Solar-Planetary Relationships: Magnetospheric Physics

Aaron Barnes

Reprinted from
REVIEW OF GEOPHYSICS AND SPACE PHYSICS
Volume 17  Number 4  June 1979
Solar-Planetary Relationships: Magnetospheric Physics

PHYSICS OF THE SOLAR WIND

For the 1975-1978 IUUG Quadrennial Report

Aaron Barnes

Theoretical and Planetary Studies Branch, Ames Research Center, NASA, Moffett Field, California 94035

Introduction

The quadrennium 1975-1978 was a period of great advance for solar-wind studies, a period that combined exploration of new regions with increased maturity in established fields of study. The Helios, Pioneer, and Voyager spacecraft have been exploring the inner and outer regions of the solar wind. There has been a rebirth of the study of possible relations between solar variability and Earth's climate and weather, stimulated largely by Eddy's [1976] investigation of the Maunder Minimum; the solar wind may well prove to be a significant link in solar-terrestrial relations. Unique coronal data from the SKYLAB 1973-1974 mission, in combination with satellite and ground-based observations, provided the basis for identification of coronal holes as the main source of high-speed solar wind. The interplanetary medium has continued to serve as a laboratory for the study of plasma processes that cannot yet be studied in terrestrial laboratories, providing insights of potential importance both for controlled fusion research and for astrophysics. It is ironic that such a productive period, the legacy of many past space missions, was also a time of severely limited opportunity for new space investigations; the outlook for the future is equally austere. Especially regrettable is the dearth of career opportunities for young scientists in this field; comparison of the bibliography of this report with that of its predecessor 4 years ago shows few new names. Despite such problems, research has continued with enthusiasm and much has been learned.

The present report will survey selected topics related to the origin, expansion, and acceleration of the solar wind and the plasma physics of the interplanetary medium. Companion reports [Papadopoulos, 1979; Scherrer, 1979; Smith, 1979] deal with a number of closely related topics, including the heliocentric distance and latitude variation of the solar wind and its fluctuations, topology of the interplanetary magnetic field, morphology of solar-wind streams and shocks, sunweather studies, and interplanetary manifestations of type-III bursts. Of the subjects that fall within the scope of this report, the study of the relationship between coronal holes and solar-wind streams, and the associated revision of our ideas about solar wind acceleration and heating, have had the most impact; hence I review these topics in considerable detail. In addition, I discuss the topics of hydromagnetic waves and turbulence, and interplanetary electrons, as items of particular importance during the past quadrennium. Limitations of time and space require the omission of a number of important topics from the text (the alternative is to try to cover everything, and thus produce a completely superficial report); however, the omitted topics are thoroughly covered in the bibliography.

Besides the archived periodical literature, important papers are to be found in the proceedings of two major international conferences, The International Symposium on Solar-Terrestrial Physics [Williams, 1976], and Solar Wind 4 [Rosenbauer, 1978]; the individual papers are not listed in the bibliography unless specifically cited in the text. The proceedings of the SKYLAB Coronal Hole Workshop [Zirker, 1977a] is also a valuable source. For more detailed reviews of the topics discussed in this report, the reader may consult a number of review articles [Dobrowolny and Moreno, 1976, 1977; Hollweg, 1975a, 1978a; Holzer, 1977b, 1978; Zirker, 1977b; Barnes, 1978]. For related reports for the previous quadrennium, see Gosling [1975], Hirshberg [1975], Thomas [1975], and Barnes [1975].

Coronal Holes and the Solar Wind

The concept that coronal holes are the primary sources of fast solar wind streams grew and attained wide acceptance during 1975-1978. The rapid growth of this notion is due in large part to the research associated with SKYLAB Solar Workshop I, which took place in 1975-1976 [Zirker, 1977a,b]. The workshop activities were centered on a unique set of data from the Apollo Telescope Mount (ATM), a battery of advanced solar telescopes aboard SKYLAB, during the 9-month mission (May 1973-February 1974). About 80 scientists with diverse backgrounds in optical, radio and theoretical solar astronomy, interplanetary physics, and cosmic ray physics participated in the workshop research activities. This research resulted in a number of seminal papers that subsequently have had a broad influence on the entire field of solar wind studies.

