associated with the decay of the ring-current in the late stages of magnetic storms, contains useful energy at periods up to 3 days and has been used in induction studies to depths of order 700 km (Lahiri and Price, 1939; Matsushita, 1975).

8.2 Techniques of Induction Studies

Data from the world-wide network of geomagnetic observatories are used in the study of the radial variation of electrical conductivity, the principal planetary-scale feature of the earth's electrical structure. The source fields are principally the daily variation, D_{st}, and fields of longer periods (Bailey, 1973).

In the magnetotelluric method, sources such as magnetic pulsations and magnetospheric substorms provide input in the period range 10^{-1} to 10^{4} seconds or more, to a single station at which two orthogonal horizontal components of magnetic field, $H_x(T)$ and $H_y(T)$, and the corresponding electric field components $E_x(T)$ and $E_y(T)$ are recorded simultaneously. The magnetic field detectors are commonly coils or fluxgate magnetometers; the electric fields are detected as potential differences between separated grounded electrodes. At the surface of a half-space made of isotropic plane layers, the apparent resistivity at a particular period T is

$$\rho_a = 0.2\ T \left|\frac{E_x}{H_y}\right|^2 = 0.2\ T \left|\frac{E_y}{H_x}\right|^2$$

(Cagniard, 1953) and ρ_a (T) can be fitted to layered models of ρ(depth) which are non-unique but may approximate the true structure. The real earth seldom resembles a set of isotropic plane layers, and in general magnetotelluric data

define a resistivity tensor whose maximum and minimum values are interpreted in terms of resistive structure. A typical depth range for magnetotelluric soundings is from the surface to 100 km depth. Contemporary magnetotelluric techniques are reviewed by Hermance (1973).

In magnetometer array studies, an array of three-component magnetometers records magnetic fields in the period range 10-1440 minutes, at a number of points on the earth's surface. In studies reported to date, the number of magnetometers has ranged from four to eight, often placed along a straight line, or from 15 to 46 covering an area. The area covered was 10^6 km^2 in the largest array yet used (Camfield et al., 1971). Large, two- dimensional arrays record a single magnetic event over the whole area, and highly conductive structures can be mapped in terms of amplitudes and phases of spectral peaks supplied by the source field. In some cases the fields of internal, induced currents can be separated from those of the external source, and the internal fields can be fitted to calculated fields from induction in model structures (Porath et al., 1970; Jones and Pascoe, 1971). In all cases, anomalous conductors can be mapped with a precision limited by the magnetometer spacing and can be assigned maximum depths.

An alternative approach to interpretation expresses the electromagnetic response of the earth, at

each station, as a transfer function from the horizontal components (in which the external source field is largest) to the vertical component (in which the fields of internal, induced currents often predominate). Such transfer functions, calculated at several periods in which several variation events supply the inputs, allow adjacent small linear magnetometer arrays to be combined into a partial approximation to a two-dimensional array.

The problems of the quantitative interpretation of magnetometer arrays are severe (Gough, 1973a); one of the most serious is discussed here. This most intractable difficulty affects all types of investigation of the earth by electromagnetic induction, and therefore deserves mention. Anomalous currents in the region under study often flow in a long, highly conductive channel in the crust or upper mantle which connects large, unknown conductive volumes of the earth elsewhere to the area being studied. In such cases the induction may occur mainly in the large, distant conductors where the source field may differ greatly from that recorded by the array. Such channelled-current anomalies are more numerous than those in which induction occurs mainly in the structures covered by the array, and are extremely difficult to interpret beyond the point of locating the conductive channel. An example is shown in the next section.

8.3 Some Representative Results

The principal feature of the electrical resistivity within the Earth is a very rapid fall to values below 1 ohm-m in the depth range 400-800 km (Banks 1972). This increase in conductivity is certainly related to temperature but may be related also to pressure and to seismically observed phase changes at depths near 390 km and 650 km. Since the skin depth at ρ = 1 ohm-m for the 24 hour term in the daily variation is only 150 km, the fields of external current systems penetrate negligibly beyond a depth of $\sim$1000 km. In effect, electrical sounding using induction by magnetospheric-ionospheric currents is limited to the depth range 0-800 km. In the top 100 km the resistivity varies through four orders of magnitude, from 0.25 ohm-m for sea-water to 10^4 ohm-m for dry crystalline crustal rock.

Local anomalies in geomagnetic variation fields represent currents induced in conductive structures within 800 km of the earth's surface. Many of them can be classified as continental-edge anomalies, subduction-zone anomalies, the anomalies of western North America, rift valley anomalies or crustal anomalies. A review with many references up to 1973 has been given elsewhere (Gough, 1973b). Only later references are quoted here.

Large anomalies, mainly in the vertical components of fields with periods $1h \leq T \leq 24h$, appear near continental edges. Largely, and in some cases wholly, these can be

ascribed to induced electric currents in the sea-water. There may also be a contribution due to a step in the conductive structure of the upper mantle below the continental edge, but existing information is indecisive on this point.

Prominent conductivitiy anomalies in zones of active subduction are known from magnetometer array studies in Honshu Island, Japan, and in the Peruvian Andes. These probably result from the ascent of molten material with ionic conductivity, from partial melting of the downgoing plate. No such conductor is found beneath the Andes of central Chile.

Large magnetometer arrays have revealed complex conductive structures in the upper mantle under western North America: the results are summarized in Figure 8.1. Highly conductive mantle material estimated to be at least 100 km thick underlies the Basin and Range Province, with still greater development below the Wasatch Fault Belt and the Southern Rockies. North of the line XX (near 43°N in the Figure) the conductive layer is much thinner ∿10-20 km for ρ = 5 ohm-m. The electrical conductivity, the heat flow from beneath the crust, and seismic wave velocities in the upper mantle are correlated (Gough, 1974), and are all consistent with anomalously high temperatures producing partial melting in the upper mantle.

Local anomalies associated with induced currents in the upper mantle and crust have been discovered beneath the eastern Rift Valley of Kenya and have been studied by a magnetometer array and by the magnetotelluric method. High conductivities rise to crustal depths in Iceland and it is highly probable that an array study across a mid-ocean ridge would contribute to identifying the location of molten material. The technical problems of ocean-floor magnetometer arrays have yet to be overcome, however.

Local induction anomalies within continents are sometimes associated with sedimentary basins (Porath and Dziewonski, 1971) in which the conductivity is probably due to saline pore-water. In other cases long, very narrow anomalies of the channelled-current type seem to be related to fracture zones, in which graphite is commonly present. An example of a crustal anomaly of this type is illustrated in Figure 8.2. This North American Central Plains conductive structure was discovered by a magnetometer array near the Black Hills of South Dakota and it was suggested that the currents might be channelled by graphite in metamorphic rocks of the basement (Camfield et al., 1971). A later array study (Alabi et al., 1975) mapped the structure from the Canadian Shield of Saskatchewan, where graphite sheets are common in exposed metamorphic rocks in a large fracture zone, to another major fracture zone 1800 km to the south in Wyoming. The

conductive structure is believed to lie in a great fracture zone which may mark a continental collision 1.7 billion years old. Figure 8.2 shows how the induced currents perturb the field of a magnetospheric substorm, in the Fourier spectral term of period 68 minutes.

Further references to the literature of induction studies using the current systems of the magnetosphere and ionosphere can be found in recent reviews by Frazer (1974) and Gough (1973b, 1974).

8.4 References

Alabi, A. O., P. A. Camfield and D. I. Gough, The north American central plains conductivity anomaly, Geophys. J. R. Astr. Soc., 43, 815, 1975.

Bailey, R. C., Global magnetic sounding - methods and results, Phys. Earth Planet. Interiors, 7, 234, 1973.

Banks, R. J., The overall conductivity distribution of the Earth, J. Geomagn. Geoelectr., 24, 337, 1972.

