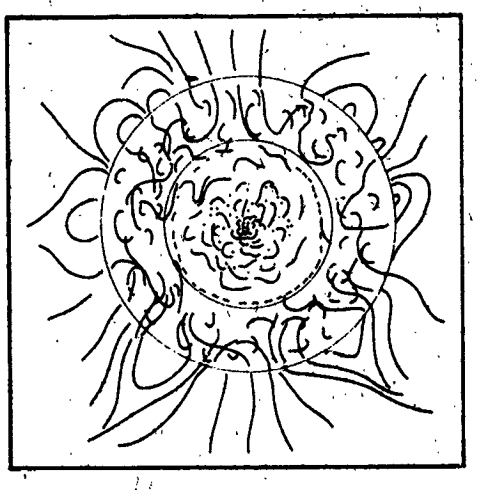
JPL PUBLICATION 78-70

A Close-Up of the Sun



September 1, 1978

National Aeronautics and Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

JPL PUBLICATION 78-70

A Close-Up of the Sun

Edited by
M. Neugebauer
R. W. Davies

September 1, 1978

National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

ABSTRACT

NASA's long-range plan for the study of solar-terrestrial relations includes a Solar Probe Mission in which a spacecraft is put into an eccentric orbit with perihelion near 4 solar radii (0.02 AU). The scientific experiments which might be done with such a mission are discussed in this report. Topics included are the distribution of mass within the Sun, solar angular momentum, the fine structure of the solar surface and corona, the acceleration of the solar wind and energetic particles, and the evolution of interplanetary dust. The mission could also contribute to high-accuracy tests of general relativity and the search for cosmic gravitational radiation.

CONTENTS

SECTION I.

FOREWORD

FOREWORD

M. Neugebauer

Jet Propulsion Laboratory
Pasadena, California 91103

This book is, to first order, the Proceedings of the Solar Probe Science Workshop held at the California Institute of Technology on May 22 and 23, 1978. This volume is not, however, a transcript of the Workshop. Some presented papers were not written up or were not accepted for publication. Some papers in this book were written by people who were unable to attend the meeting. Finally, formal and informal discussions at the Workshop inspired some people to write down their new ideas and submit them for inclusion in the Proceedings. This, I feel, is one of the marks of a successful symposium.

The topic of the Workshop was the scientific uses of a Solar Probe. A Solar Probe Mission is part of the long-range plan of NASA's Solar Terrestrial Program Office. The trajectory of the planned Solar Probe is a very eccentric ellipse with perihelion near 4 solar radii (0.02 AU) and aphelion at or beyond 1 AU. A plot of the near-solar part of the trajectory is shown in Figure 1. During the solar flyby, the spacecraft bus and most of the scientific instruments will be totally shaded by a heat shield; the perihelion distance of 4 solar radii was selected on the basis of a preliminary heat-shield performance calculation. Possible trajectories and a tentative spacecraft design are discussed in more detail in Section IX in papers by Randolph (JPL's Solar Probe Study Leader), Bender, Yen, and Roxburgh.

At the Workshop, Harold Glaser, the head of NASA's Solar Terrestrial Program Office, asked, "What can the Solar Probe do that no other mission can do?" The answer to this question is, hopefully, made clear by this report. I would summarize the very broad, multi-faceted answer as follows:

(1) Only with a Solar Probe is it possible to obtain data on the distribution of mass and angular momentum in the solar interior. Recent solar neutrino experiments indicate that either our understanding of the

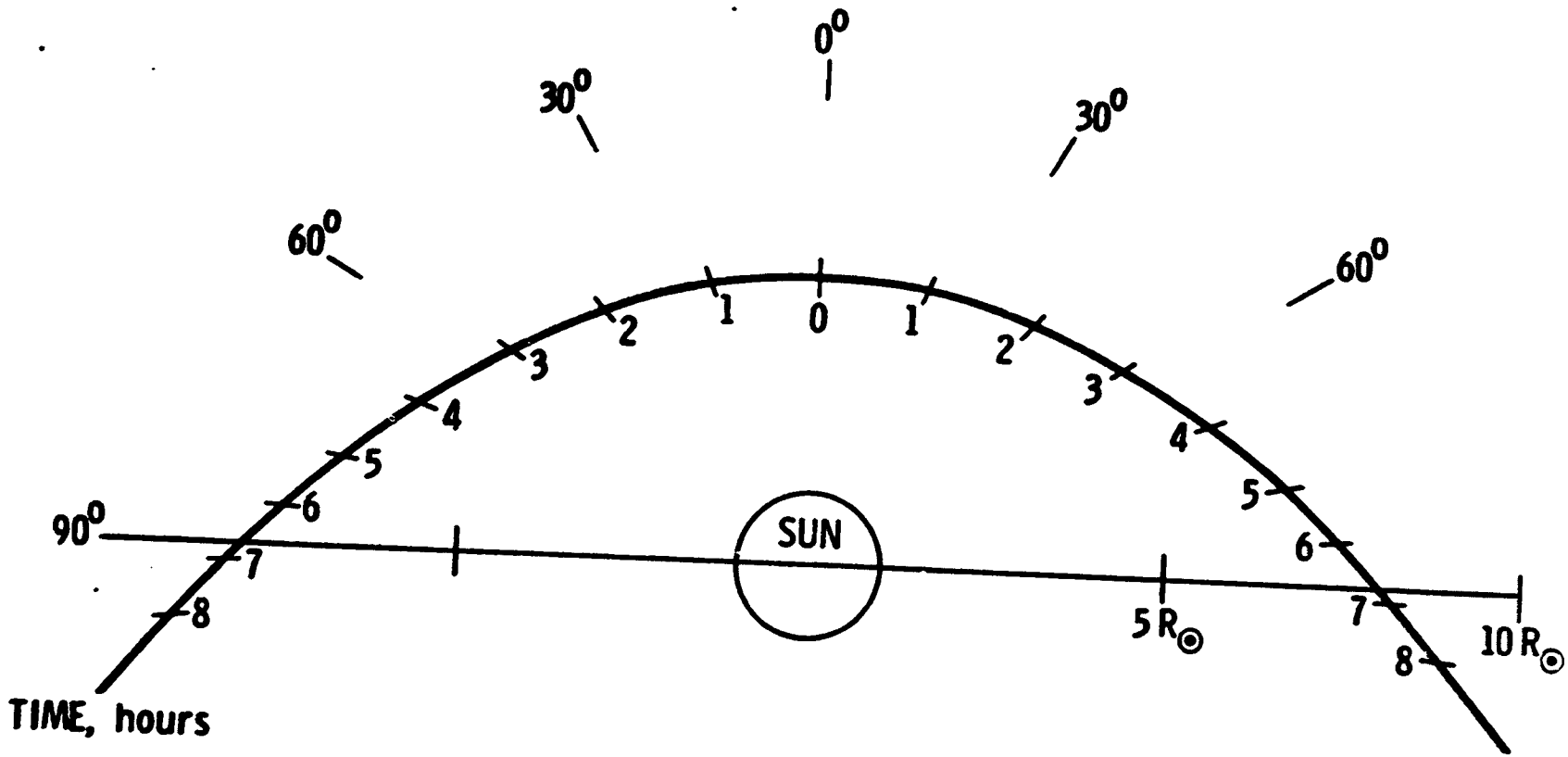


Figure 1. The near-Sun portion of the Solar Probe trajectory.

solar interior (and thus of stellar evolution) or our understanding of neutrino physics must be incorrect. A properly instrumented Solar Probe could measure the solar gravitational quadrupole moment J_2 (Papers by Roxburgh; Reasenberg and Shapiro; Anderson and Lau; Nordtvedt; Everitt and DeBra; Forward.), oscillations in J_2 (Paper by Douglass.), and perhaps even the solar angular momentum vector (Paper by Nordtvedt.). High-accuracy tests of general relativity would be an unavoidable output of these experiments.

(2) With a small telescope, a Solar Probe could obtain higher-resolution pictures (of a very limited portion of the solar surface) than can be obtained with the largest solar telescopes planned for Earth orbit. Furthermore, the Solar Probe sweeps around the Sun so rapidly that stereoscopic views of many features can be obtained. The ultimate limit of the sizes of solar structures is unknown. Known problems which can be attacked with the resolution and stereoscopic viewing allowed by the Solar Probe include: What is the temperature and density structure of a coronal loop and what is the role of loops in coronal heating? What are the scale sizes relevant to turbulent viscosity? What is the horizontal structure of the transition region between the chromosphere and corona? Do weak magnetic fields exist or are all fields concentrated into high-field filaments of 1200 Gauss? What is the structure and dynamics of strong field flux tubes? (Papers by Hudson; Zirin; Ulrich; Davis and Krieger; Datlowe; Title; Baum and Bratenahl.)

(3) Only a Solar Probe can sample solar energetic particles before their spectra are distorted by propagation effects. Such measurements are necessary to understand the acceleration, possible storage, release, and propagation of energetic ions and electrons. (Papers by Gloeckler; Lin; Van Hollebeke; Roelof et al.; Stevens et al.)

(4) A Solar Probe can get close enough to the Sun to detect neutrons released in energetic, nuclear processes occurring at or near the solar surface. (Papers by Simpson; Ahluwalia.) Although solar neutron observations could also be made from a spacecraft orbiting the planet Mercury, a Solar Probe is needed to investigate the lowest energy neutrons.

(5) Only a Solar Probe which goes well inside 20 solar radii can determine how the solar wind is accelerated. Other very important questions which can be answered by solar-wind measurements at this distance include: What are the sources of the low-speed solar wind? Why is the helium concentration so variable and its velocity often greater than the hydrogen velocity? What is the rate of angular momentum loss by the Sun? (Papers by Feldman; Gurnett; Barnes; Belcher et al.; Ogilvie; Kohl.)

(6) Only a Solar Probe can experimentally test the validity of the many models now in use for interpretation of remotely observed solar phenomena and interplanetary phenomena observed near 1 AU. (Papers by Vesecky et al.; Schulz et al.; Dryer et al.) Aside from theoretical models, our knowledge of what happens between 0.3 AU and the visible solar atmosphere is based on the interpretation and modelling of radio data. (Papers by Coles et al.; Croft.)

(7) Only a Solar Probe can obtain the data needed to understand the dynamics and physical changes of interplanetary dust particles near the Sun. These data have implications beyond the solar system in understanding the evolution of all circumstellar dust. (Papers by Hanner; Torr et al.; Weinberg and MacQueen.)

(8) Finally, given the communications and tracking systems and, perhaps, the on-board hydrogen maser needed to determine the solar quadrupole moment and angular momentum vector, the Solar Probe will be by far the best platform in space during the 1980's (and probably the 1980's) to use in the search for low-frequency gravity waves from cosmological sources. (Papers by Thorne; Estabrook; Callahan; Vessot and Levine; Douglass.)

No consensus was reached at the Workshop on the optimum orbital parameters for a Solar Probe. For the measurement of J_2 and for communication with the spacecraft, a high-inclination orbit is best. (Papers by Anderson and Lau; Bender.) However, an in-ecliptic, retrograde orbit has many advantages for observations of fields and particles. (Paper by Stevens.) Ideally, the Solar Probe Mission would include more than one spacecraft, at least one of which was spin stabilized.

The general tenor of the Workshop was one of high enthusiasm for a Solar Probe Mission and of hope that NASA will continue with the development program needed to precede the implementation of the mission. Key development items are: the adaption of a drag-free control system to a spacecraft of this type (Paper by Everitt and DeBra.); the further development and testing of heat shields, possibly with small holes in them (Papers by Davis and Krieger; Datlowe.); further analysis of the accuracy with which J_2 and the solar angular momentum vector can be determined (Papers by Anderson and Lau; Reasenberg and Shapiro; Douglass.); development of an X-band transponder and, perhaps, a K-band communications system; possible development of an on-board hydrogen-maser clock for such a long mission and of a 4-link communication system (Paper by Vessot and Levine.); study of alternative power supplies which don't interfere with solar neutron observations (Paper by Simpson.); and changes to scientific instruments to meet the environmental requirements and observing strategies of the Solar Probe Mission.

Acknowledgments: The Solar Probe Science Workshop and the publication of this book were sponsored by the National Aeronautics and Space Administration under contract NAS7-100 with the California Institute of Technology. I thank H. S. Hudson, L. A. Fisk, and R. W. Davies for organizing the sessions on the Sun, energetic particles, and gravity, respectively. The logistics and creature comforts of the Workshop were very ably taken care of by Carol L. Snyder of JPL. I also wish to thank Dianne Weitzel for her superb, rapid efforts of correcting or typing most of the papers in this report.