Coronal holes are regions of abnormally low density in the corona, which show up most clearly in observations made above the Earth's atmosphere. They lie within open magnetic configurations whose footpoints are contained in large regions of the solar surface with one dominant magnetic polarity. The coronal holes of the SKYLAB period have been intensively investigated; their properties are thoroughly summarized in the review articles of Bohlin and Hulburt [1977], Krieger [1977], Levine [1977], Withbroe [1977], and Zirker [1977b].

Observation of the solar wind near the ecliptic plane over one solar cycle indicates that stable, large-amplitude fast streams (peak velocity \( \geq 700 \) km/sec) are more common in years of declining and minimum solar activity than near solar maximum [Bame et al., 1976; Gosling et al., 1976b]; the broadest streams occurred near solar minimum in
of fast solar wind is illustrated schematically in Figure 1.

The tenuous regions at the solar poles during solar minimum are presumably coronal holes that last throughout the solar cycle, perhaps disappearing at solar maximum. The polar holes would then be expected to produce fast solar wind, and the solar wind should be faster at high latitude than in the ecliptic plane over much of the solar cycle. Interplanetary radio scintillation observations from 1971 to 1975 showed the average solar wind speed to increase with solar latitude, with a mean gradient of 2.1 km/sec per degree of latitude [Giles and Rickett, 1976]. However, in situ measurements [e.g., Bame et al., 1977a; Rhodes and Smith, 1975, 1976a,b] give both larger and smaller gradients at particular epochs, and studies of comet tail observations over a 75-year interval show no systematic variation of speed with latitude [Brandt et al., 1975]. These questions have been reviewed by Dobrowolny and Moreno [1976] [see also Nerney and Suess, 1975; Pneuman, 1976]. The variation in the observational results may be due largely to temporal variations in the corona and solar wind. An additional ambiguity arises from the fact that most inferred gradients are based on averages over solar longitude [Hundhausen, 1978]. A clear resolution of these issues will probably not be attained until the International Solar Polar Mission of the mid-1980's. The general question of interplanetary gradients is discussed in greater detail in a companion report [Smith, 1979].

Since coronal holes emanate from <20% of the solar surface, the SKYLAB Workshop picture implies that the solar wind comes from a relatively small fraction of the solar surface. Simple conservation arguments then lead to surprisingly large values of magnetic field (~8 Gauss) at the coronal base [Hundhausen, 1977]. Potential-field calculations of the coronal magnetic field based on maps of the photospheric field, when correlated with interplanetary magnetic measurements, confirm that only a relatively small fraction of the photospheric area connects via open field lines to the interplanetary field [Levine et al., 1977a,b]. Levine et al. found that those regions which do connect lie beneath coronal holes (sometimes at high latitude!), that the areas of open flux tubes expand much more (X40-120) than would radial tubes, and that the fastest solar wind streams seem to come from those tubes that expand least in cross-sectional area [see also Nolte et al., 1976a,b]. Moreover, the best-fit models indicate that only about 10% of the solar surface is open to the solar wind, in which cases even coronal holes contain some regions of closed field. Burlaga et al. [1978a] extrapolated the observed values of the interplanetary field inwards, and used potential-field models to infer the photospheric field in the source region, which could then be compared with the measured photospheric field. The inferred photospheric field (average 9 Gauss) was consistent with the measured field in range and average, but not in all detail. In some cases the interplanetary field seems to correlate with open-field regions not obviously related to coronal holes [Burlaga et al., 1978a; Levine, 1978].

If the solar wind comes from a fairly small fraction of the surface, conservation arguments also lead to high values of mass and energy flux...
in the low corona [Hundhausen, 1977]. The high mass flux implies acceleration of the wind lower in the corona than had previously been expected, but does not greatly change the expected rate of total particle output. The inferred energy flux \( (\sim 5 \times 10^5 \text{ erg sec}^{-1}) \) implies a total power input to the solar wind several times larger than previously estimated, and the flux is too high to be consistent with thermal conduction as the sole energy transport mechanism (at least for coronal temperatures currently regarded as acceptable). Moreover, for such high fluxes the solar wind must be the dominant energy loss mechanism for coronal holes [Hundhausen, 1977; Kopp and Orrall, 1977].