Banks, R. J. and P. Ottey, Geomagnetic deep sounding in and around the Kenya Rift Valley, Geophys. J. R. Astr. Soc., 36, 321, 1974.

Cagniard, L., Basic theory of the magnetotelluric method of geophysical prospecting, Geophysics, 18, 605, 1953.

Camfield, P. A., D. I. Gough and H. Porath, Magnetometer array studies in the northwestern United States and southwestern Canada, Geophys. J. Roy. Astr. Soc., 22, 201, 1971.

Frazer, M. C., Geomagnetic deep sounding with arrays of magnetometers, Rev. Geophys, and Space Phys., 12, 401, 1974.

Gough, D. I., The interpretation of magnetometer array studies, Geophys, J. Roy. Astr. Soc., 35, 83, 1973a.

Gough, D. I., The geophysical significance of geomagnetic variation anomalies, Phys. Earth and Planetary Int., 7, 379, 1973b.

Gough, D. I., Electrical conductivity under western north America in relation to heat flow, seismology and structure, J. Geomag. Geoelect., 26, 105, 1974.

Hermance, J. F., Processing of magnetotelluric data, Phys. Earth and Planetary Int., 7, 349, 1973.

Jones, F. W. and L. J. Pascoe, A general computer program to determine the perturbation of alternating electric currents in a two-dimensional model of a region of uniform conductivity with an embedded inhomogeneity, Geophys. J. R. Astr. Soc., 24, 3, 1971.

Kisabeth, J. L. and G. Rostoker, Modeling of the three dimensional current system associated with magnetospheric substorms, Geophys, J. Roy. Astr. Soc., 49, 1977.

Lahiri, B. N. and A. T. Price, Electromagnetic induction in non-uniform conductors and the determination of the conductivity of the Earth from terrestrial magnetic variations, Phil. Trans. Roy. Soc. London, Ser. A., 237, 509, 1939.

Matsushita, S., Morphology of slowly-varying geomagnetic external fields, Phys. Earth and Planetary Int., 10, 299, 1975.

Porath, H., Magnetic variation anomalies and seismic low-velocity zone in the western United States, J. Geophys. Res., 76, 2643, 1971.

Porath, H. and A. Dziewonski, Crustal resistivity anomalies from geomagnetic deep-sounding studies, Rev. Geophys. and Space Phys., 9, 891, 1971

Porath, H., D. W. Oldenburg and D. I. Gough, Separation of magnetic variation fields and conductive structures in the western United States, Geophys. J. R. Astr. Soc., 19, 237, 1970.

Rostoker, G., 1972. Polar magnetic substorms, Rev. Geophys. and Space Phys., 10, 157-211.

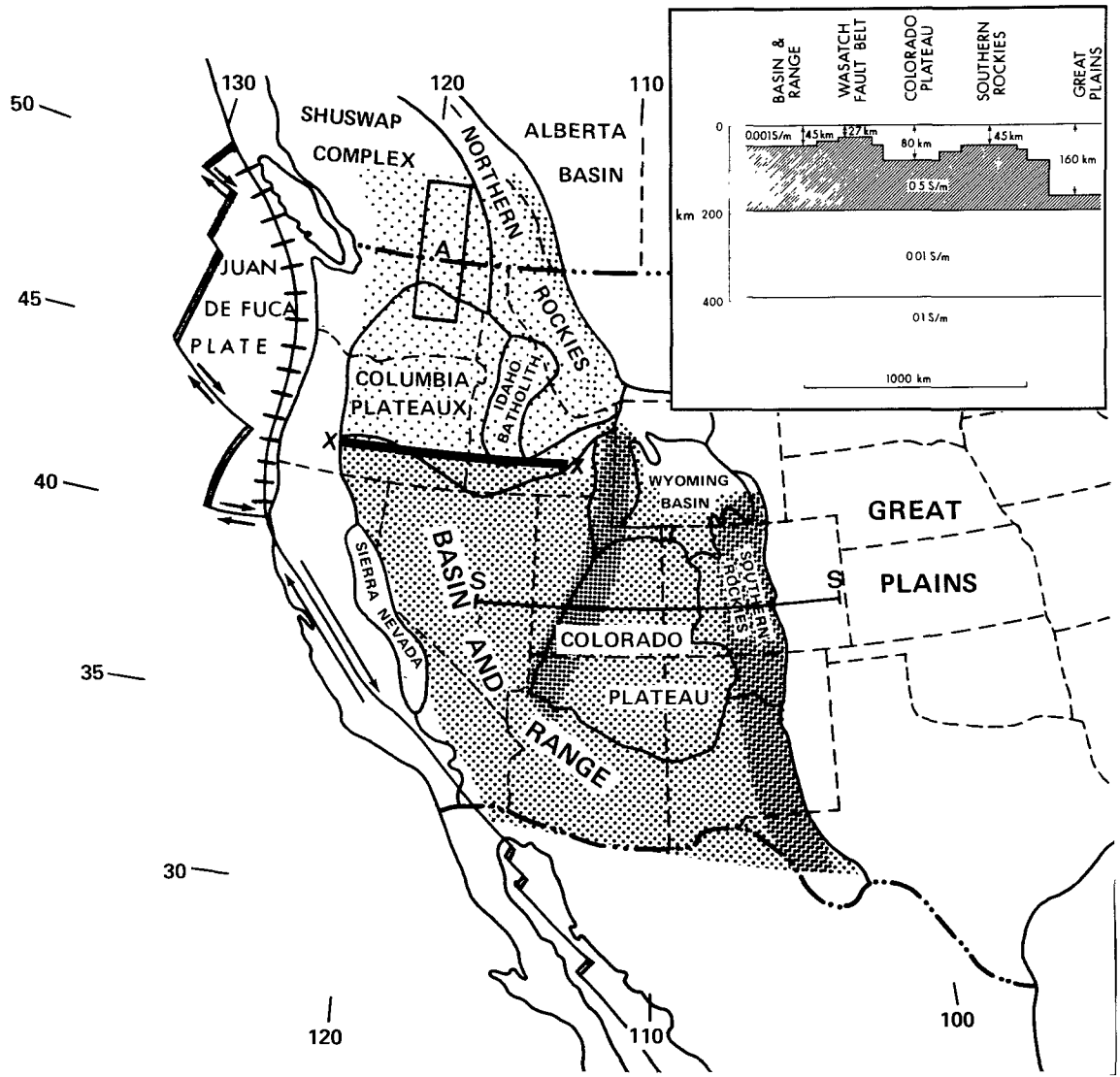


FIGURE 8.1 Distribution of anomalously high conductivity in the upper mantle of western North America as mapped by three large magnetometer arrays. The density of stippling indicates variations of the thickness-conductivity product in an upper mantle layer. The inset shows a vertical section through a model by Porath (1971) which fits observed fields along the east-west profile SS at 38°N, with conductivities expressed in siemens per meter (1 siemens unit ≈ 0.94 ohm).