SECTION II.

THE SOLAR INTERIOR AND GENERAL RELATIVITY

THE IMPORTANCE OF DETERMINING THE SOLAR QUADRUPOLE MOMENT

N78-32965

Ian W. Roxburgh

Queen Mary College
University of London
Mile End Road
London, E14NS, England

"...it is reasonable to hope that in the not too distant future we shall be competent to understand so simple a thing as a star."

A. S. Eddington, Internal Constitution of the Stars 1926

This concluding remark to Eddington's treatise on stellar structure sets the scene; fifty-two years later we are still struggling to understand such a simple star as the Sun, with average mass, average age, average in most properties. Yet it cannot be said that we have achieved Eddington's goal; indeed, just the contrary. Because we are able to observe not only the integrated light from the surface, but fine structure, and with the rise of neutrino astronomy, the deep interior as well, current theoretical modelling of the Sun is in difficulty. We need additional observations to place constraints on the plethora of conjectures that have been advanced to explain the low neutrino flux and to guide the solar model builder in his search to understand the deep interior of the Sun. A determination of the magnitude of the solar quadrupole moment, J_2, is one such constraint; it places limits on the density, angular velocity and magnetic field distributions inside the Sun.

Solar Models The first model of the Sun was developed by J. Homer Lane in 1869. Following Kelvin's work on the structure of the Earth's atmosphere, Lane proposed that the Sun was sufficiently well mixed by convection currents to be

10
9

adiabatic, $P = K \rho^{5/3}$. Combining this with the condition of hydrostatic support, he determined the internal density and temperature distribution in the Sun (Figure 1). While we no longer believe that the Sun is adiabatic, if it were and rotated uniformly with its surface angular velocity, the quadrupole moment would be $J_2 = 2 \times 10^{-6}$.

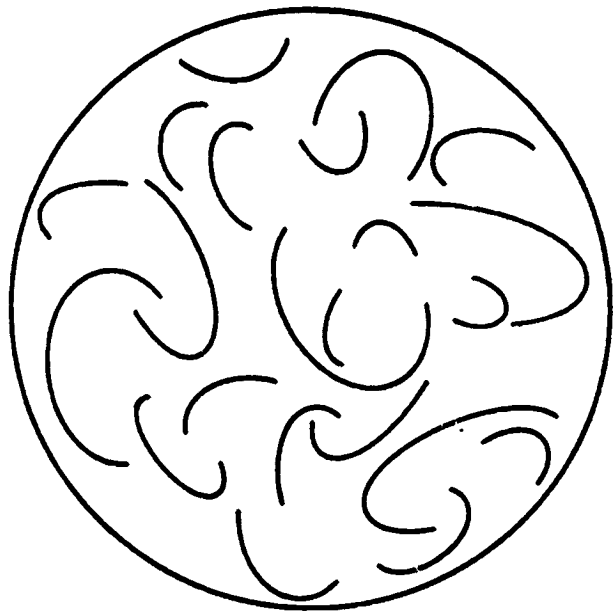


Figure 1 First Solar Model by J. Homer Lane (1869). This model was adiabatic throughout.

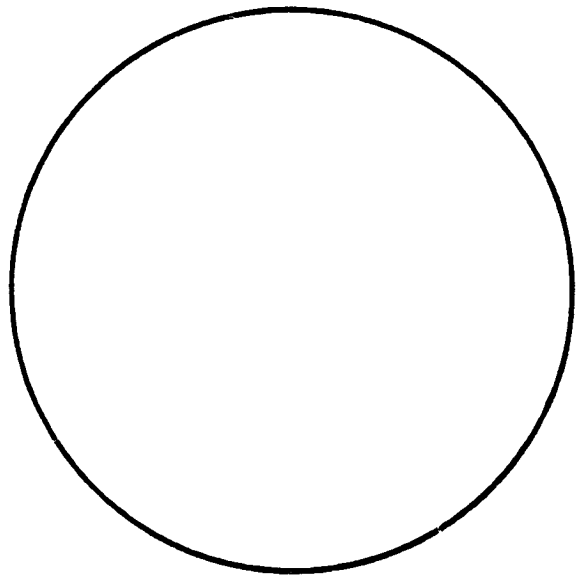


Figure 2 The Eddington Model (1926). In this model all energy was carried by radiation.

The next step came with the application of the theory of radiative transfer to the solar interior by Arthur Eddington (1926). Eddington's standard model was completely radiative (Figure 2) and the pressure and density were related by $P = K \rho^{4/3}$. For such a model, uniform rotation gives a quadrupole moment $J_2 = 2 \times 10^{-7}$. We now realize that both convection and radiation contribute to energy transport in st[illegible]nd that there is a balance between energy production by nuclear reactions in the center and the ability of the star to transport this energy. In the cool surface layers of the Sun, the energy cannot be carried by radiation without such a steep temperature gradient that convection sets in which, in the present solar models, occupies the outer 20% of the radius. In the inner regions of the Sun, the energy transport is by radiation, and the energy is produced by the conversion of hydrogen to helium through the proton-proton chain (Figure 3).

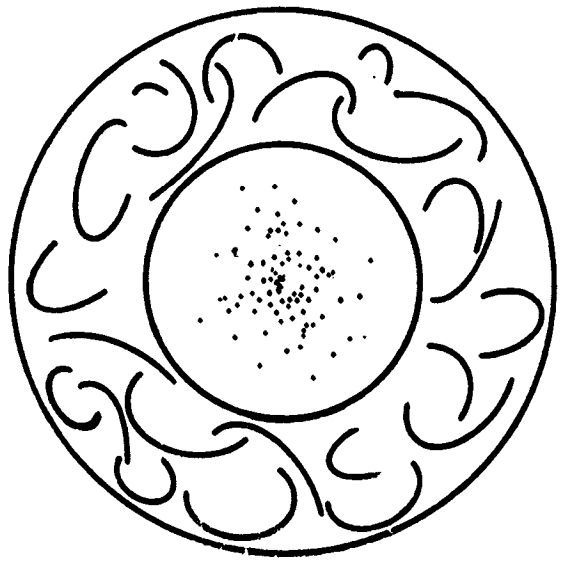


Figure 3 The 'Standard solar model'. The outer 20% of the radius is convective; the inner regions are helium rich due to the conversion of hydrogen to helium.

It is perhaps worth emphasizing the actual procedure used to build a model of the present Sun. While we have detailed knowledge of the surface

abundance of most elements, the abundance of helium, Y, is unknown: The 'theory' of turbulent convection used to model the surface convection zone contains a free parameter α, the ratio of the mixing length (or mean free path of turbulent eddies) to the density scale height. These two paramenters α and Y are then adjusted so that an initially homogeneous spherical Sun of constant mass has the observed luminosity and radius at the present age of the Earth (4.7×10^9 yr). Not surprisingly, with two free parameters, the models can be adjusted to fit two observations! It is this model that predicts a neutrino flux of 4.7 SNU to be compared with Davis's (1978) determination of 1.6 ± 0.4 SNU. If the Sun rotates uniformly, this gives $J_2 = 9 \times 10^{-8}$.

The obvious inference from the neutrino experiment is that the Sun is cooler in the center than predicted by the standard model, but the dilemma cannot be resolved by just lowering the central temperature since the total energy produced would not then be equal to the observed luminosity. Since the neutrinos are produced by the ^8B branch of the p-p chain, which is more temperature sensitive than the main energy producing branches, what is needed is a change in the relative temperature profile - cooler in the center, hotter further out. This has been achieved by a number of esoteric models (c.f. Rood, 1978), some of which have a rapidly spinning central core, or a large central magnetic field, some of which have central mixing increasing the central hydrogen content, and some of which have low heavy element abundance. In all these models the variation of density with radius is different from the standard solar model to give a different quadrupole moment, and in some there is a distorting force, rotation, or magnetic fields that will also affect the quadrupole moment. Clearly, any information we can obtain from the solar probe will provide a useful constraint on such models.

Is the Sun rotating uniformly? There is no convincing answer to this at the present time. It might be conjectured that the Sun is threaded by a magnetic field which ensures approximately uniform rotation, but this is unlikely to be the case. The surface field changes polarity on a short (11 year) time scale and will therefore be severed from any field in the interior. It is quite probable therefore that beneath the convective zone is a shear layer where weak turbulence couples the exterior to the interior region, thus transporting angular momentum from the interior to the surface where it is lost in the solar wind. At the same time, the shear layer gradually mixes surface material down to regions where Lithium is burnt, thus explaining the low surface abundance of Lithium. This would fit in with observations of the decrease in rotation speed and Lithium abundance with age as observed in solar type stars in the Hyades, Pleiades and solar neighborhood. A simple estimate (Dicke 1972, Roxburgh 1976) would suggest that the bulk of the Sun, $r < 0.6\ R_\odot$, is rotating with about twice the surface speed; the quadrupole moment would then be about 7×10^{-7}.

If the interior and exterior are not magnetically coupled, the possibility exists that the quadrupole axis is not coincident with the rotation axis of the surface layers. This possibility should be allowed for in orbit analysis of the solar probe.

Does the interior have a strong magnetic field? We know very little about the dynamo generation of magnetic fields in solar conditions except that we observe magnetic activity in the solar surface layers, so it is reasonable to conjecture that magnetic fields are produced in the other turbulent convective zones, and solar model calculations predict that such zones existed in the young Sun. How strong would such fields be? What would be the mean direction of the field? We just don't know, but fields of the order of 10^5 Gauss could produce a measurable quadrupole moment.

What if the Sun had no interior magnetic field? In this case the angular momentum lost in the solar wind would slow down the convective zone and the shear layer underneath would be unstable to weak turbulence (Goldreich and Schubert 1967, Fricke 1968). If this turbulence transports angular momentum, then the angular velocity must increase inwards at least as fast as the marginally stable state and this gives a quadrupole moment as large as 10^{-5}.

What then is the Sun really like inside? It should be clear from the above discussion that the honest answer is "we don't know" and this in itself is a justification for determining the quadrupole moment. However, I would like to conjecture that it is similar to Figure 4. The inner regions are mixed by weak turbulence driven by a combination of the ^{3}He instability and magnetic oscillations; this gives a higher central hydrogen abundance and a small convective core. The moment of inertia is therefore larger than the standard solar model. The inner bulk of the Sun is magnetically connected and rotates at about twice the surface speed; in between the outer convective zone and the inner region is a shear zone of weak turbulence. Such a model has a quadrupole moment of about 7×10^{-7}. The values of J_2 for different models are given in the following table:

Table 1

PREDICTED VALUES OF THE SOLAR QUADRUPOLE MOMENT

Lane (1869)	Eddington (1926)	Standard Sun	Mixed Core	Non-Standard Model
2×10^{-6}	2×10^{-7}	1×10^{-7}	2×10^{-7}	7×10^{-7}

Conclusions

The expected quadrupole moment of the Sun lies between 10^{-8} and 5×10^{-7}, but could even be larger. A determination of its magnitude and orientation would be a valuable constraint on solar models and perhaps help us along that road that Eddington thought would be so easy to travel but has turned out to be so difficult.