An abundance of evidence points to coronal holes as the source of fast solar wind; this is the viewpoint that will be taken in our subsequent discussion of solar-wind energetics. However, it should be borne in mind that most of the evidence on this question is data from about one year of one solar cycle. Even during that period the correspondence between fast streams and coronal holes is not one-to-one [Burlaga et al., 1978a; Rickett et al., 1976]. The apparent exceptions may reflect the disappearance of coronal holes as optical (but not as magnetic) features as they evolve [Levine, 1978], limitations of analysis techniques and/or available data, or simply the fact that regions other than coronal holes can produce high-speed wind. Moreover, the fast streams of 1973-1975 may not be typical of all solar minima [Gosling et al., 1977b]. The idea that coronal holes are the main source of high-speed streams, and that such streams regularly appear in the ecliptic plane near solar minimum, must be tested over subsequent solar cycles.

Accretion and Heating of the Solar Wind

Most recent theoretical work on the acceleration and heating of the solar wind has centered on the working hypothesis of coronal holes as the source of the wind. As reported above, elementary arguments strongly suggest that thermal conduction cannot be the main heat transport mechanism in flows out of coronal holes. This point was made strikingly clear in the empirical model of a polar coronal hole due to Munro and Jackson [1977]. They used observations from the SKYLAB white light coronagraph to determine the three-dimensional density structure within a polar coronal hole from 2 to 5 R\(_S\). They found the increase of the hole's cross-sectional area from the surface to 3 R\(_S\) to be 7 times greater than that of a radial cone. Under the assumption that the solar wind coming out of this hole was similar to that of high-speed streams in the ecliptic, they inferred radial profiles of flow speed and effective pressure. The velocity profile was much steeper than in radial-flow models, the sonic point lying between 2.2 and 3 R\(_S\). One important consequence of the rapid expansion is that the flow becomes completely collisionless within \( \sim 4 \text{ R}_S \).

The effective pressure profile can be converted to a profile of "effective temperature," which increases out to \( \sim 4 \text{ R}_S \) or beyond, implying extended energy deposition [either heating or work due (for example) to hydromagnetic wave pressure]. In particular, thermal conduction could not produce an increasing effective temperature profile. These conclusions are generally consistent with the empirical model of Rosner and Vaiana [1977], which was based on X-ray, EUV, and radio observations rather than white light data. In particular, they find the sonic point to lie above the temperature maximum, and conclude that an additional energy supply (besides thermal conduction) is required to drive streams from coronal holes.

Much work has centered on developing theoretical models of the flow in such a diverging field geometry [see the review by Suess, 1978]. Kopp and Holzer [1976] calculated models of one-dimensional expansion in prescribed flux tubes whose cross sections increase faster than \( r^2 \). They showed that for fixed wind speed (thus fixed energy per particle), two or more critical points can arise in the dynamical equations if the flux tube expands rapidly enough. The flow must pass through the first critical point, suggesting that the sonic point may lie low in the corona for expansion from coronal holes. However, the polytropic approximation implies different energy deposition profiles for the different models, so that the calculations of Kopp and Holzer do not clearly separate the effects of flow divergence from heating. Their work was extended by Steinolfson and Tandberg-Hanssen [1977], who replaced the polytropic equation of state with the energy equation, with thermal conduction as the only heat transport process. Their numerical solutions of the dynamical equations give results qualitatively similar to those of Kopp and Holzer. They conclude that their results are a poor representation of high-speed interplanetary streams, reinforcing the conclusion that streams from coronal holes are not driven by thermal conduction alone. (Radial-flow models of Nerney and Barnes [1977, 1978] also support this conclusion.) Holzer [1977a] has studied the problem in a very general way, investigating various forms of flux tube divergence, energy deposition, and momentum addition. He showed that in certain flow-tube geometries, the heating by thermal conduction can be enhanced, possibly to the point of eliminating the need for another heating mechanism. However, he points out that this is a qualitative conclusion, and that it is not clear that a realistic conduction-driven wind model is possible. Holzer also emphasized the importance of an accurate description of electron heat transport in wind models. The problem is especially difficult in the context of coronal holes, because the rapid expansion leads to collisionless flow near the Sun. In particular, the electrons must be described kinetically, and it is possible that local descriptions of electron dynamics are inadequate [Scudder and Olbert, 1978]; this problem will be discussed in a later section.