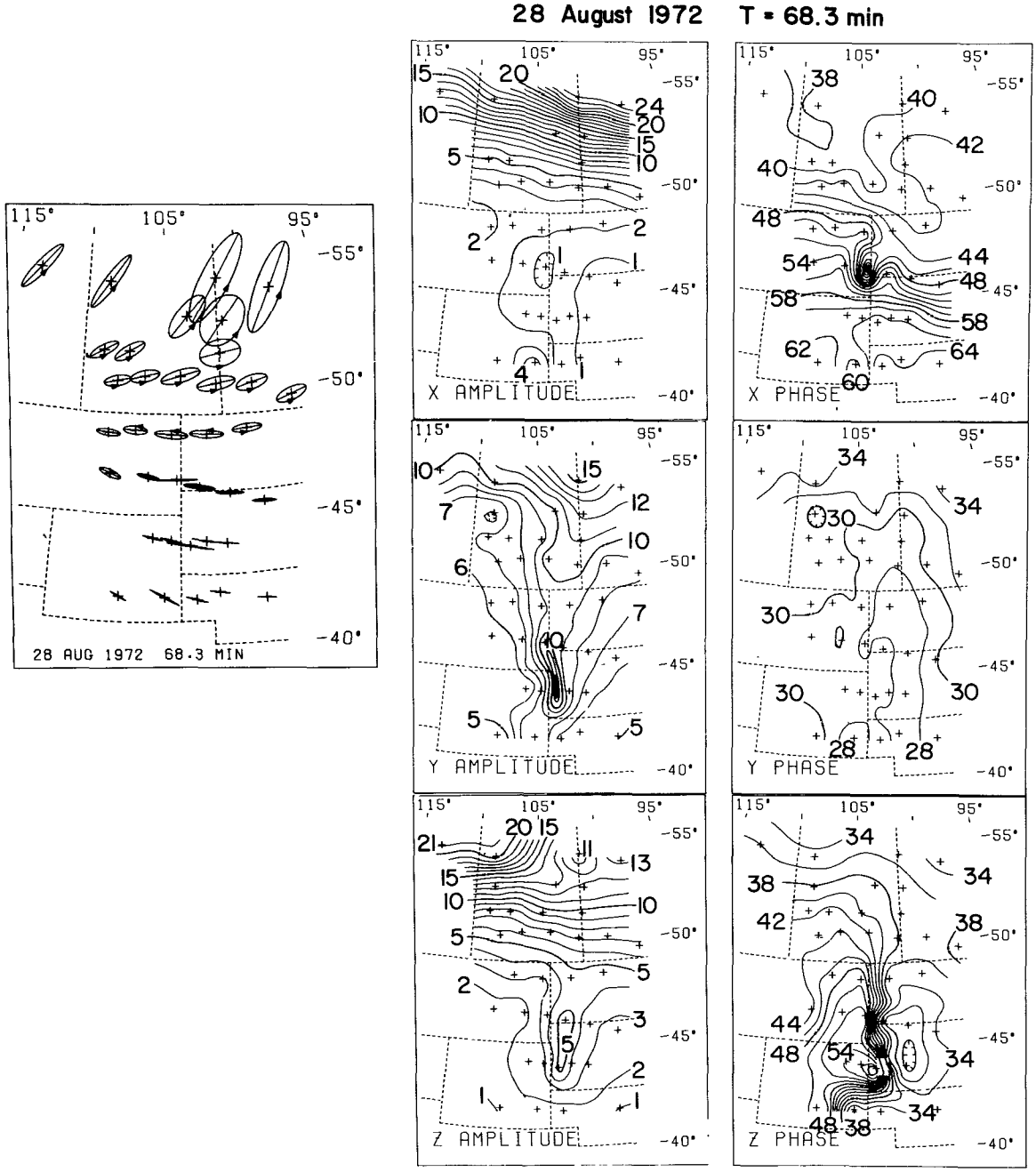


FIGURE 8.2 Perturbation of the field of a magnetospheric substorm by the North American Central Plains conductivity anomaly. In each map the crosses indicate the locations of magnetometers in an array. The polarization of the horizontal field at a period of 68.3 min. is mapped on the left. The other six maps show amplitudes and phases of the Fourier transforms of the substorm field components at the spectral period of 68.3 min. The northern border of the array was close to the ionospheric part of the source current, as shown by the X amplitude and polarization maps. The conductivity anomaly is best seen in Y and Z amplitudes and Z phase. Amplitudes are in nT, phases in minutes. After Alabi *et al.* (1975).

9. CONSIDERATIONS OF TELLURIC CURRENT EFFECTS ON PIPELINES

9.1 Pipelines at Middle to Low Latitudes

(A. W. Peabody, Ebasco Services, Inc., New York, N.Y., 10006)

9.1.1 Introduction

The rapid expansion of the transmission pipeline industry in the decades following the nineteen thirties caused increasing attention to be devoted to controlling the corrosion of pipeline steel. One of the problems that faced pipeline corrosion engineers involved an unusual form of variable "stray" direct currents and potentials on the pipelines. These could not be correlated with stray direct currents and potentials of a man-made nature. The phenomenon became particularly apparent on long cross-country electrically-continuous steel pipes which normally had a protective coating of an insulating material used to assist in the mitigation of corrosion effects. These pipelines were, in effect, long insulated electrical conductors subject to inductive effects from variable magnetic fields (see, for example, Gideon et al., 1970). The nature of the non-man-made stray effects on pipelines was found to be similar to effects on long distance telegraph lines during magnetic storms associated with severe sunspot activities. As a result, any otherwise-unexplainable stray effects observed on pipelines by pipeline corrosion engineers came to be known as telluric current effects.

9.1.2 Corrosivity of Telluric Currents

The frequency and severity of occurrence of telluric currents are quite erratic with, often, long "quiet" periods when there is no significant effects at all. Further, the pipeline locations where the telluric currents discharge to earth (with the resulting possible consumption of pipeline steel) tend to move from place to place with no long-term concentration of a corrosive effect at any one location. Both of these effects arise from the variable nature of geomagnetic activity (see, for example, Campbell, 1976).

The pipeline corrosion effect arises from the discharge of current from the pipe to the ground through small holes in the non-conducting protective coating of the pipe. These holes develop from damage during the laying of the pipe. A short-term study of the peak values of induced currents on pipelines in several states made by Gideon et al. (1970) showed typical values in the range of one ampere. Using extreme values of the exposed pipe area in the pipe coating, they found a negligible 10^{-7} m/yr. corrosion for 0.5-0.75m diameter pipes (see also Campbell, 1977).

In addition to their erratic occurrence pattern, telluric currents can reverse direction from time to time as well. This also operates to reduce their corrosivity

compared to a steady, uninterrupted, unidirectional flow of direct current. McCollum and Ahlborn (1916) reported the corrosion rate as a function of period and their results are shown in Fig. 9.1 (this work was carried out in connection with the operation of street railway systems). It can be seen from Fig. 9.1 that for telluric currents with periods in the range of 5 mins. to several hours (typical of geomagnetic variations) that the conclusions of Gideon et al. (1970) would be reduced by ∿75%.

There has been some evidence to support the possibility of a concentration of telluric current discharge at a specific location where subsequent severe damage will result. One such instance, studied by the writer's associates in connection with an early 1940's examination of corrosion damage on a new pipeline system extending from the Atlantic to the Pacific across the Isthmus of Panama, found that currents reaching values in the order of 30 amperes would at times flow from ocean to ocean, with the direction of flow changing from time to time. Current discharge to earth occurred at the ocean terminals with corrosion failures developing on some terminal facilities prior to being placed in service. The currents could not be related to man-made effects.

Although the writer does not have specific documentation to the effect, he has been advised that there has

been at least one other location where known telluric current discharge effects at an ocean pipeline terminal have caused corrosion problems. From this very limited knowledge base, it does appear reasonable to expect a concentration of the corrosion problem where a pipeline terminates at a location where the resistance to the flow of discharge currents would be very low. An ocean terminal is most apt to fulfill this condition.

Because of the erratic nature of the telluric current effect, essentially no active corrective measures, such as the active "cathodic protection" techniques (Peabody, 1967), are taken that are specifically directed towards the effects of telluric current activity. In the rare instances where a problem is identified (such as at an ocean terminal on a pipeline particularly subject to telluric effects), the telluric current can be drained from the pipe to earth through a separate ground connection designed for the purpose.