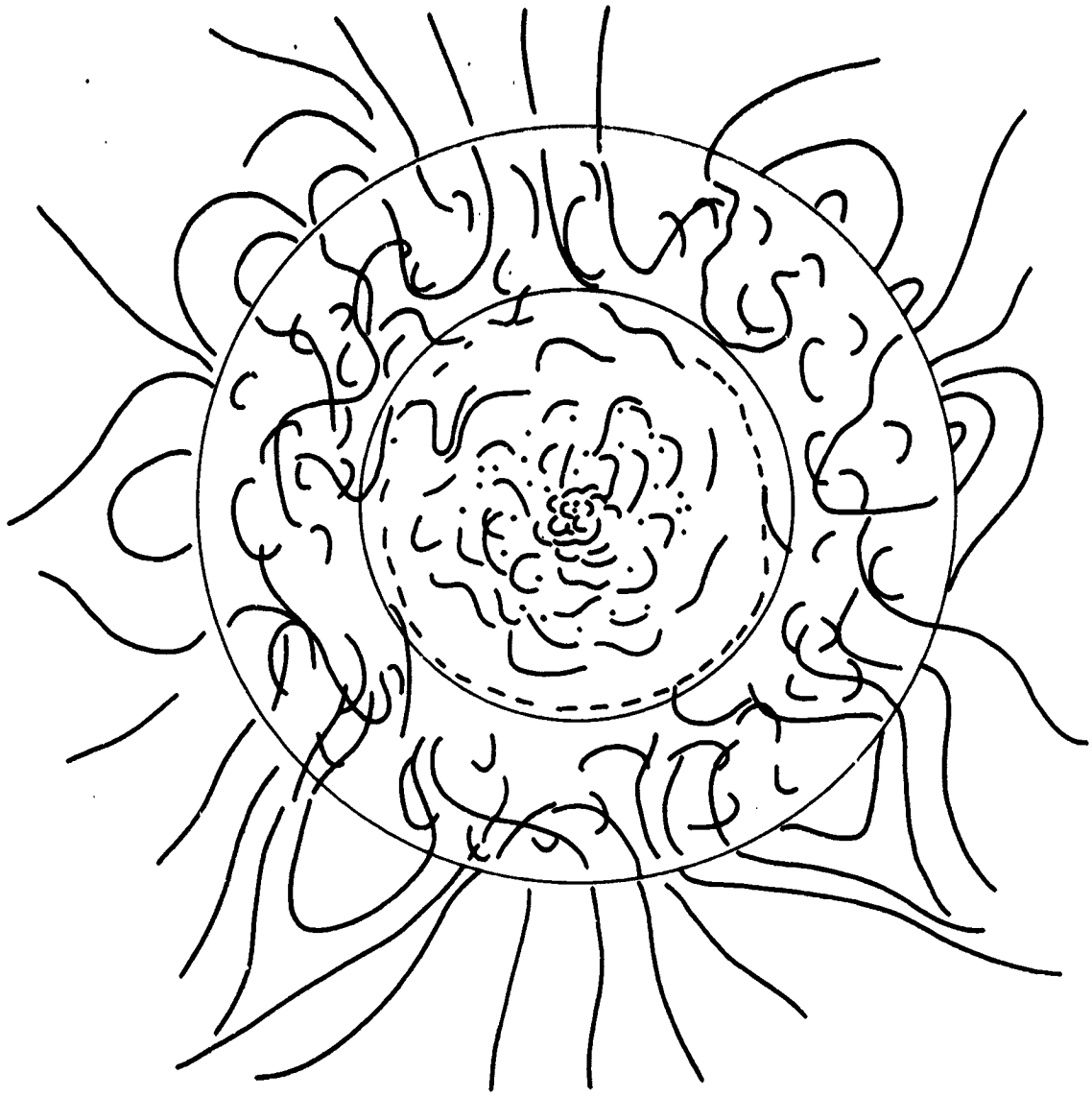


Figure 4. A conjectured solar model. Slow diffusion mixes the inner core increasing the hydrogen content, lowering the temperature, and producing a small convective core. The inner region is magnetically linked and rotates about twice as fast as the surface convective zone. The inner region is separated from the convective zone by a shear layer with weak turbulence transporting angular momentum outwards.

REFERENCES

Davis, R. (1978) in Brookhaven Conference on Solar Neutrinos, ed. G. Friedlander (in press).

Dicke, R. H. (1972) Astrophysical Journal, 171, 331.

Eddington, A. S. (1926) Internal Constitution of the Stars, p. 393, Cambridge.

Fricke, K. (1968) Zeit. fur Astrophysik, 68, 317.

Goldreich, P. and Schubert, G. (1967) Astrophysical Journal, 150, 571.

Lane, H. J. (1869) Amer. J. Sci., 50, 57.

Rood, R. (1978) in Brookhaven Conference on Solar Neutrinos, ed. G. Friedlander (in press).

Roxburgh, I. W. (1976) in Basic Mechanisms of Solar Activity, ed. Bumba and Kleczek, p. 453, Reidel.

Roxburgh, I. W. (1978) in Brookhaven Conference on Solar Neutrinos, ed. G. Friedlander (in press).

Possible Measurement of $J_{2\odot}$ with a Solar Probe

N78-32966

R. D. Reasenberg and I. I. Shapiro
Department of Earth and Planetary Sciences
Massachusetts Institute of Technology
Cambridge, Mass. 02139

Abstract

We have examined the possible use of radio tracking of a solar probe to estimate accurately the coefficient, $J_{2\odot}$, of the second zonal harmonic of the sun's gravitational field. Our preliminary conclusion is that $J_{2\odot}$ can be estimated with a standard error, $\sigma(J_{2\odot})$, of 10^{-8} or less, provided that the probe is equipped with a suitable "drag-free" system to compensate for the effects of non-gravitational accelerations. For signal paths that pass near the sun, dual-band ranging to the probe can provide the tracking accuracy needed to insure that $\sigma(J_{2\odot}) < 10^{-8}$. It might even be possible to achieve such accuracy with a single-band radio uplink and a dual-band downlink. Doppler-shift measurements may not be sufficiently reliable in view of the possibility of frequent "loss of lock" of the signal when the probe is near the sun. Such a condition, brought about mainly by turbulence in the corona, could be a serious impediment to the useful analysis of measurements of Doppler shift, but would have little effect on the analysis of measurements of range.

A determination of $J_{2\odot}$ with this accuracy could provide an important test of theories of the solar interior and would be of very substantial indirect benefit for some solar-system tests of theories of gravitation.

I. Introduction

A spacecraft that passes close by the sun offers a unique opportunity to probe the structure of the solar interior. Our present experimental knowledge of the interior is based, almost exclusively, on measurements of its radius, mass, luminosity, surface composition, rotation rate, and neutrino flux. The remainder of our knowledge rests purely on theory. The predictions of the neutrino flux, however, are in very serious discord with theoretical models based on the other observations: The observed flux is less than one quarter (90% confidence limit) of the flux expected on the basis of these standard models of the sun's interior (Bahcall and Davis, 1976).

With a solar probe that passes close to the sun, one has the opportunity to obtain an additional experimental datum: the coefficient, $J_{2\odot}$, of the second zonal harmonic of the sun's gravitational field. This coefficient is determined largely by the radial distributions, within the sun, of the density and of the rotation rate. What light could a measurement of $J_{2\odot}$ shed on the solar neutrino problem? The answer to this crucial question is not yet clear. Astrophysicists concerned with the theory of stellar interiors have not had to consider experimental constraints on $J_{2\odot}$. Thus they have not devoted substantial effort to a determination of its utility as a discriminator between various characteristics of theoretical models, for example, the fraction of heavy elements within the sun. The possibility of soon obtaining a value of $J_{2\odot}$ with an uncer-

tainty of $\sim 10^{-8}$ or less, will encourage theorists to examine this question carefully, not only with respect to the solar neutrino problem per se, but with respect to all other possibly relevant aspects of models of the solar interior. This uncertainty would be about 10% of the value expected theoretically for J_{20}, based on a solar rotation rate independent of radius. Our simplified calculations, for standard values of the sun's density as a function of radius, yielded values for J_{20} between 5×10^{-8} and 2×10^{-7} (see, also, Colombo et al., 1977; and Roxburgh, 1978).

By passing close to the sun, a solar probe also will be subject to relatively large relativistic effects. However, if the probe operates through only a single flyby of the sun, the direct measurement of relativistic effects on the trajectory will likely not be of sufficient accuracy to add significantly to experimental verification of these predictions of general relativity.* The indirect benefit will nonetheless be large. At present, the main limitation in the interpretation of the perihelion advance of the orbit of Mercury in terms of a test of general relativity (or more generally as a test of non-Newtonian gravitational effects) concerns the uncertainty in the contribution to this advance of the second zonal harmonic of the sun's gravitational field. This contribution is proportional to J_{20}. Thus an accurate determination of J_{20} via a solar probe would have the indirect effect of improving at least several-

*Note, however, that a significant redshift experiment is possible with the placement of a hydrogen-maser frequency standard on a solar probe, as discussed for example by Vessot (1975) and by others in these Proceedings.

fold the stringency of this test of relativity based on measurement of the perihelion advance of Mercury's orbit.

On the other hand, should it be possible to track the solar probe through several close approaches to the sun, the direct contribution to tests of the predicted relativistic effects on particle motion could be substantial provided that the "process-noise" contribution to the spacecraft's motion could be kept sufficiently small or monitored sufficiently well.

What conditions must be set on a solar probe mission in order to realize the opportunity to measure $J_{2\odot}$ to within 10^{-8}? Precise statements can not yet be made reliably, but certain general constraints can be noted. The trajectory of the probe must be sufficiently sensitive to the effects of $J_{2\odot}$. This condition can be met if the probe passes within a few solar radii of the center of the sun. But the trajectory must not be perturbed in a manner that obscures the effects of $J_{2\odot}$. It seems clear from crude estimates of the orders of the magnitudes of the nongravitational accelerations involved that some type of "drag-free" system will have to be employed as described in more detail below. Finally, we must be able to track the probe with sufficient accuracy to determine the effects of $J_{2\odot}$ at the desired level. This tracking requirement can be met with a radio system provided that dual-band capability exists at least on the downlink (probe-to-earth path). If the standard frequencies within the S and X bands are used, ranging capability will also be required to ensure that the solar corona does not prove a serious impediment to

tracking. Although Doppler tracking becomes intermittent, because the receiver loses lock near the sun, ranging is relatively immune to the effects of the corona and has been accomplished successfully at S band within about 0°5 of the sun's center (Shapiro et al., 1977).

In the next section, we discuss in more detail, but still qualitatively, these effects of trajectory and tracking uncertainties on the accuracy of the estimation of $J_{2\odot}$. Section III contains a quantitative, but simplified, analysis for the determination of the standard error $\sigma(J_{2,\odot})$ in the estimate of $J_{2\odot}$ for a particular case of probe trajectory and tracking schedule. In the last section, we make several suggestions for further study.

II. Trajectory and Tracking Constraints

The accuracy of the estimation of $J_{2\odot}$ from tracking a solar probe depends critically on two factors: the characteristics of the non-gravitational accelerations that affect the trajectory of the probe and the effects of the solar corona on the tracking signals. Here we include several remarks on each factor.

A solar probe will experience nongravitational accelerations as the result of many processes. Solar radiation pressure will provide the largest, externally driven, nongravitational acceleration. A large fraction of this acceleration will be relatively easy to model as it will be proportional to $\vec{r}/|r|^3$, where $\vec{r}$ is the sun-spacecraft vector. However, that fraction is not large enough for our purposes because of the enormous increase in the radiation pressure near the sun. As examples of the difficulties in the development of an adequate model, we note that: (1) The tem-

perature and time dependence of the reflection characteristics of the heat shield on the probe may be difficult to model adequately; (2) The radiation reaction from the back of this shield will be difficult to calculate accurately; and (3) The effects of changes in the orientation of the spacecraft, within its limit cycle, on the radiation-pressure acceleration may also be difficult to calculate with sufficient accuracy.

In addition to the radiation pressure, other possibly important nongravitational accelerations include: (1) the solar wind, which is expected to be highly variable and nonradial, especially as seen by the moving spacecraft; (2) spacecraft gas leaks, especially from the attitude control system; and (3) effects due to electric or magnetic forces on a possibly charged spacecraft.

To combat the very serious deleterious effects of the nongravitational accelerations, the spacecraft should include a drag-free system (Everitt and DeBra, 1978). However, the drag sensor and thruster servo system will be imperfect. Some possible mechanisms are: (1) The servo loop could contain nonlinearities such that the spacecraft would experience a net acceleration proportional to the external nongravitational acceleration; (2) The servo sensor could cause random accelerations of the proof mass which in turn could cause corresponding spacecraft accelerations which need not represent a white noise process; (3) Through a combination of processes, the spacecraft could experience a <u>random</u> acceleration proportional to the net

external nongravitational acceleration; and (4) Environmental influences, such as from cosmic rays, could affect adversely the sensor-servo system (see Everitt and DeBra, 1978). We also note that the drag-free system must be programmed to operate over a wide dynamic range, unlike the situation of the DISCOS in earth orbit (Staff, 1974), since the magnitude of the non-gravitational accelerations will vary over several orders of magnitude during the crucial, near perihelion, parts of the mission alone.

How will the imperfect compensation of nongravitational forces affect the estimation of $J_{2\odot}$? First we note that a drag-free system is generally described in terms of a bound on the servo error, i.e., on its acceleration error, sometimes referred to as process noise. However, this description is not sufficient for our purposes. We require the power spectral density of this process noise. A few examples will make this need clear: (1) Suppose that the process noise is merely a constant acceleration bias. Then, if this fact were known, we could estimate that bias accurately, along with the other relevant parameters such as the initial conditions of the orbit of the solar probe, from tracking data obtained when the probe was far from the sun. The effects of such process noise on the accuracy of the estimate of $J_{2\odot}$ would clearly be benign; (2) Suppose that the spectral density of the process noise is uniform up to a cutoff frequency high compared to the highest significant frequency component attributable to $J_{2\odot}$ in the data obtained from tracking the probe. Suppose further that the total power of this process noise is comparable to that of $J_{2\odot}$ in its effect on the

trajectory of the probe. Then the net effect of the process noise on the estimation of $J_{2\odot}$ will be small; under these conditions, most of the process noise is not in the vicinity of the significant spectral components of the signature of $J_{2\odot}$ and is therefore relatively harmless; and (3) Suppose, finally, that the power of the process noise is concentrated around the principal frequencies of the signature of $J_{2\odot}$. This is the worst case. Thus, if the total power of this process noise were comparable to that in Example (2), the experiment would be useless.