Suess et al. [1977] calculated a series of magnetohydrodynamic models of the Munro-Jackson coronal hole, approximating the flow as quasi-radial [Suess and Nerney, 1973b], and using a polytropic equation of state. The density boundary condition (a latitude profile at fixed radius) was inferred from the Munro-Jackson empirical model, and was not varied. Corresponding boundary conditions on temperature and magnetic field were varied (under certain empirical constraints) and served as "free parameters" defining the series of models. Each calculated model yielded the spatial distribution of density in the hole and the geometry of the hole boundary, which were then compared to the Munro-Jackson model. The best-fitting MHD model corresponded to boundary...
conditions (at 2 R\(_S\)) of temperature decreasing from 2.5 \times 10^6 K at hole center to 1.25 \times 10^6 K at the edge, and magnetic field decreasing from 1 G at the center to 0.5 G at the hole center. The flow speed calculated at hole center increased from 150 km/sec at 2 R\(_S\) to 350 km/sec at 5 R\(_S\). The profile inferred by Munro and Jackson was somewhat steeper (<100 km/sec to 450 km/sec), but this latter profile was based on the area expansion of the entire hole, so that the agreement between the empirical and MHD models is probably adequate. Suss et al. found that velocity and temperature have their maximum near hole center at all distances. The heating implicit in the polytrope model occurs mostly near the center of the hole, with little or none at the edge, and varies in proportion to the field strength at 2 R\(_S\). This model is generally a satisfactory picture of flow in a coronal hole. However, it should be noted that the model may not be unique (the authors point out a difficulty in reconciling the model with photospheric field observations), and that polytropic models are not necessarily reliable for drawing conclusions about energy deposition.

As discussed above, the association of fast streams with coronal holes seems to require extended energy deposition (heating or work) in the wind far beyond the coronal base. It is reasonable, and currently fashionable, to suppose that hydromagnetic waves of solar origin account for the extended acceleration and/or heating, although little firm observational evidence exists to confirm or deny this hypothesis. It has long been recognized that wind models including Alfvén or magnetoacoustic waves have higher flow speeds than analogous models without waves. Jacques [1977a,b; see also Hollweg, 1978a] recently showed that wind speeds of ~700 km/sec can readily be attained in Alfvén wave driven radial flow models. Hollweg [1978a] studied the propagation of small-amplitude Alfvén waves in a realistic model of the solar atmosphere, for a postulated wave source near the top of the convection zone. He found that the wave power output is strongest near a series of resonant periods (<1.6 hr), and that enough power to significantly affect the solar wind may reside in these peaks. Hollweg [1978b] has pointed out that the transverse wave numbers associated with observed photospheric motions would not send magnetoacoustic waves into the corona if transmission from the photosphere were correctly described by the linearized theory of transmission through a thin boundary. However, this model may not be realistic because of geometrical effects and nonlinear modification of the wave spectrum. Wentzel [1977a] has argued (essentially by dimensional analysis) that the Alfvén mode is the hydromagnetic wave mode least affected by nonlinear dissipation in the outer corona; however, magnetoacoustic waves may be generated as well as dissipated in the corona [Hollweg, 1978a; Wentzel, 1978]. Auer and Rosenbauer [1977] suggested that measurements of proton thermal anisotropy and its variation are consistent with fast--wave heating of the outer corona, but pointed out that this interpretation may not be unique [cf. Hollweg, 1978b].