9.1.2 Telluric Currents and Pipeline Corrosion Control

When making field measurements to ascertain the level of cathodic protection on a pipeline, the cognizant corrosion engineer must make observations of pipeline-to-earth potential and direct current flow measurements accurate to at least ±5 millivolts and ±100 milliamps, respectively. The inability to make observations to such levels of accuracy could result in misleading or completely meaningless data. Telluric current effects, when active,

have the capability of completely screening the true steady-state conditions. The writer has observed both pipeline potentials and currents of a telluric current origin which were several times the magnitude of the normal steady-state conditions.

It is obvious that without advance warning of possible telluric current activity, pipeline corrosion survey work could be scheduled and performed at a time when the results obtained are meaningless insofar as interpretation of normal conditions are concerned. Recognizing this, the NOAA Space Environment Services Center in Boulder, Colorado, initiated regular publication of predictions for pipe-induced current activity (Campbell and Doeker, 1974). Pipeline corrosion engineers can obtain this information (updated daily) by telephone in order to plan their survey activities. Additionally, in order to make it possible to evaluate past incidents, the Space Environment Services Center periodically publishes past activity data in Materials Performance, a publication of the National Association of Corrosion Engineers.

9.1.3 Summary

The effect of telluric currents on pipelines at low and middle latitudes is more of a nuisance than a serious problem for corrosion. The nuisance arises from the telluric current interference with normal pipeline

corrosion survey engineering work. This problem is a sufficient nuisance to have caused the implementation of a regular service for predicting pipe-induced current activity.

9.2 The Alaskan Pipeline*

(W. H. Campbell, U. S. Geological Survey, Denver, Colo., 80225)

9.2.1 Introduction

Associated with every occasion that society finds reason to build long lines of conducting material there is a reawakened interest in the induced currents in these lines associated with geomagnetic disturbances (Burbank, 1905). Typically such interest mounts near the 11-year maxium in geomagnetic activity (Bartels, 1963).

The Alaska oil pipeline, a colossus in size with respect to those within the lower 48 states, traverses the auroral regions, regions in which the geomagnetic disturbance energy could be at least 100 times that troubling engineers in the mid-U.S. (Campbell, 1973). The principal geomagnetic disturbances seen at high latitudes arise from currents within the ionized layers of the atmosphere about 100 km above the earth. The surface manifestations of these ionospheric currents are electric and magnetic fields which induce current flow in the conducting earth. Such current is channeled into regions of the highest conductivity, either of geologic origin or of man-made origin, as in the case of a pipeline.

The Alaska oil pipeline (Raymer, 1976) is essentially a long, surface-grounded conductor extending from about 69.3° geomagnetic latitude at Prudhoe Bay on

---

*Abstracted by the editor from Campbell (1977).

the Arctic Ocean, generally southward near the region of auroral zone maximum, to about 61.6° geomagnetic latitude at Valdez on the Pacific Ocean (Fig. 9.2).

Assuming the pipe steel to have a resistivity of about 1.4×10^{-7} ohm-meters, the length to be 796.4 miles (1.28×10^{3}km), the mean diameter 48 inches (1.22 m), and the mean wall thickness 0.512 inches (1.30 cm), then the end-to-end resistance of the pipeline is 3.6 ohms. This assumes that the insulating connections at the 12 pumping stations are shorted by conducting jumpers. The resistance per unit length of pipe, R_m, is thus 2.81×10^{-6} ohm/m.

9.2.2 Effects of the Auroral Currents

Because of the high pipe conductivity with respect to the ground to which it is electrically connected via zinc grounding cables at regular intervals (∿150m) and because the pipeline is long with respect to the scale of the inducing fields, the currents flowing in the pipeline will be just of that magnitude necessary to eliminate the induced electric field parallel to the pipeline. The Alaskan pipeline route runs generally in a geomagnetic north-south direction for the northern and southern sections and east-west in the central section (Fig. 9.2). The ducted induced electric fields that are of concern are those parallel to this pipeline route. At maximum they would represent the horizontal electric component $E_{Hor} = E_{NS} \csc \theta$, where θ, the angle

between the horizontal electric field component and geomagnetic North, varies between about 1.0 and 1.6.

The scale to the right side of Fig.9.3 gives the computed equivalent pipeline current for a geomagnetic east-west electric field parallel to the pipeline (left-hand scale) as a function of the period of the electric variations (Campbell, 1977). To obtain the equivalent horizontal components such values should be multiplied by csc θ. The electric field values are deduced from analyses of magnetic field variations measured at several places along the pipeline route and then using a three-layer model of the earth to compute the induced electric field. The magnetic variation studies are carried out for different levels of geomagnetic disturbance as represented by the geomagnetic disturbance index Ap.

The statistics of occurrence of Ap over several solar cycles show that 50% of the index values are $\sim$7 or less (see, for example, Campbell, 1977). An Ap value of 7 would correspond to currents near a 1-hr. period of $\sim$2.7 amps for the sample conditions of Fig. 9.3. Multiplied by csc θ = 1.07 (for the average pipeline angle with respect to the auroral current system) a value of $\sim$2.9 amps is obtained. This value represents a daily average current of $\sim$1 amps (the average daily pipeline current level is about 1/3 the magnitude at the maximum auroral latitude

and about 1/5 the magnitude for an average of all stations along the pipeline). Individual surges in current amplitude between about 30 min. and 2 hr. could be about 5 to 10 times the size indicated in Fig. 9.3 (see Campbell, 1977), corresponding, the above example, to individual fluctuations reaching $\sim$14 amps.

Similar calculations for small and large magnetic disturbance conditions (Ap $\geq$ 52 and Ap $>$ 150, values that can statistically occur approximately once every two months and every two years; see Campbell, 1977) give current surges $\geq$ 95 and $\gtrsim$ 280 amps, respectively. Such large transient currents could cause problems at electronic monitoring and control points along the pipeline.

The increased pipe area of the 1.2m Alaska pipe would reduce corrosion to $\sim$0.4 to $\sim$0.7 the values quoted in Section 9.1.2 from Gideon et al (1970) for a one-ampere current. Increasing the current by a factor of ten would mean that the comparable Alaskan corrosion should be $\lesssim 10^{-6}$ m/yr., a truly negligible amount for a 1.3 cm wall pipe thickness. In addition, the computed corrosion rate should be reduced by the appropriate factor shown in Fig. 9.1.

9.2.3 Summary

Corrosion in the Alaskan pipeline from the auroral current system appears to be a negligible concern. However, it appears that large transient currents can be

established under geomagnetic disturbance conditions. These currents could greatly disturb, or prevent, corrosion survey engineering studies on the pipeline. They could also produce severe problems for the pipeline monitoring and control electronics.

9.3 References

Bartels, J., Discussion of time-variations of geomagnetic activity indices Kp and Ap, 1932-1961, Ann. Geophys., 19, 1, 1963.

Burbank, J. E., Earth currents and a proposed method for their investigation, Terr. Mag., 10, 23, 1905.

Campbell, W. H., Spectral composition of geomagnetic field variations in the period range of 5 min. to 2 hr. as observed at the earth's surface, Radio Sci., 8, 929, 1973.

Campbell, W. H., An analysis of the spectra of geomagnetic variations having periods from 5 min to 4 hrs., J. Geophys. Res., 81, 1369, 1976.

Campbell, W. H., Induction of auroral currents within the surface conductivity anomaly presented by the Alaskan pipeline, Geophysics, submitted, 1977.

Campbell, W. H., and R. B. Doeker, Pipe induced current activity indices, Materials Performance, 13, 9, 1974.

Gideon, D. N., A. T. Hopper, and R. E. Thompson, Earth current effects on buried pipelines; analysis of observations of telluric gradients and their effects, Am. Gas Assn., cat. no. L30570, 77 pp., Apr., 1970.