When we attempt to analyze the entire experiment, the situation becomes more complicated. Each parameter to be estimated in general contributes a distinct signature to the data, and these signatures can be Fourier analyzed. Should the process noise dominate the signature of a parameter in one region of the spectrum, but not in another, then the estimator should extract information predominantly from the latter spectral region. This manner of extraction is an essential characteristic of a Kalman filter that includes a correct representation of the spectral density of the process noise. Thus one requires that (1) the power spectral density of the probe's stochastic acceleration be "sufficiently small" in the regions of interest, and (2) this spectral density be known.

The tracking signal is, of course, affected by the solar plasma such that the increase (decrease) in group (phase) delay is approximately proportional to N/f^2, where N is the columnar electron content along the ray path, and f is the signal frequency. Unfortunately, N can not be predicted with sufficient accuracy. To reduce the plasma effects on the radio-tracking

observations to a benign level, in principle either f can be made sufficiently large or the tracking can be accomplished simultaneously at two suitable frequencies. Unfortunately, the equipment required for use of a sufficiently high frequency is not now available for spacecraft tracking. The second approach appears viable, however, even with the use of the conventional S and X band frequencies. With this dual-band tracking, the group delay and the changes in the phase delay can be monitored at two frequencies simultaneously. Over a short period of time, the changes in phase delay can usually be measured more accurately than those of the group delay. However, when the tracking signal passes near the sun, it is difficult to maintain two-way "lock" of the signal, especially for S band. Each loss of lock invalidates the associated Doppler measurement and, under conditions of severe coronal turbulence, these losses may become so frequent as to render Doppler data nearly useless. On the other hand, the group-delay, or range, measurements are relatively immune to loss of lock. Each range measurement stands alone, independent of neighboring measurements. Moreover, the ability to make accurate range measurements when the signal passes near the sun has already been adequately demonstrated with the Viking spacecraft (Shapiro et al., 1977). Even though the dual-band capability was available only on the downlink from the Viking orbiters, it was still possible to correct for the plasma effects in the S-band measurements of range to the Viking landers. With the signals passing as close as four solar radii from the center of the sun, the corrections for plasma were sufficiently accurate to allow the equivalent "vacuum" range

between the earth and the landers to be determined to within $\sim$0.2 μsec. The corrections for the effects of the solar corona could be made more accurately for a solar probe than was possible for Viking for two reasons. First, the paths of the uplink and downlink signals will be very close together in space and time near the sun because the probe will be much closer to the sun than were the Viking spacecraft. Second, the downlink signal from the probe itself will be dual band whereas, for Viking, the lander S-band range measurements had to be corrected on the basis of the dual-band measurements of range to the orbiters. We are assuming here that (i) the coronal turbulence will be no more severe for the tracking of the solar probe than it was for the tracking of the Viking spacecraft; and (ii) the perhaps severe thermal effects on the ranging transponder in the probe can be calibrated with uncertainties well under 0.1 μsec.

In summary, a successful measurement of $J_{2\odot}$ seems to depend critically on the use of an adequate drag-free system whose required process-noise characteristics have yet to be specified quantitatively. The Viking experience seems to indicate that with an S-band uplink signal and with the dual-band downlink signals, vacuum-equivalent ranges to the solar probe could be obtained with uncertainties no greater than $\sim$0.1 μsec, even for signals that pass about four solar radii from the sun's center. This accuracy, as indicated in the next section, should be adequate to insure that $J_{2\odot}$ can be estimated with an uncertainty of $\sim 10^{-8}$ or less, provided a suitable trajectory is chosen and an adequate drag-free system is employed.

III. Accuracy of Estimation of $J_{2\odot}$

In early 1976, a preliminary study of a possible solar-probe mission was conducted at M.I.T. to determine the accuracy with which $J_{2\odot}$ might be determined from the analysis of radio-tracking observations of the probe. A single spacecraft trajectory and a single schedule of observations were considered. These are described in Tables 1 and 2, respectively. The spacecraft was assumed to be drag free, i.e., to follow a purely gravitational trajectory, and the observations were assumed to be disturbed by a corona that was time invariant -- both oversimplifying assumptions.

The spacecraft was assumed to pass by Jupiter and then the sun at a perihelion distance of one solar radius ("sun grazer"). The orbit was chosen so that the earth-spacecraft line was inclined by about 45° to the orbital plane of the spacecraft at the time of perihelion passage. The radio-tracking observations were assumed to be distributed from just past the time of the spacecraft's passage by Jupiter until ten days past the spacecraft's passage through perihelion. These observations were scheduled at a rate of one per five days until 23 days before the time of perihelion passage and at a rate of five per day thereafter, in a simple attempt to represent the more dense tracking coverage that would undoubtedly occur with the spacecraft near the sun.

Each observation was assumed to be made at an X-band radio frequency and to consist of a measurement of range, with a 1 μsec standard error, and a measurement of Doppler shift, with a 1 mHz standard error. The latter error is equivalent, for example, to

Table 1

Characteristics of Solar-Probe Trajectory

Orbital Element	Value (1950.0 Equator and Equinox System)	
Semimajor Axis	2.6	a.u.
Eccentricity	0.9982	
Inclination	113.5	deg
Argument of Perihelion	0.0	deg
Longitude of Ascending Node	0.0	deg
Time of Perihelion Passage	4 February 1972	
Minimum Distance from Sun	1.006	solar radii
Inclination to Ecliptic	90	deg
Angle Between Earth-Probe line and Sun-Probe line at Probe Perihelion	~45	deg

Table 2

Schedule of Probe Observations

Rate (observations/day)	Total Number		Time Interval
	Range	Doppler Shift	
0.2	148	148	763 to 23 days before encounter
8	265	265	23 days before to 10 days after encounter

that from a series of 30 closely spaced, Doppler-shift measurements, each having an independent error with a standard deviation of 0.1 mm/sec (Randolph, 1978). The standard error assumed for the range measurements is quite conservative, even when the ray paths pass near the sun, in the light of the Viking results obtained a year after this error analysis was completed (see section II). On the other hand, the standard error assumed for the Doppler-shift measurements may well be too optimistic for observations made with the spacecraft near the sun because of the possibility of frequent loss-of-lock during radio tracking.

Table 3 shows the 18 parameters that were estimated in the covariance analysis of these observations. Of these, twelve parameters represent the initial conditions for the orbits of the spacecraft and the earth. The pair of parameters, T and S, represent the relativistic effects on the spacecraft's trajectory and the radio signal, respectively. These parameters are equivalent to the usual pair, β and γ, except that, as disclosed by our preliminary analysis, the estimates of T and S are much less highly correlated than are those of β and γ. The plasma constant is the coefficient of a simple model of the radial dependence of the coronal electron density; this parameter was included to represent crudely the effect of imperfect calibration of the plasma from dual-band measurements. Similarly, the parameter for range measurement bias was included to account for errors in the calibration of the range measurement system. This bias was here assumed to remain constant throughout, an

Table 3

Model Parameters

Initial Conditions of Probe Orbit	6
Initial Conditions of Earth's Orbit	6
Light-time Equivalent of the Astronomical Unit	1
Coefficient of Model of Solar Corona	1
Effect of General Relativity on Trajectory	1
Effect of General Relativity on Signal Propagation	1
Bias in Range Measurements	1
J_{20}	1
Total	18

assumption which, in practice, would be invalid. The remaining two parameters are the light-time equivalent of the astronomical unit and $J_{2\odot}$. No a priori constraints were applied to any of these parameters in the analysis.

Based on these measurement and parameter sets, we performed three sets of covariance analyses: one with the standard errors for the measurements as described above ("nominal"), one with the standard error for the Doppler measurements increased to 1 Hz ("range dominated"), and one with the standard error for the range measurements increased to 1 msec ("Doppler dominated"). Since no a priori constraints were applied, the results of these analyses can be scaled; i.e., multiplication of the measurement standard errors in any of the analyses by a given factor results in a multiplication of the standard errors in the parameter estimates by the same factor.

Some of the results of the covariance analyses are summarized in Figures 1 and 2. In Figure 1, we display the evolution with time of the standard error, $\sigma(J_{2\odot})$, in the estimate of $J_{2\odot}$ for each of the three sets of measurement standard errors. As expected, $\sigma(J_{2\odot})$ decreases dramatically as the spacecraft nears the sun. (Note that the discontinuity in the slope of each curve is due to the abrupt increase in the rate at which measurements are made, starting 23 days before the spacecraft reached perihelion.) The values of $\sigma(J_{2\odot})$ for the solutions with range-dominated measurements are the largest. However, the assumed standard error for the range measurements was unrealistically high. If this standard error is lowered to a value of, say,

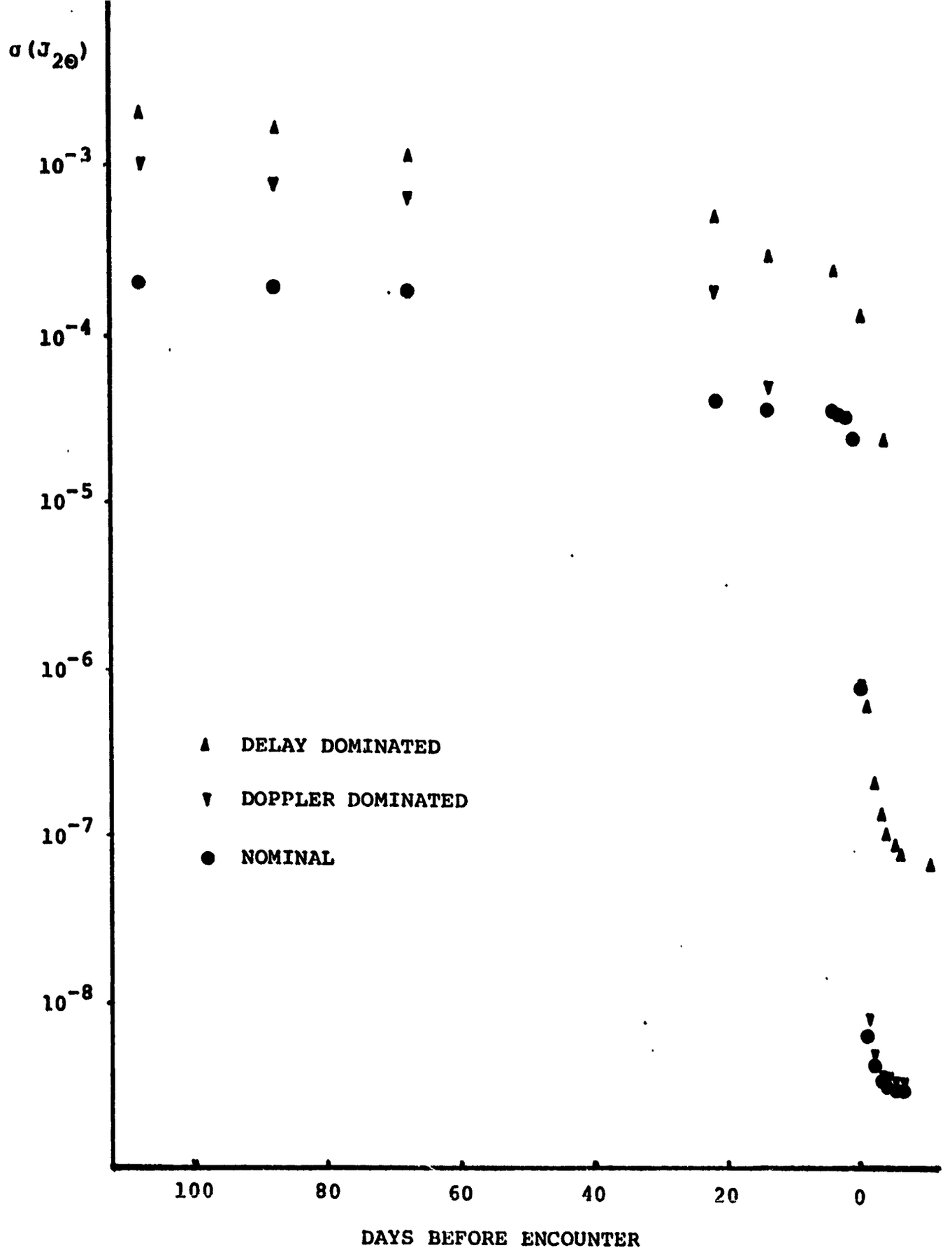


Figure 1. Time evolution of the uncertainty in the estimate of $J_{2\Theta}$ for various measurement standard errors.