Interplanetary Hydromagnetic Waves and Turbulence

The interplanetary medium provides the best presently available laboratory for the study of large-amplitude hydromagnetic waves and turbulence. This fact, and the interest in hydromagnetic waves as a possible wind acceleration or heating mechanism (see preceding), have motivated a number of theoretical and observational studies of interplanetary hydromagnetic fluctuations [see the review of Barnes, 1978]. It has long been established that much of the solar wind (especially high-speed streams and their trailing edges) exhibits fluctuations that are nearly Alfvénic (constant magnetic field strength and plasma density), propagating outward from the Sun. Five years ago, most theorists envisaged these fluctuations as nearly plane waves whose spatial variation should be describable by hydromagnetism. This viewpoint now appears to be inconsistent with observation; measured directions of minimum magnetic variance do not agree with predictions of the eikonal (geometrical) theory [Burlaga and Turner, 1976; Solodyna and Belcher, 1976], and two-spacecraft constraints on direction of wave normals do not agree with minimum-variance directions [Burlaga and Neurot, 1976]. Although these studies were based on limited data sets, and cannot be regarded as definitive, they do strongly suggest that the eikonal description is inadequate.

If the eikonal approximation fails, constant-phase surfaces must be curved on a scale comparable to the fluctuation "wavelength." This curvature and the characteristic power law spectra of interplanetary fluctuations suggest that it may be more useful to think of interplanetary Alfvénic fluctuations as turbulence rather than waves. The two viewpoints are of course equivalent in the small-amplitude limit. However, in large-amplitude turbulence, part of the fluctuation spectrum may well be associated with nonlinear phenomena that are not easily incorporated in a wave picture. Thus, attempts to analyze interplanetary turbulence in terms of a superposition of, for example, Alfvén and magnetoacoustic waves are not strictly self-consistent. On the other hand, there is at present no adequate theory of nonlinear turbulence, and mode superposition has generally been adopted as a working hypothesis. Sarf and Valley [1976] and Neugebauer et al. [1978] have carried out data studies under the mode-superposition hypothesis, and find evidence for a small magnetoacoustic component in some data periods. The small but nonzero fluctuations of magnetic field strength found by Burlaga and Turner [1976] even during the purest Alfvénic periods are also consistent with the presence of a magnetoacoustic component. These investigations were based on data sets too small for general conclusions about the character of compressive components of turbulence, and much more work is needed. The large amplitude character of interplanetary waves presents formidable theoretical problems. In this area the greatest progress in recent years is probably the numerical simulation of test particles in large-amplitude magnetoacoustic waves [Matsumoto, 1977]. Matsumoto found that as wave amplitude increases, particle trapping becomes important and apparently can inhibit the Landau damping process. The processes of trapping and Landau damping are expected to compete with nonlinear steepening of magnetoacoustic waves, so that magnetoacoustic shock generation in the solar wind may be quite different from that envisaged by magnetohydrodynamics [Barnes and Chao, 1977].
These studies represent first steps toward understanding the evolution of simple waves in collisionless plasma. Future progress is likely to require simulation studies using large computers.

Theoretical studies of collisionless turbulence (as opposed to simple waves) so far have been limited to the case of weak turbulence. Goodrich [1978] and Hollweg [1978a] have used the quasi-linear approach to investigate the development of velocity distributions in the presence of turbulence, and to analyze the force exerted on the plasma by the turbulence field. Jacques [1977a] used a somewhat different (Lagrangian) technique to study this kind of force, but so far only in the MHD approximation.