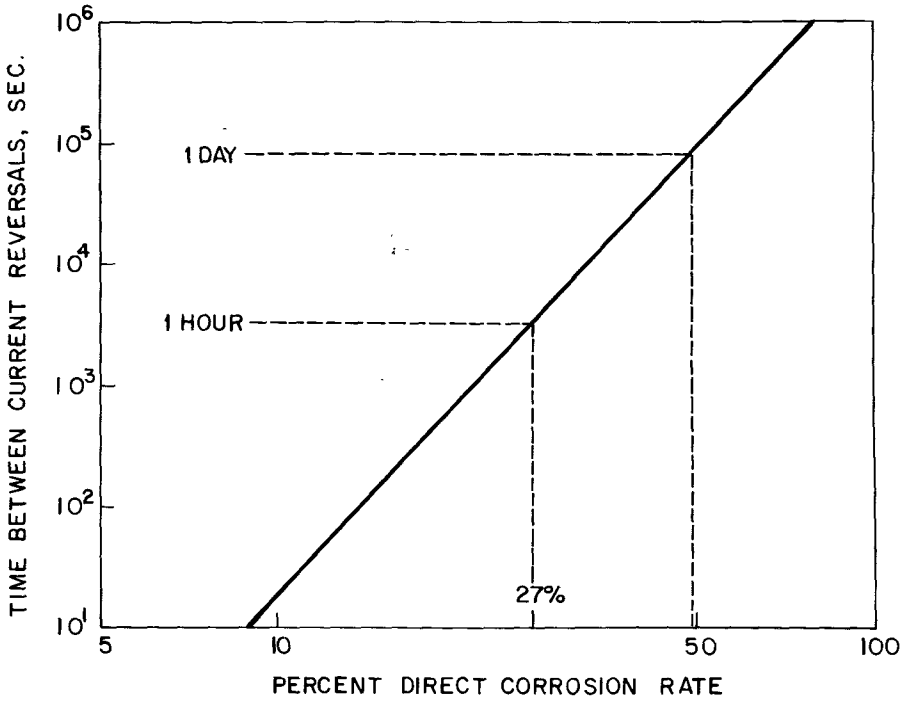


FIGURE 9.1 Percentage of direct current corrosion rate as a function of the period of the induced current (from McCollum and Ahlborn, 1916).

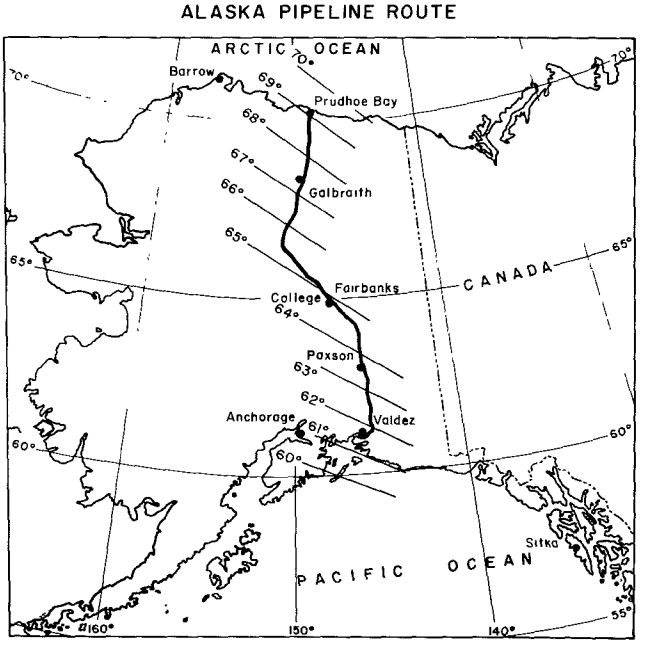


FIGURE 9.2 Alaska oil pipeline shown as a dark line connecting Prudhoe Bay to Valdez. Geomagnetic latitudes along this route are indicated by line segments every degree from 60° to 70°. Location of geomagnetic observatories at Barrow, College, and Sitka are appropriately marked.

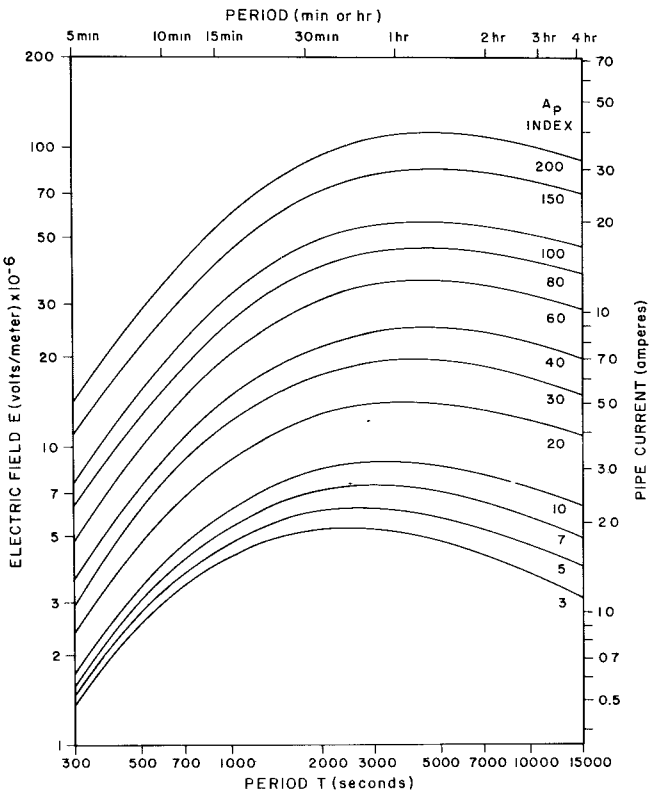


FIGURE 9.3 Electric fields (10^{-6} V/m) as a function of period T (sec) obtained from a layered earth conductivity model and a spectral analysis (four hour samples) of high latitude geomagnetic disturbances. Separate curves represent values for geomagnetic activity level indices, Ap, shown at the right. On the right margin is the equivalent pipeline current, I (amperes). Momentary surges of field and current at the time of maximum disturbance reach amplitudes 5 to 10 times the magnitudes shown in this figure.

10. PROBLEMS OF MAGNETIC FLUCTUATIONS IN GEOPHYSICAL EXPLORATION

(M. S. Reford, Geoterrex Ltd., Ottawa, Ontario, Canada)

10.1 Introduction

In exploration geophysics, magnetic surveys are conducted to obtain information about subsurface rocks. Most of the surveys involve measurements of the total intensity of the earth's magnetic field across the survey area. Of necessity, the measurements include time fluctuations of the field, and one of the problems that must be dealt with is the separation of the unwanted time variations from the desired spatial variations. The nature and seriousness of the problem, and its solution, depend on the nature of the geophysical survey.

Most magnetic surveys are made with airborne magnetometers, and probably more than one million line-kms are flown every year all over the world. The most common objectives of a survey are to assist geological mapping and mineral exploration. Typically, a survey may involve flying lines one km apart, measuring the field intensity in units of one gamma (1γ = 1 nanotesla), at intervals of one second, and compiling the results as contour maps with a basic contour interval of ten gammas. Greater precision than this is required in petroleum exploration surveys, which are aimed at determining the depth to the magnetic basement rocks

(buried beneath several kms of almost non-magnetic sedimentary rocks) and at detecting local structures on the basement rock surface. High-resolution magnetometers measuring in units of 0.01 gammas are often used for such surveys, and local anomalies of one gamma relief may be important.

Another type of survey is common over the oceans, where the structure of the oceanic crust is revealed by the magnetic anomaly patterns recorded in data obtained from towing magnetometers behind ships, and sometimes near the ocean bottom. Finally, there are ground magnetometer surveys, usually restricted today to local detailing of individual mineral prospects.

Figure 10.1 shows a sample magnetic contour map from a high-resolution airborne survey in the Canadian Arctic. The contour interval is one gamma, with half-gamma contours in places, and the survey was flown with east-west lines one mile apart and north-south tie lines three miles apart. Magnetic surveys such as this are interpreted in geological terms, both from contour maps, and from profiles along the survey lines.