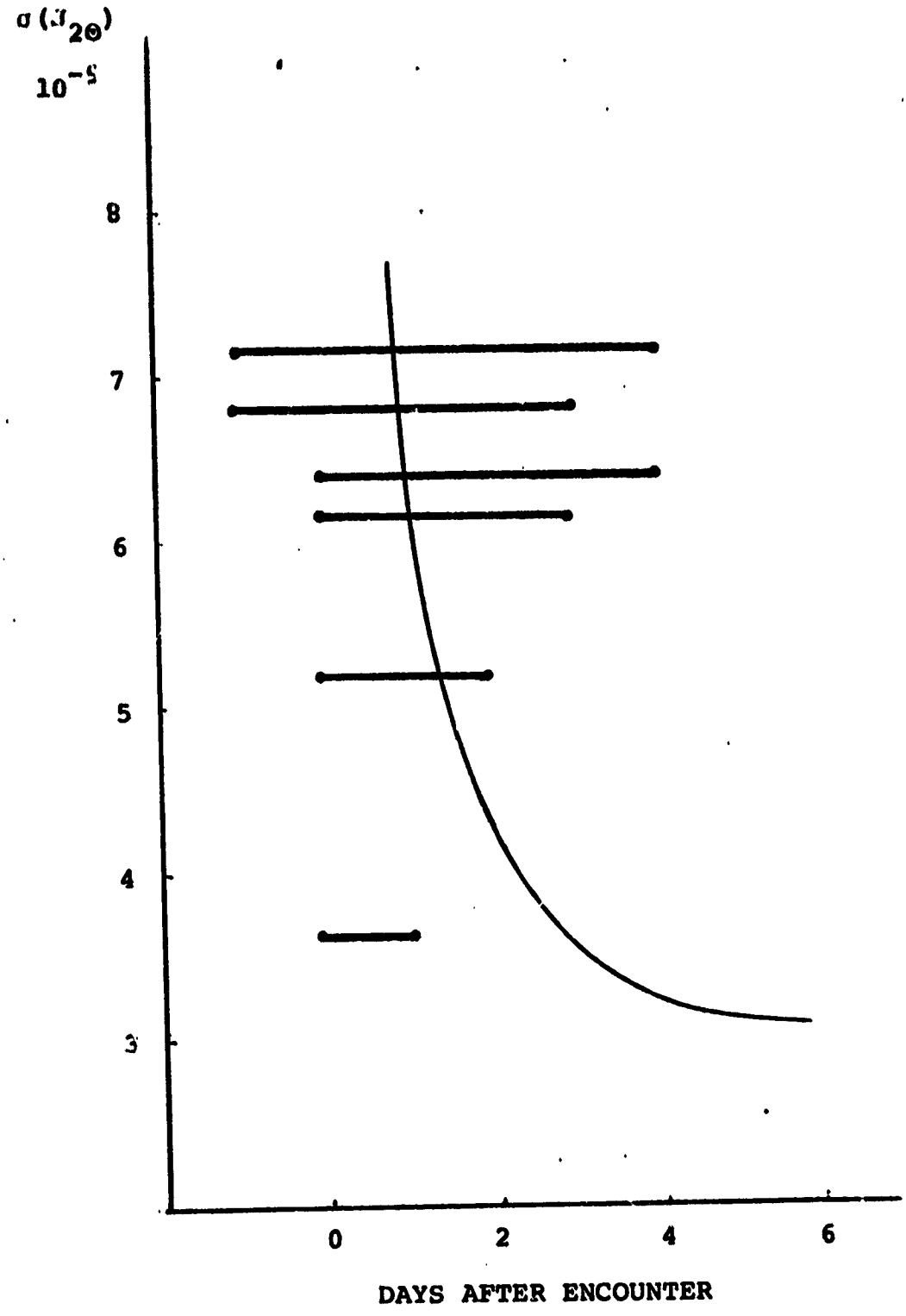


Figure 2. Effect of loss of data on the uncertainty in the estimate of $J_{2\odot}$ for the nominal standard errors.

0.1 μsec (see Section II), we find that $\sigma(J_{2\odot})$ is comparable to that found from the analysis of the Doppler-dominated set of measurements with the 1 mHz standard error. However, as stated earlier, the solution based on the Doppler dominated measurements may be unrealistic since no account was taken of possible loss of lock of the tracking signal due to turbulence in the corona.

Figure 2 shows the evolution of $\sigma(J_{2\odot})$ for the nominal standard errors as data are added after the passage of the spacecraft through perihelion. The horizontal bars show the value of $\sigma(J_{2\odot})$ obtained from the inclusion of all data save those measurements scheduled during the time interval covered by the bar. Thus, the horizontal extent of a bar indicates the omitted span of data and the vertical position of the bar indicates $\sigma(J_{2\odot})$. This figure suggests that a useful estimate of $J_{2\odot}$ can be obtained even if the tracking system performance is substantially degraded as the spacecraft passes by the sun. Of course, this result has validity only if it is possible to obtain accurate measurements <u>after</u> perihelion passage. Thus prudence would require that every effort be made to insure a minimum loss of data before perihelion passage.

The correlations among the estimates of the parameters exhibit no particular surprises. For example, the estimates for the initial conditions of the earth's orbit are only weakly correlated with the estimate of $J_{2\odot}$ from the full data set since the earth almost stands still during the time the

spacecraft is near the sun. Because of the very close passage of the spacecraft by the sun, we also find that the estimates of T and S, and, hence, of β and γ, are only weakly correlated with the estimate of $J_{2\odot}$. The estimates of T and S themselves show that, for the nominal standard errors of measurement, no significant advance over current knowledge of β and γ will be forthcoming. As stated in Section I, the main advantage for testing theories of gravitation through the tracking of a solar probe will lie in its indirect effect -- removal to a large degree of the uncertainty in the value of $J_{2\odot}$ and, in particular, in its contribution to the precession of the perihelion of Mercury's orbit.

For a probe on a trajectory with a larger perihelion than the sun grazer we considered, the value of $\sigma(J_{2\odot})$ will, of course, be greater if all other circumstances are equivalent. The scaling law, however, is not simple; it depends on the schedule of observations as well as on the radial and latitudinal dependences of the trajectory perturbations attributable to $J_{2\odot}$. Without having carried out the relevant, albeit simple, analyses, we cannot conclude reliably that ranging with an S-band uplink and the dual-band downlink would enable a determination of $J_{2\odot}$ to be made with $\sigma(J_{2\odot}) < 10^{-8}$. However, we do feel confident that sufficient accuracy could be achieved if the uplink were also dual band. The presence of a very stable frequency standard on the probe (Vessot,

1978) could enable the tracking to be accomplished with a somewhat higher accuracy. In any event, with dual-band tracking, the main limitation on the accuracy of the determination of $J_{2\odot}$ would probably reside with imperfections in the drag-free system although, clearly, more must be known about this system before any reliable conclusions can be reached.

In summary, the simplified error analyses described above indicate that a solar-probe mission could lead to a determination of $J_{2\odot}$ with an uncertainty of $\leq 10^{-8}$. However, this conclusion depends critically on the use of a suitable drag-free system.

IV. Suggestions for Further Study

The present state of our predictions for $\sigma(J_{2\odot})$ is illustrated in Figure 3. It is clear than an improvement is required in a number of areas, especially since the orbital characteristics of a solar probe are so dramatically different from those of previously flown spacecraft. We mention several such improvements: (1) Computer programs should be developed to treat adequately in error analyses the process noise that may accompany a drag-free system. Further, in view of the great complexity of all of these error-analysis programs, it is essential that they be developed independently by at least two groups and compared to insure that identical results are obtained from the analysis of the same data; (2) The performance characteristics, and their uncertainties, of feasible drag-free systems should be determined by detailed engineering study; (3) The characteristics of feasible tracking systems should be delineated; and (4) Error analyses should

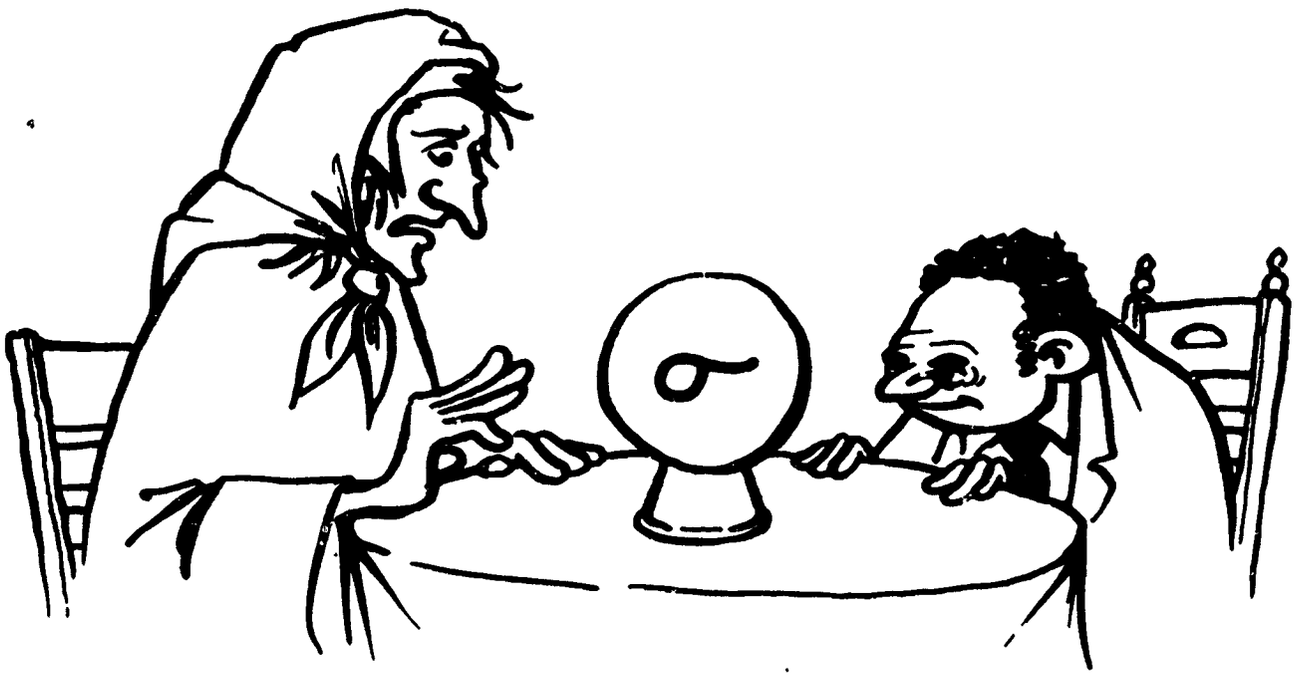


Figure 3. Present state of determination of $\sigma(J_{2\odot})$.

be carried out based on these characteristics. In particular, various trajectories, schedules of observations, extreme coronal conditions, etc. should be considered, including, for example, the possible benefits of auxiliary data, such as differential very-long-baseline-interferometry observations of the spacecraft and of compact, extragalactic, radio sources that are nearby in the plane of the sky. Such auxiliary data might be useful to reduce correlations among parameter estimates and thereby to improve substantially the determination of $J_{2\odot}$. One objective of these studies would be to determine whether a robust experiment could be designed that could produce useful results for $J_{2\odot}$ even under adverse conditions.

V. Acknowledgments

We thank A. Forni for her aid with the covariance analysis and the Jet Propulsion Laboratory for a grant that supported this analysis.

References

Bahcall, J. N., and Davis, R., Science, 191, 264 (1976).

Colombo, G., Lautman, D. A., and Pettengill, G. H. in Probleme Attuale di Meccanica Teorica e Applicata, Accad. Turin, p. 291 (1977).

Everitt, C. W. F. and DeBra, D. B. (1978). These Proceedings.

Randolph, J. (1978). These Proceedings.

Roxburgh, I. (1978). These Proceedings.

Shapiro, I. I., et al., J. Geophys. Res., 82, 4329 (1977).