Solar-Wind Electrons

The past quadrennium has brought a number of discoveries and insights about the electron component of the solar wind. It is now clear that the electron velocity distribution is not a straightforward consequence of either collisional-dominated or exospheric flow, but is rather in a subtle intermediate state. The electron velocity distribution can be separated into a low-energy (≤60 eV) "core" and high-energy "halo" [Feldman et al., 1975]. The nearly isotropic Maxwellian core distribution is strongly influenced by Coulomb collisions [Feldman et al., 1975, 1976b]. The halo component, on the other hand, varies from bi-Maxwellian [Feldman et al., 1975] in low-speed solar wind, to a strongly beamed (or "strahl") state [Rosenbauer et al., 1976, 1977; Feldman et al., 1978b] in high-speed streams, and is not dominated by Coulomb collisions (at least not in a local sense). Heat conduction in the wind is primarily due to the halo, either by convection of the halo relative to the core [Feldman et al., 1975] or by the beaming of the "strahl" [Rosenbauer et al., 1976].

The simplest model of a core-halo distribution is that of a "fluid" core and "exospheric" halo; in this picture the halo contains a record of conditions at the "exobase." Feldman et al. [1978b] infer an exobase lying between 10 and 30 R_S from their high-speed stream data. Ogilvie and Scudder [1978] studied the run of core and halo temperatures between 0.45 and 0.05 AU and showed that extrapolation inward gives equality of the two temperatures at radial distance ~2-15 R_S. These results generally support the mixed fluid-exosphere model. However, the halo population can be relatively isotropic, especially in low-speed streams, strongly suggesting that its evolution is not purely exospheric. It is possible that the velocity distributions become unstable; the resulting instability could isotropize the halo velocity distribution and regulate heat flux [Evlatar and Schulz, 1976; Feldman et al., 1976b; Gary, 1978a,b; Gary et al., 1975a,b; Lakhina, 1977; Schwartz, 1978; Singer, 1977; Singer and Roxburgh, 1977]. The concept of heat flux regulation by microinstability, and a related method of closing the plasma moment equations, has been discussed in detail by Hollweg [1976, 1978a].

Scudder and Olbert [1978] have examined the electron distributions from a somewhat different viewpoint. They studied the global form of the electron velocity distribution predicted by the Krook kinetic equation, assuming only Coulomb collisions. This model recovers the core-halo form, the core being dominated by collisions as in previous models. The halo population is more nearly collisionless, but is nevertheless governed to a large degree by rare Coulomb collisions. In particular, halo electrons with sunward velocities have been scattered backward by Coulomb collisions occurring beyond the observer (1-10 AU). Thus, the halo population has a memory not only of conditions near the Sun, but also of regions beyond 1 AU. The predicted velocity distributions are reasonable representations of observed ones. The work of Scudder and Olbert shows that the halo particles can be understood without invoking wave-particle interactions, and indicates that a purely local description of electron dynamics may not be adequate.

Acknowledgments

I wish to thank J. Mihalov and S. Nerney for their comments on the first draft of this report.

References


Barouch, E., Properties of the solar wind at 0.3 AU from measurements at 1 AU, J. Geophys. Res., 82, 1493, 1977.


Behannon, K. W., Observations of the interplanetary magnetic field between 0.46 and 1 AU by the Mariner 10 spacecraft, Ph.D. thesis, Catholic Univ. of America, Washington, D.C., 1976. (Also NASA/GSFC X Doc. 692-76-2, Jan. 1976.)


Bell, B., and G. Noci, Intensity of the FeXV emission line corona, the level of geomagnetic activity, and the velocity of the solar wind, J. Geophys. Res., 81, 4508, 1976.


Brandt, J. C., R. S. Harrington, and R. G. Roosen, Interplanetary gas. XX. Does the radial solar
Burlaga, L. F., and J. D. Scudder, Motion of shocks through interplanetary streams, J. Geophys. Res., 80, 4004, 1975.


Gary, S. P., W. C. Feldman, D. W. Forslund,


Intrillgator, D. S., In situ observations of the scale-size of plasma turbulence in the asteroid belt (1.6-3 astronomical units), Astrophys. J., 196, L87, 1975a.


Ivanov, K. G., A type of hydromagnetic waves of finite amplitude in magnetohydrodynamics with anisotropic pressure, Geomagnetism and Aeronomy, 15, 664, 1975.


Nolte, J. T., A. S. Krieger, E. C. Roelof, and


Wu, C. S., and J. D. Huba, Low frequency fluctu-