10.2 Acquiring Survey Data

The problems of time variations may be tackled at several stages of a survey. For a start, the problems are minimized by collecting data which includes the least

possible effects of fluctuations. Two very different approaches are used at present. One of these approaches is the measurement of the vertical magnetic gradient as well as the total field intensity. Gradiometers, measuring the difference in intensity between two high-resolution magnetometers, have been in regular survey use for ten years. Both magnetometers are affected equally by time fluctuations, leaving the measured gradient unaffected. However, the resolution of such a gradiometer is a function of the vertical distance between the sensors, which cannot be increased much beyond 70m without introducing serious operational problems; i.e., noise effects from relative movement between the sensors in turbulent air. Within these limits gradiometer measurements solve the problem of time fluctuations, but do introduce other difficulties in its place. One of the major of these is the increased costs. Thus, gradiometer surveys represent a very small proportion of the total volume of magnetic survey work presently being done, and have only partially replaced total intensity surveys, even in petroleum exploration.

A second approach is to use a fixed, ground-based magnetometer to record time fluctuations, and to accept survey data only when these fluctuations fall within specified limits. Since variations which are linear for a number of minutes can be corrected in leveling the data, the

specifications are keyed to short periods. For instance, a specification in high-resolution petroleum surveys might be to reject data obtained when the ground magnetometer trace showed departures exceeding 2 gammas from any chord two minutes long. This type of specification works well in most parts of the world, but can be very difficult to apply where irregular fluctuations are common (for example, near the auroral zones) or in areas affected by the equatorial electrojet. If the application of such a specification is too rigid, an aircraft can wait forever on the ground for the ideal conditions which never arrive. It is better to obtain data, even though compromised, and re-fly certain lines to check for possible errors by comparison of results from both flights.

A third approach might be suggested--to schedule surveys for periods when fluctuations are forecast to be minimal. Unfortunately, this is not very practical. Exploration and budget flows cannot be modified to fit the sunspot cycle. Also, many surveys involve a data collection period of one to three months, which is usually fixed by the optimum flying conditions. Forecasting of magentic fluctuations is no more certain than forecasting weather, and an unpredicted period of several weeks excessive magnetic activity is simply one of the hazards of working near the auroral zone.

10.3 Processing Survey Data

Having obtained the survey data, fluctuations must be removed, as far as possible, in the subsequent data processing. Two approaches are generally used. First, the fixed ground magnetometer measurements may be subtracted from the airborne magnetometer measurements. For success, this pre-supposes certain conditions. Obviously, time synchronization between ground and mobile units must be good, and the ground magnetometer must be carefully sited to avoid man-made interference, such as trucks moving near the sensor. Second, the fluctuations at both places must be nearly identical. This makes it difficult, if not impossible, to apply such a subtraction in many situations. The survey area may be several hundred kms away from the base station where the ground magnetometer is established. Differences in the earth electrical conductivity at the base station and at the survey points are probably the most critical factor. For instance, fluctuations over resistive basement rocks are generally stronger and contain higher frequency components than those over a nearby sedimentary basin, or ocean. Additional complications occur along coastlines, where the conductivity contrast can increase the amplitudes of the fluctuations relative to points off the shore or island (Vacquier, 1972).

At the present time, it does not seem practical to predict how successful subtraction may be in all circumstances,

nor how many ground magnetometers would be needed at what locations for optimum results. Establishing independent recording ground magnetometers at remote locations can be very expensive. Hence, the method is applied usually on a limited and empirical basis, and not in areas such as the auroral zones, where it is likely to introduce more problems than it solves. As a rule of thumb in middle latitudes, ground stations may be located so that the mobile magnetometer is not taken more than 200 kms away. The effect of subtraction is then tested on the mobile data: if recognizable fluctuations are removed, or at least reduced, by subtraction, then it may be applied at an early stage in the data processing. An example is shown in Figure 10.2, from a test aeromagnetic survey in the Gulf of Mexico aimed at detecting anomalies with a relief of less than two gammas. Both ground and airborne magnetometer (traces 1 and 3) showed "micropulsations" of about 0.5γ relief, which were removed by subtracting ground from air measurements (trace 4).

Even if subtraction is applied, and more importantly if it is not, the survey data must be leveled together before it can be contoured. "Tie lines" are run for this purpose, usually perpendicular to the lines. At each intersection of a line and tie line, the difference between the two magnetic measurements is calculated. These differences are analyzed, and adjustments are applied to reduce them to zero, at the

same time keeping the pattern of adjustments smooth along each line. In effect, the adjustments represent an attempt to reconstruct the time fluctuations. Study of the adjustments can show any lines of data which might be badly affected by erratic time fluctuations.

The distance between tie lines is partly set with regard to errors that could arise from unrecognized time fluctuations (Reford and Summer, 1964). However, the tie lines improve the definition of magnetic gradients which are nearly parallel to the lines, and this consideration also affects the choice of tie line spacing.

10.4 Discussion

The problems of time fluctuations are more severe for marine than for airborne surveys for several reasons. Because of the differences in speed, the same fluctuation will be spread over a greater distance along the airborne measurements and will affect the true shape of magnetic anomalies proportionately less. Also, as a result of speed, the time elapsed between intersections is much smaller for airborne than for marine profiles, so that fluctuations will be better corrected in leveling the data together. Finally, as a practical consideration, marine magnetometers are most often used as auxiliary instruments, not the main survey tool, so that there may be little chance of repeating survey lines affected by magnetic disturbances.