Staff of the Space Department, Johns Hopkins Applied Physics Laboratory, and the Staff of the Guidance and Control Laboratory, Stanford University, J. Spacecraft & Rockets, 11, 637 (1974).

Vessot, R. F. C. in Atomic Masses and Fundamental Constants 5, Plenum Press, New York, Edited by J. H. Sanders and A. H. Wapstra, p. 312 (1975).

Vessot, R. F. C. (1978). These Proceedings.

D3

N78-32967

GRAVITATIONAL EXPERIMENTS ON SOLAR PROBE

by

John D. Anderson
Eunice L. Lau

Jet Propulsion Laboratory
4800 Oak Grove Drive
Pasadena, CA 91103

ABSTRACT

A covariance analysis has been performed for a solar probe trajectory which encounters the sun at four solar radii. The unknown parameters in the analysis are the six initial cartesian coordinates for the probe, six initial cartesian coordinates for the earth, the astronomical unit, the solar gravitational quadrupole coefficient J_2, and two PPN relativity parameters (β,γ). Errors in the unknown parameters have been computed as a function of standard errors on the radio tracking data and on the nongravitational forces which act on the probe. Results have been obtained for several tracking geometries and for several orbital inclinations to the ecliptic.

The results of the analysis show that the principal scientific result from the radio tracking of a solar probe would be the determination of the quadrupole moment, which would place a constraint on models of the solar interior. We find that Solar Probe is capable of yielding a measurement of J_2 to an accuracy of $\pm 10^{-8}$, and so at this level of accuracy, a number of currently plausible interior models would no longer be viable. However, the measurement requires that nongravitational forces be eliminated to a level of 10^{-10}g, and that a dual frequency tracking capability be made available for purposes of determining the effect of free electrons on the radio propagation.

A secondary scientific result of the tracking would be a determination of the general relativistic precession of the orbital perihelion to an accuracy of ±0.1%. It is likely that a continued analysis of data on planetary dynamics will yield results of greater interest to experimental relativity, but Solar Probe can yield a single measurement of the perihelion advance with just a few hours of encounter data.

I. INTRODUCTION

We were first introduced to the concept of a Solar Probe by G. Colombo (1975) who sugested that such a spacecraft could be thermally shielded and that it could approach the sun at a close enough distance to perform significant gravitational experiments. At about the same time L. Friedman initiated a small preliminary engineering feasibility study of Solar Probe at JPL, and we joined with Colombo in support of that study. As a result, a paper was presented at an experimental gravitation symposium in Pavia (Anderson et al., 1977), which concluded that 1) a solar probe mission with a perihelion distance of four solar radii (4 $R_\odot$) was feasible, 2) the primary gravitational experiment on such a mission would be a determination of the solar gravitational quadrupole moment, and 3) the spacecraft would have to be drag-free to a level of 10^{-10}g. Preliminary telecommunication studies indicated that a Doppler tracking accuracy of 0.1 mm/s could be achieved near the sun, and hence it would be possible to measure the coefficient J_2 in the Legendre expansion of the sun's external gravitational potential energy V to an accuracy of two parts in 10^8.

In all studies to date, it has been assumed that V is static and that all other harmonics besides J_2 are negligible, so that

$$V = -\frac{GM_\odot}{r}\left[1 - J_2\left(\frac{R_\odot}{r}\right)^2 P_2(\sin\phi)\right] \qquad (1)$$

By observing the motion of the probe near the sun with Doppler tracking, it is possible to determine the size of J_2 by means of its orbital effects through the time variations in the heliocentric radius r and latitude ϕ of the probe. Future studies should consider the possibility of a time variation in J_2 to account for normal modes in the sun (Roxburgh; Douglass, this volume), but for now all our studies to date depend on the static potential (Eq. 1) and consider the determination of J_2 along with two post Newtonian relativistic parameters (β,γ) in the following post Newtonian one-body metric (Robertson, 1962).

$$ds^2 = \left[1 - 2(m/r) + 2\beta(m/r)^2\right]c^2dt^2 - \left[1 + 2\gamma(m/r)\right](dx^2 + dy^2 + dz^2) \qquad (2)$$

During 1978, a more detailed engineering study of Solar Probe was initiated at JPL under the technical direction of J. Randolph. We agreed to participate in that study and to use the computer software developed for the Pavia paper to gain a better understanding of the gravitational experiments at solar encounter. Results are reported here, but it should be understood that this is a status report and more work will be carried out in the future in support of the engineering studies. In addition to including a time dependent gravitational potential in the analysis, more work is needed to characterize the nongravitational forces and to specify in more detail the requirements on the drag-free system. Also the implications of adding an on-board clock as proposed by Vessot (this volume) to enhance gravitational experiments needs further study. Finally, the possibility of performing null tests of general relativity, such as bounding the magnitude of preferred frame effects or non-metric terms in the equations of motion, will be considered.

The conclusions that are drawn here from the current studies do not differ substantially from those in the Pavia paper. These are:

1. The coefficient J_2 can be determined to $\pm 10^{-8}$ or better with two-way Doppler tracking and an S and X band dual frequency capability on both the up and down legs of the transmission.

2. The accuracy of the proposed measurement of J_2 is significant, and hence Solar Probe could provide an important boundary condition on theoretical models of the solar interior (Roxburgh, this volume).

3. A meaningful design goal for the drag compensation system is 10^{-10} g, but a significant measurement of J_2 would still be possible even if that goal were not met exactly.

4. The general relativistic precession of the perihelion of Solar Probe could be measured with an accuracy of 0.1 percent or better from a single flyby with drag compensation.

5. The dual frequency (S/X) telecommunication system on the up leg and down leg of the two-way transmission would result in an improved Doppler tracking system over that on the Galileo and Solar Polar missions, with an associated increased sensitivity for

gravitational wave detection far from the sun (Estabrook; Douglass, this volume).

II. GRAVITATIONAL EXPERIMENTS AT SOLAR ENCOUNTER

The primary objective of a gravitational experiment at solar encounter is to determine the quadrupole coefficient J_2 to an accuracy* on the order of $\pm 10^{-8}$. At this level of accuracy, several currently plausible interior models would no longer be viable (Roxburgh, this volume). Uniform solar rotation and a polytropic model of index 3 implies a value for J_2 of 2×10^{-7}, which is probably an upper bound on this coefficient. The most extreme centrally condensed models yield values of J_2 near 8.5×10^{-8}. Therefore a measurement of J_2 to an accuracy of $\pm 10^{-8}$ is very important. Other proposals to measure this coefficient (e.g. Mercury Orbiter) are at least an order of magnitude less accurate (Anderson et al., 1977), so Solar Probe provides a unique opportunity for a definitive measurement of J_2.

Secondary objectives at solar encounter involve the determination of various post-Newtonian gravitational effects on the orbit and the tracking signal (see Nordtvedt, this volume). At least the precession of the perihelion of the orbit is determinable with the two-way tracking and drag-free system needed for the determination of J_2. The measurement of other effects would benefit from a more sophisticated telecommunication system, such as the on-board clock proposed by Vessot (this volume).

III. ASSUMPTIONS FOR ERROR ANALYSIS

Tracking Assumptions

A tracking arc of two months centered about solar encounter is used for the error analysis. A Doppler measurement is simulated every 10 minutes and a range measurement every 20 minutes over this interval. Two-way tracking is assumed, and the errors in each Doppler and range measurement are assumed to be independent. The range error is set at 15 m, a figure that is achievable now with telecommunication systems of the Viking and Voyager class. The Doppler error is expressed as an error in velocity along the line of sight

*All error statements in this paper are expressed in terms of one standard deviation.

(range rate), and is set equal to various values with a range of estimated errors for two-way tracking configurations. Estimates of error are summarized in the following table, which was provided by the DSN Radio Science Systems Engineer (A. L. Berman) as input for the current engineering study under J. Randolph.

DOPPLER ACCURACY ASSUMPTION

Two-way Configuration	Error (mm/s)	Error ($\Delta f/f = 2\ \Delta v/c$)
S band up and down	25	1.7×10^{-10}
S band up, S/X down	5	3×10^{-11}
X band up and down	2.5	1.7×10^{-11}
X band up, S/X down	0.5	3×10^{-12}
S/X bands up and down	0.1	7×10^{-13}

Dual-frequency S/X phase scintillation measurements of the solar wind conducted with the Viking spacecraft in 1976 indicate that the S band fractional frequency stability $\Delta f/f$ (square-root Allan variance) for a closest approach distance of 4 $R_\odot$ and an integration time of 1000 s is $\sim 7 \times 10^{-11}$ (Armstrong et al, 1978). To apply this result to the Solar Probe case, we divide by $\sqrt{2}$ to account for the fact that the spacecraft is at the closest approach point rather than at Mars, and multiply by $(3/11)^2$ to extrapolate to X band. Thus, even without any charged-particle calibration, the fractional frequency stability at X band is $\sim 4 \times 10^{-12}$. Berman's estimate of 1.7×10^{-11} is conservative and his error estimate of 0.1 mm/s (7×10^{-13}) is certainly reasonable for a dual frequency (S/X) calibration on both legs of the transmission. Note, that for this two-way frequency configuration, the error in the Doppler is considerably greater than the limit imposed by hydrogen maser timekeeping. Therefore, far from the sun, the expected error in Doppler should be two orders of magnitude more precise than our assumption for solar encounter, where plasma calibration errors dominate. This fact makes the detection of gravitational radiation a promising experiment for solar probe because that experiment will be conducted at or near opposition (Estabrook; Douglass, this volume).

Non-gravitational Forces

It is assumed that there are non-gravitational forces acting on the probe, primarily from solar radiation pressure, and that they are removed to some degree by an active drag-compensation system, by passive accelerometers, by a sequential adaptive filtering of the data, or by some combination of the three. The best engineering strategy for removing nongravitational forces which contaminate the gravitational experiments at solar encounter has not been determined at this time. However, for purposes of an error analysis, it is not necessary that all engineering and spacecraft control systems be understood in detail. It is sufficient to incorporate a stochastic model for the nongravitational forces in the error analysis, and then to determine the sensitivity of the parameters of interest (e.g., J_2, β,γ) to the magnitude of the error in the stochastic model. To this end, we have assumed that all systematic errors in the nongravitational forces have been reduced to an insignificant level by some combination of the three techniques listed previously. Then what remains is white acceleration noise with standard deviation σ_a which corrupts the determination of the position and velocity of the probe at each data time t_i according to the following rule.

$$\sigma_x = \sigma_y = \sigma_z = \frac{1}{2}\sigma_a (t_i - t_{i-1})^2 \tag{3}$$

$$\sigma_{\dot{x}} = \sigma_{\dot{y}} = \sigma_{\dot{z}} = \sigma_a (t_i - t_{i-1}) \tag{4}$$

The covariance analysis, and hence the error in the determination of each parameter, then becomes a function of the acceleration noise parameter σ_a. In effect the autocorrelation time of the acceleration noise is assumed equal to the time interval between data points, which is 10 minutes for the analysis of this paper. Admittedly, the real autocorrelation time is likely to be considerably longer, perhaps on the order of a few hours (Everitt and DeBra , this volume). However, the white noise model is probably a better representation of the realistic error of the experiment than is a worst case error analysis based on long autocorrelation times, particularly in view of the likelihood of an adaptive sequential filter for the analysis of real data. In the future, as the spacecraft systems become better defined, and

as their noise characteristics become known, it will be worthwhile to perform a more realistic (and much more costly) error analysis with colored acceleration noise. For now, the white noise model is sufficient for preliminary design purposes. We think it very unlikely that the autocorrelation function of the drag compensation system will be equal to the autocorrelation in the Doppler data introduced by the sought after J_2 or relativity signals.

IV. ERROR ANALYSIS

A total of 16 unknown or dependent parameters have been included in the error analysis. These are:

1. Six heliocentric cartesian coordinates $(x_p, y_p, z_p, \dot{x}_p, \dot{y}_p, \dot{z}_p)$ of the probe at epoch (encounter minus one month).

2. Six heliocentric cartesian coordinates $(x_E, y_E, z_E, \dot{x}_E, \dot{y}_E, \dot{z}_E)$ of the earth at epoch

3. The astronomical unit (AU) or equivalently GM for the sun ($GM_s \propto AU^3$).

4. The quadrupole coefficient J_2.

5. Two relativity parameters (β,γ).

Partial derivatives of the range and range rate of the probe with respect to these 16 parameters have been obtained by first numerically integrating the second order ordinary differential equations of the variations of the probe and earth accelerations, and then by computing the range and range rate derivatives by the geometrical definitions of those quantities. Range ρ is defined by:

$$\rho = \left[(\vec{r}_p - \vec{r}_E) \cdot (\vec{r}_p - \vec{r}_E) \right]^{\frac{1}{2}} \tag{5}$$

so that

$$\frac{\partial \rho}{\partial q} = \frac{(\vec{r}_p - \vec{r}_E)}{|\vec{r}_p - \vec{r}_E|} \cdot \left(\frac{\partial \vec{r}_\rho}{\partial q} - \frac{\partial \vec{r}_E}{\partial q} \right) \tag{6}$$

where q is any one of the sixteen parameters.

The range rate partial is obtained by differencing two range partials,

a technique which yields a realistic representation of integrated Doppler.

$$\frac{\partial \dot{\rho}}{\partial q} \equiv \left(\frac{\partial \rho_i}{\partial q} - \frac{\partial \rho_{i-1}}{\partial q} \right) / \left(t_i - t_{i-1} \right) \tag{7}$$

In the case of the parameter γ, a signal propagation effect is added to the range partial to account for the general relativistic time delay (Shapiro, 1964). This additional term is:

$$\frac{\partial \rho}{\partial \gamma} = \frac{GM_s}{c^2} \ln \frac{r_p + r_E + \rho}{r_p + r_E - \rho} \tag{8}$$

Hence, the general relativistic time delay is also included in the range rate partial because of the differencing by Eq. 7.

The units of length and time in the computer program (km, sec) are used here to quote results. Actually, the unit of length in radio propagation experiments is time ($c \equiv 1$), but it is convenient to express results in familiar metric units. Errors in conversion between km and light-seconds are inconsequential to the two or three significant digits of concern.

Baseline case for error analysis.

For purposes of carrying out parametric studies of the sensitivity of the covariance matrix on the sixteen parameters to variations in the assumptions on tracking conditions and spacecraft systems configurations, it is useful to define a baseline case for the error analysis, and then to vary independent variables one at a time from their baseline values. The baseline case is defined in the following table.

BASELINE CASE

Independent Variable	Value
1. Doppler error at 10 min sample	0.1 mm/sec
2. Range error at 20 min sample	15 m
3. Length of tracking arc	E-1 mo to E+1 mo
4. Standard deviation σ_a of acceleration noise	10^{-10} g = 10^{-9} m/sec^2
5. Perihelion distance at solar encounter	4 $R_\odot$
6. Inclination of probe orbit to ecliptic	90°
7. Angle between sun-probe line and sun-earth line at encounter	45°
8. Argument of probe perihelion with respect to ecliptic	0°

The standard deviation for each of the 16 dependent variables for the baseline case are given in the following table. The column labeled a-priori error represents the assumed error on each parameter before the introduction of the first simulated measurement. The a-posteriori error refers to the computed standard deviation after the sequential processing of all simulated measurements.

RESULTS FOR BASELINE CASE

Parameter	A-Priori Error	A Posteriori Error
x_E	1.5 km	0.4 km
y_E	1.5 km	0.7 km
z_E	1.5 km	1.5 km
$\dot{x}_E$	0.1 mm/s	0.09 mm/s
$\dot{y}_E$	0.1 mm/s	0.08 mm/s
$\dot{z}_E$	o.1 mm/s	0.09 mm/s
x_p	10^6 km	0.4 km
y_p	10^6 km	1.1 km
z_p	10^6 km	3.2 km
$\dot{x}_p$	10^3 km/s	0.2 mm/s
$\dot{y}_p$	10^3 km/s	0.1 mm/s
$\dot{z}_p$	10^3 km/s	0.8 mm/s
AU	1.0 km	0.2 km
β	0.01	0.0088
γ	0.01	0.0047
$J_2 \times 10^8$	10^3	0.5

The parameter J_2 can be determined to better than one part in 10^8 for the baseline case. The results on the relativity parameters are disappointing, until one considers the full 2 x 2 covariance on β and γ. That submatrix of the 16 x 16 covariance matrix is:

$$\begin{array}{c} \\ \beta \\ \gamma \end{array} \begin{array}{c} \begin{array}{cc} \beta & \gamma \end{array} \\ \begin{pmatrix} 7.829 & 4.052 \\ 4.052 & 2.200 \end{pmatrix} \end{array}$$

which yields a correlation between β and γ of 0.976. The eigenvalues of the matrix yield independent errors of 0.01 and 0.0009 with associated eigenvectors which are nearly $\beta + \gamma/2$ and $\beta - 2\gamma$. Thus, one relativity parameter

can be determined to a significant accuracy, and this parameter is a linear combination of β and γ in the ratio $\beta - 2\gamma$. The physical interpretation is that the relativistic precession of the perihelion can be determined from the data, but not both β and γ. The post Newtonian precession $\dot{\omega}_{PN}$ is given in terms of the Einstein precession $\dot{\omega}_{GR}$ by (Robertson, 1962),

$$\dot{\omega}_{PN} = \frac{1}{3}(2 - \beta + 2\gamma)\dot{\omega}_{GR}$$

and the error on the precession parameter $(2 - \beta + 2\gamma)/3$ as determined from the covariance matrix on β and γ is 0.0007.

Variation of assumptions from baseline case.

A limited number of variations in the baseline assumptions are considered here. First of all, the standard deviation on the acceleration noise is allowed to vary between 10^{-7}g and 10^{-12}g. The results for J_2 are shown in Fig. 1 and the results for the relativistic precession parameter $(2 - \beta + 2\gamma)/3$ in Fig. 2. Taken at face value, it would seem that an acceleration noise of 10^{-9}g would yield a meaningful measure of J_2, though the degradation in the relativity test is more serious. However, it is unlikely that the error curves are calibrated with enough precision to draw firm conclusions from them. At this point it is prudent to use 10^{-10}g as a design goal for the drag compensation system, particularly if post Newtonian terms are a scientific goal for Solar Probe. In any event, an acceleration noise at 10^{-8}g is insufficient for the gravitational experiments at encounter. On the other hand, improvements beyond 10^{-10}g are difficult to justify unless the Doppler error can be reduced below 0.1 mm/sec at the same time.

The dependence of the error in J_2 to various assumptions on Doppler accuracy is shown in Fig. 3. The dotted portion of the curve indicates extrapolation into a region where computer simulations were not actually made. However, it is the region of the curve between 1 mm/s and 0.1 mm/s that is most important. The error in J_2 rises sharply with decreased tracking accuracy. A telecommunication system capable of delivering a Doppler accuracy of 0.1 mm/s at solar encounter is essential.

Finally, the inclination of the orbit is varied from 90° to 45°. There is little effect on the determination of J_2 until inclinations of 60° or so are reached as shown in Fig. 4. Again the dotted curve represents an extra-

polation. The conclusion here is that the experiment is not very sensitive to the orbital inclination, but as expected, high inclination orbits are better because there is a variation in the latitude with time, and more of the potential function (Eq. 1) is sampled during the encounter tracking.

References

Anderson, J. D., Colombo, G., Friedman, L. D., and Lau, E. L., An Arrow to the Sun, in Proceedings of the International Symposium on Experimental Gravitation, Accademia Nazionale Dei Lincei, Roma, 1977.

Armstrong, J. W., Woo, R. and Estabrook, F., paper in preparation.

Colombo, G., Lautman, D. A., and Pattengill, G. H., in Problemi Attuali di Meccanica Teorica e Applicata, Accad. Turin, p. 291, 1975.

Robertson, H. P., Relativity and Cosmology, in Space Age Astronomy, ed. A. J. Deutsch and W. B. Klemperer, Academic Press, New York, 1962.

Shapiro, I. I., Fourth Test of General Relativity, Phys. Rev. Lett., 13, 789-791, 1964.

We wish to thank T. Woo for some important information on the effects of plasma on the tracking signal and for his assessment of the Doppler noise at the solar encounter of the Solar Probe.

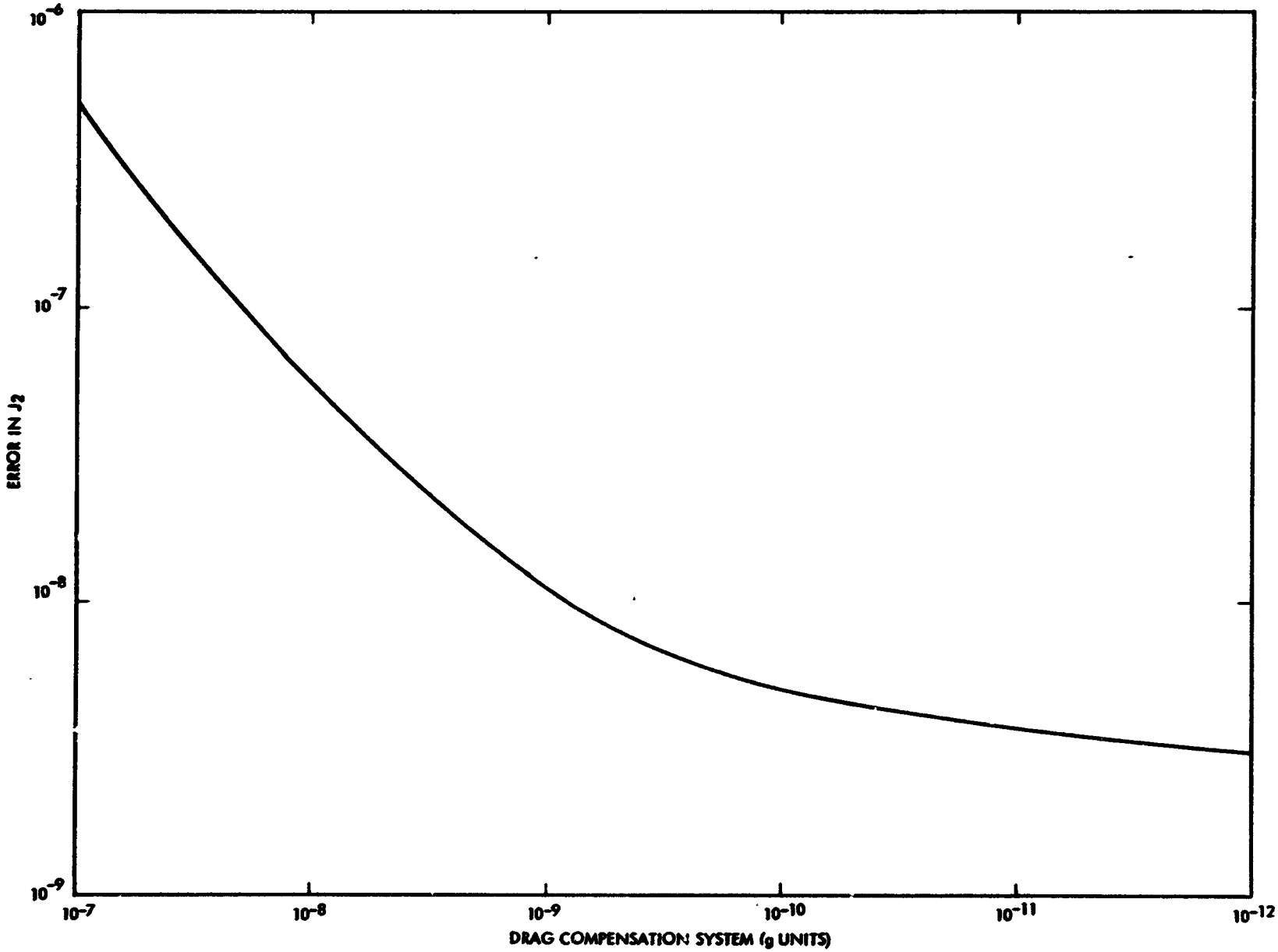


FIGURE 1. Error in J_2 as a Function of Acceleration Noise.