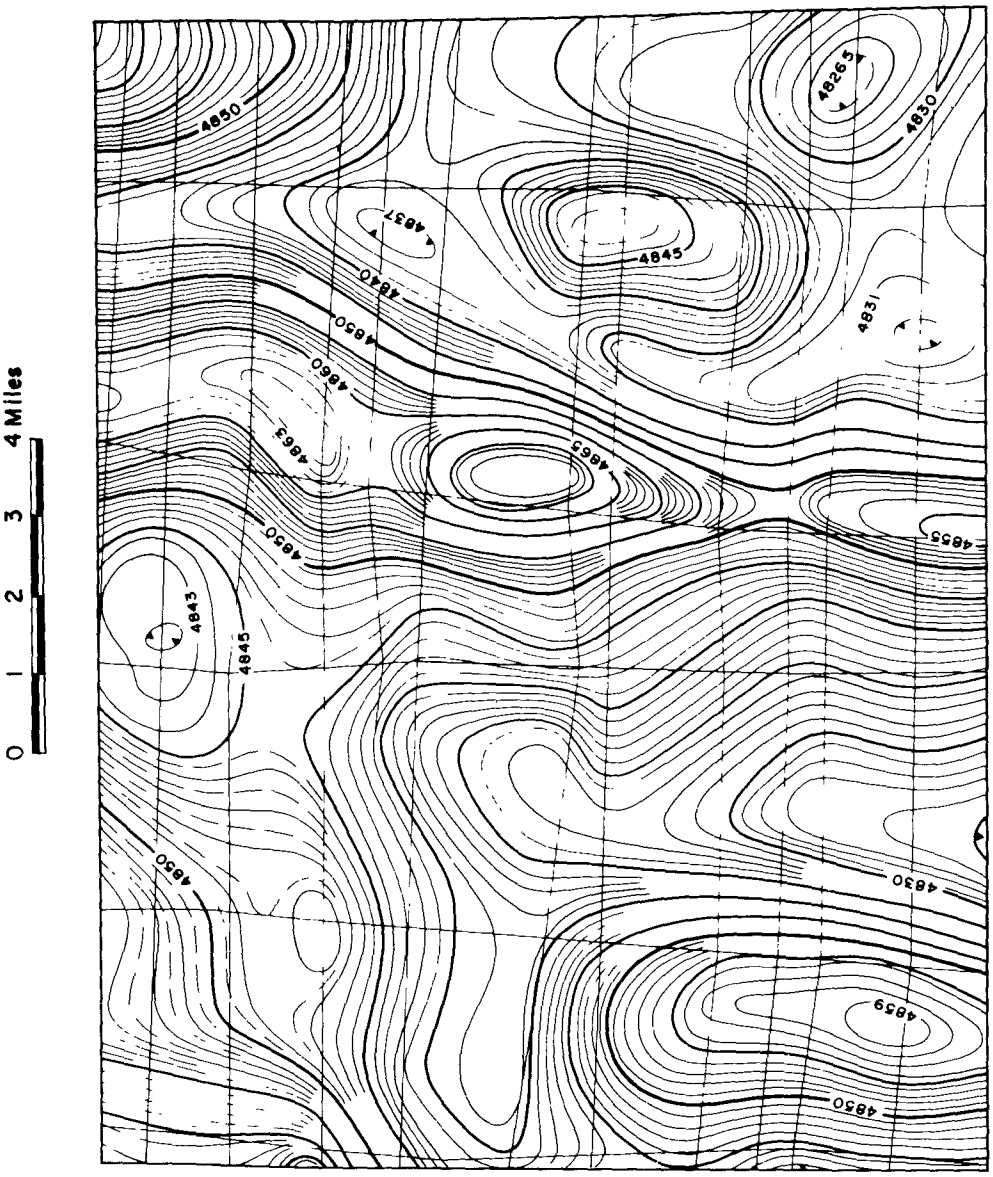


FIGURE 10.1 Total horizontal magnetic intensity contour map from a high resolution airborne survey in the Canadian Arctic.

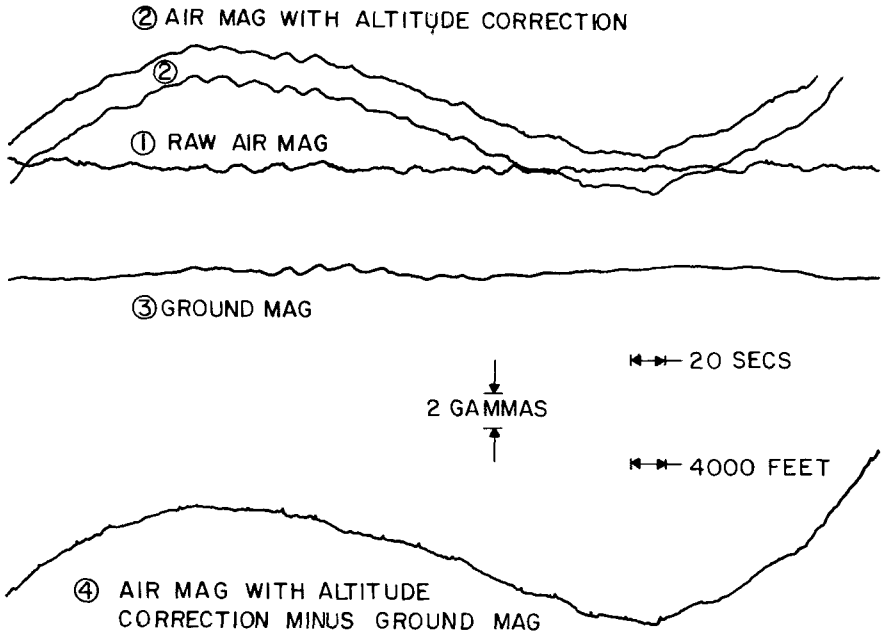


FIGURE 10.2 Aeromagnetic survey over the Gulf of Mexico and ground "base" station data taken to help eliminate the magnetic fluctuations.

IMPACTS OF SOLAR SYSTEM ENVIRONMENT ON MAN AND MAN ON THE ENVIRONMENT

by

P.A. Sturrock

The first seven chapters of this report review various plasma processes which are known to occur in the magnetosphere and the solar system. Some of the known technological effects of these plasma processes were discussed in the previous section. Any one who had been out of touch with these topics for more than a few years would no doubt be surprised at some of these effects. It was a distinct surprise when the association between magnetic storms and power outages was first realized. Moreover, it is unlikely that engineers planning cable communications or petroleum exploration realized, from the outset, that the functioning of their equipment might be affected by fluctuations of the geomagnetic field.

This realization inevitably poses the following question: Are we now overlooking comparable effects? In other words, are there further important effects of the solar system environment on human life and technology or, conversely, does our terrestrial technology have significant effects beyond the known effects in the lower atmosphere — for instance, effects upon the ionosphere, plasmasphere or magnetosphere? Such questions are among the most intriguing and challenging issues with which scientists may be faced.

As with any other branch of science, one may initiate either an experimental (or observational) attack or a theoretical attack. In the first case, one looks for puzzling facts (or quasi-facts) which may

possibly be related to the search; in the latter case, one begins to speculate on consequences of present-day theoretical knowledge. Neither approach is by itself likely to yield firm new scientific information. To use a categorization of scientific activities proposed by Morrison[1], these approaches may yield only "new intelligence". The later stage, which he terms "consolidation of newly occupied territory", will typically involve mutually supportive observational and theoretical activities.

Perhaps the most striking example of a possible effect of the type we are considering, which originated in analysis of observational data, is a possible association between solar variability and terrestrial weather and/or climate. Wilcox[2] notes that claims for a connection between the variable sun and the earth's weather can be found in well over 1,000 papers published during the past century, many of these papers written by illustrious scientists. Nevertheless, the subject has tended to remain on the fringes of respectable science. The near coincidence, for some 10 cycles, between severe drought in the high prairies and sunspot minima strikes some scientists as remarkable, but remains unimpressive to other scientists. Rightly or wrongly, scientists tend to ask "How can there possibly be a real cause-and-effect relationship?"

An important development in the history of this controversy has been the work of Wilcox and his collaborators[3] showing that there is an apparent association between the vorticity area index in the Northern Hemisphere and sector boundaries in the solar wind. Since this correlation occurs on the time scale of days, rather than decades, one

can accumulate within a few years far more evidence on this association than one could on an association between cycles with an 11-year or 22-year period. Another important aspect of these findings is that one is associating a large-scale terrestrial index with an index of the near-earth environment. Such an association is valuable in any search for a causal chain of mechanisms relating solar variability to terrestrial weather.

The attitude of funding agencies towards the more speculative areas of scientific research is crucial. If financial support is available, this in itself tends to make the topic respectable; in addition, it clearly facilitates further research which will lead eventually to a validation or refutation of these speculations. On the other hand, it is unlikely that a Federal agency funding scientific research will give an implicit stamp of approval to a subject which is considered by most scientists to be of dubious respectability. One way to break this mutually restrictive cycle is for an agency to appoint a committee charged with the responsibility of assessing the existing evidence. For instance, such a committee recently met at the request of NASA to study the relationship between solar activity and terrestrial weather[4]. This committee concluded that there is a prima facie case for influence of the sun on terrestrial weather on a time scale of a few days and made recommendations concerning research likely to lead to a more definitive evaluation. On the one hand, further statistical analysis of solar and terrestrial variables is desirable to evaluate the proposed 11-year and 22-year associations and to search for possible causal chains. On the other hand, theoretical

study and computer modeling are desirable to evaluate possible mechanisms which may be involved in a chain of mechanisms coupling variations of the sun's atmosphere to variations of the earth's atmosphere.

Although study of the association between solar variability and terrestrial weather is in danger of becoming respectable, some other proposed correlations do not yet run this risk. For instance, a number of Russian scientists have claimed that there is a real association between geomagnetic storms and the incidence of various human diseases[5,6,7]. In particular, Gnevyshev and Novikova[6] claim that there is a correlation between solar activity and both myocardial infarction and stroke. Since such an association, if real, might point to hitherto unrecognized hazards to human health, my colleaques and I[8] have investigated this hypothesis in terms of U.S. data. We found no evidence for a correlation of the type proposed by Gnevyshev and Novikova. If further analyses show that the Russian claims are, nevertheless, well founded, it will be necessary to determine whether the effect is sensitive to geographical location or some other parameter which might distinguish the Soviet data from the U.S. data.

Over forty years ago, Dull and Dull[9] carried out an analysis claiming to demonstrate a relationship between the incidence of nervous and mental diseases and suicides, on the one hand, and magnetic storms on the other hand. Friedman, Becker and Bachman[10] decided to test the possibility of psychological disturbances associated with magnetic storms. They therefore searched for a correlation between daily

admissions to several psychiatric hospitals in New York State and magnetic activity in that area. The authors claim that their results show the existence of a statistically significant relationship between geomagnetic parameters and a gross measure of human disturbance. In a later article[11], the same scientists searched instead for a correlation between psychiatric hospital admissions and cosmic ray activity. Their claim is that cosmic ray indexes provide more significant parameters of geophysical events related to human behavior measures than do the more typical measures of geomagnetic activity, such as the K-index or Ap. These are provocative findings but, to the best of my knowledge, these important questions have not been pursued.

Most scientific research is concerned with projects of fairly high probability of success and fairly low payoff. In deciding whether or not to study topics such as those so far mentioned in this chapter, a scientist must decide whether to invest any time and effort into a topic which may appear to have a low probability of success (in terms of proving beyond doubt the existence of a significant phenomenon), when success would imply a high "payoff" in the sense that the unequivocal demonstration of such an effect would be highly significant. The true importance may not be in the established association (there is little we can do to avoid magnetic storms!), but it would surely lead to new knowledge if the problem is pursued to the point that we have a complete understanding of the association. We may find, for instance, that certain electric or magnetic fluctuations have adverse effects (or possibly beneficial effects) on human beings. On the other hand, the

association may be due to some other unsuspected environmental changes which are associated with but different from electric and magnetic fluctuations.

Scientists involved in the study of electromagnetic aspects of the earth's environment may also be alert to the inverse class of effects; possible effects on the ionosphere, plasmasphere and magnetosphere due to man's activities. We saw, in the preceding chapter, that magnetic storms have a significant influence on power systems. There is some evidence for the inverse process, disturbances of the magnetosphere by radiation from power lines. A recent search for such an effect has been carried out by Dr. A. Fraser-Smith[12] who has analyzed many years of geomagnetic data to see if the geomagnetic Ap index displays a weekly variation. His positive conclusion is that this index shows enhancement of approximately 7% at weekends, which one may reasonably associate with the fact that power consumption is lower (by 30%) on weekends.

It has been known for some time that VLF chorus emissions are influenced by power line radiation[13]. Since VLF electromagnetic waves play an important role in the precipitation of particles from the magnetosphere into the ionosphere[14,15], it is perhaps not too far-fetched to speculate that power line radiation may lead to precipitation which in turn produces ionospheric effects. Variations of the ionosphere can, of course, have important effects on radio communication systems.

The effect of power line radiation on VLF properties of the magnetosphere seem to be quite subtle. It has recently been learned that the amplification of one VLF signal in a magnetospheric duct may be drastically

suppressed by the presence of another VLF signal. Not only may power line radiations directly influence the magnetosphere, but it may change the response of the magnetosphere to electromagnetic disturbances of natural origin.

We have already witnessed large-scale changes in the earth's radiation belts due to human activity: namely, the "Starfish" experiment of exploding nuclear bombs above the ionosphere. Not only was the intensity of radiation in the inner belt substantially enhanced, but we discovered that this enhanced radiation took several years to decay. Substantially enhanced radiation may be a danger to astronauts. In addition, the enhanced rate of dumping of energetic particles into the ionosphere, which may follow such a perturbation of the radiation belts, is likely to cause changes in the D and E layers, and possibly magnetic pulsations which may interfere with power control systems and telephone communication systems.

Such questions deserve careful study before related experiments are undertaken, such as the injection of high-energy particles into the magnetosphere by an AMPS type experiment, or the injection of large amounts of electromagnetic energy which may either precipitate large fluxes of magnetospheric particles into the ionosphere or accelerate ionospheric particles, some of which may migrate into the magnetosphere. Dr. Francis Perkins of Princeton University has recently made the specific suggestion that if 3 MW of radio power at 700 kHz were beamed upwards into the ionosphere, a large fraction would be absorbed at a height of about 1400 km, principally by accelerating electrons which would then be injected into the magnetosphere.

It has recently been learned that explosive chemical combustion may have drastic effects on the ionosphere. In particular, an iono-

spheric "hole" was produced as a result of the rocket burn which put the Skylab spacecraft into orbit. Will similar, or possibly more drastic, perturbations be caused by rocket burns required to put the Space Shuttle into orbit or bring it down to earth?

Intriguing scientific questions arise not only from the rational pursuit of existing knowledge, but also from the study of possible associations of seemingly unrelated variables, and from studying the implications of man's increasing capability to perturb his environment.

REFERENCES

1. Morrison, P., 1975, Opportunities and Choices in Space Sciences, 1974 (Space Science Board, National Academy of Sciences), p. 21.

2. Wilcox, J.M., Solar structure and terrestrial weather. Science 192, 745-748, 1976.

3. Wilcox, J.M., P.H. Scherrer, L. Svalgaard, W.O. Roberts, R.H. Olson, Solar magnetic sector structure: relation to circulation of the earth's atmosphere. Science 180, 185-186, 1973.

4. Sturrock, P.A., G.E. Brueckner, R.E. Dickinson, N. Fukuta, L.J. Lanzerotti, R.S. Lindzen, C.G. Park, and J.M. Wilcox, "Study of the Relationship Between Solar Activity and Terrestrial Weather" (SUIPR Report No. 671, Stanford University), 1976.

5. Chizheviskii, A.L., One aspect of the specific bioactive or Z-radiation from the sun, in The Earth in the Universe, ed. by V.V. Fedynskii (NASA TT F-345 TT 66-51025), pp. 280-307, 1968.

6. Gnevyshev, M.N. and K.F. Novikova, The influence of solar activity on the earth's biosphere. J. Interdiscipl. Cycle Res. 3, 99-104, 1972.

7. Spreiter, J.R., Shock waves in the solar system. Astronautica Acta 17, 321-338, 1972.

8. Lipa, B.J., P.A. Sturrock, and E. Rogot, Search for correlation between geomagnetic disturbances and mortality. Nature 259, 302-304, 1976.

9. Dull, T. and B. Dull, Zusammenhange zwischen Storungen des Erdmagnestismus und Haufungen von Todesfallen, Deutsch. med Wschr., 61, 95, 1935.

10. Friedman, H., R.O. Becker, and C.H. Bachman, Geomagnetic parameters and psychiatric hospital admission. Nature 200, 626-628, 1963.

11. Friedman, H., R.O. Becker, and C.H. Bachman, Psychiatric ward behavior and geophysical parameters. Nature 205, 1050-1052, 1965.

12. Fraser-Smith, A., A weekend increase of geomagnetic activity. EOS Trans. A.G.U. 58, 470 (Abstract #SM26), 1977.

13. 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, 4249-4258, 1975.

14. Rosenberg, T.J., R.A. Helliwell, and J.P. Katsufrakis, Electron precipitation associated with discrete very-low-frequency emissions. J. Geophys. Res. 76, 8445-8452, 1971.

15. Helliwell, R.A., J.P. Katsufrakis, and M.L. Trimpi, Whistler-induced amplitude perturbation in VLF propagation. J. Geophys. Res. 78, 4679-4688, 1973.