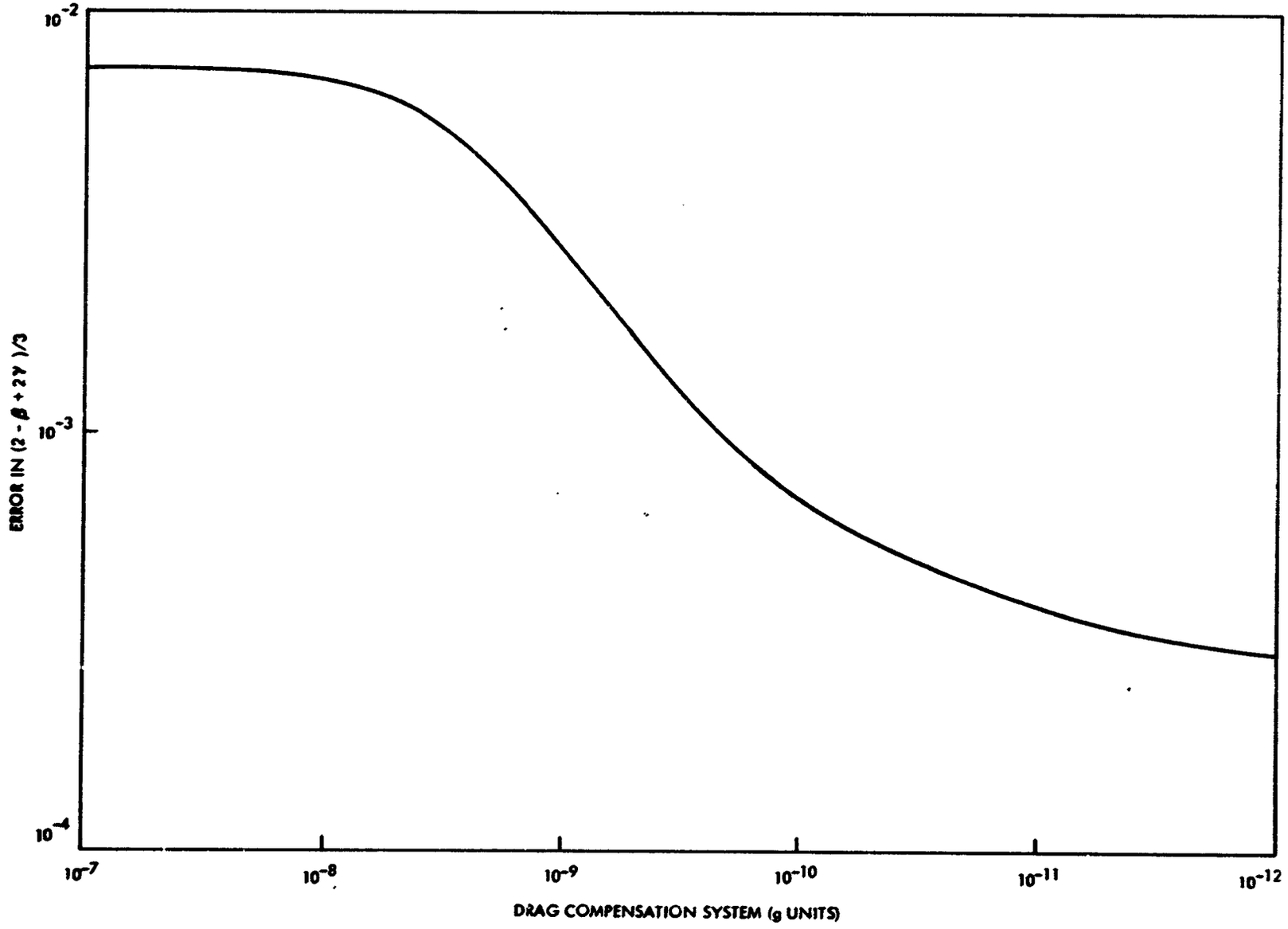


FIGURE 2. Error in Relativistic Perihelion Advance as a Function of Acceleration Noise.

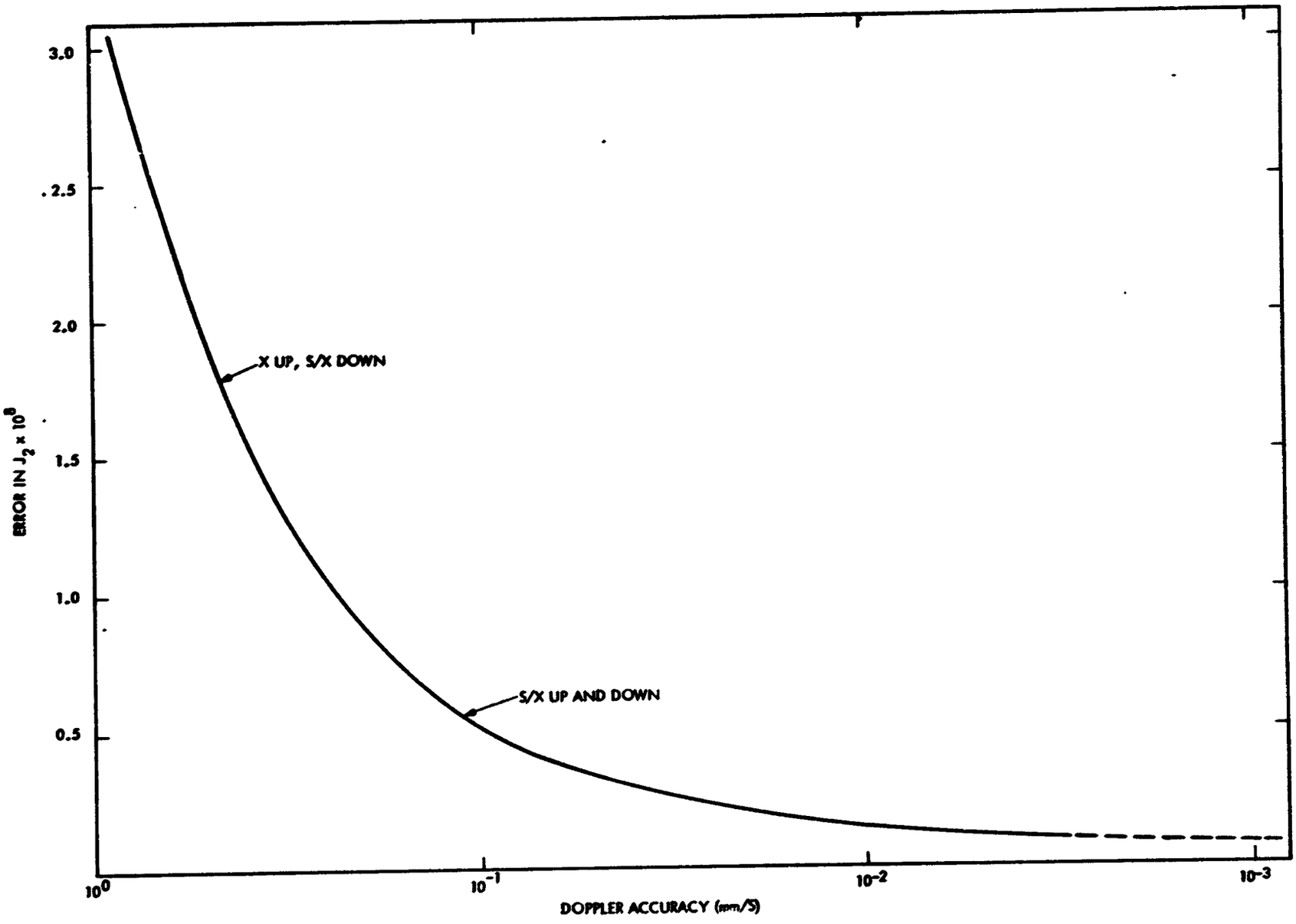


FIGURE 3. Error in J_2 as a Function of Doppler Error.

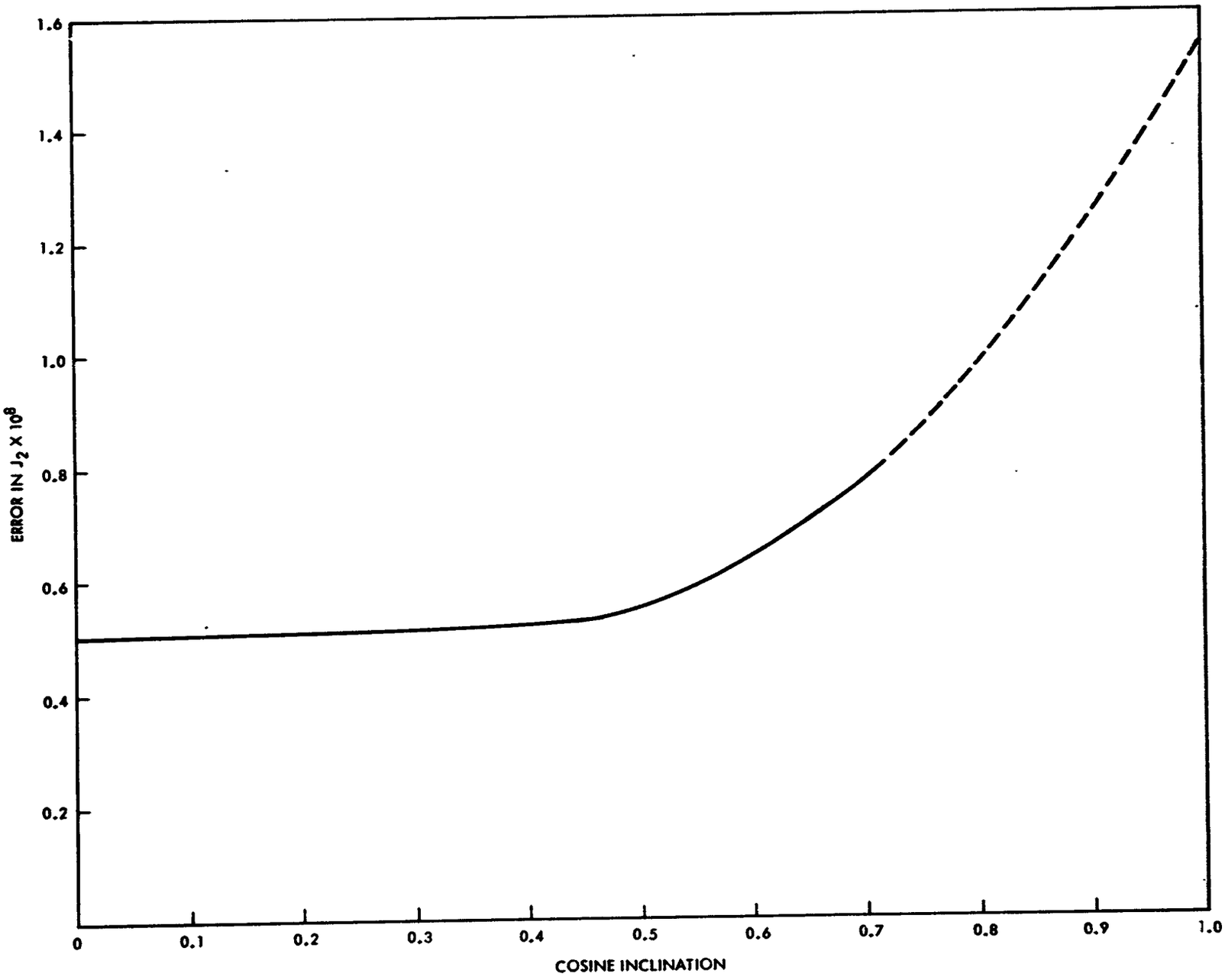


FIGURE 4. Error in J_2 as a Function of Orbital Inclination.

POSSIBLE MEASUREMENTS OF J_2 AND THE SUN'S ANGULAR MOMENTUM WITH THE SOLAR PROBE

K. Nordtvedt*

This paper presents results of a covariance analysis done at Montana State University on the detectability of the solar quadrupole moment and the solar angular momentum using Doppler tracking of a Solar Probe. It is assumed that the dynamical effects of the sun on the probe, as well as the gravitational redshift of a probe-based clock, may be totally described by the following metric tensor:

$$g_{oo} = 1 - 2\frac{m}{r} - 2J_2\frac{ma^2}{r^3} + 2\beta\frac{m^2}{r^2}$$

$$g_{ok} = -\frac{2}{r^3}\,\varepsilon_{kmn}\,r^m\,J^n$$

$$g_{mn} = -\delta_{mn}\,(1 + 2\gamma\frac{m}{r}),$$

where m is the sun's gravitational mass, r^m is the heliocentric position vector with magnitude r, a is the solar radius, J_2 is the solar quadrupole moment, and J^n is the solar angular momentum vector. The parameters β, γ, J_2 and $\left|J^n\right|$ are considered unknown, along with the orbital elements of the probe. All of these parameters are to be determined by a least squares fit to the probe tracking data.

For purposes of this study, a polar orbit was assumed with perihelion at 5 solar radii and with an earth-sun-probe angle of 135° at perihelion. Tracking was assumed to be Doppler only, with a basic uncertainty of .1 mm/sec in the radial velocity. A drag-free system which could reduce non-gravitational forces below the level of 2×10^{-12} g was found to be necessary to preserve this same level of accuracy. Both one-way and two-way Doppler were supposed in order to determine the gravitational redshift as well as the pure spacecraft dynamics. A worst case estimate was performed, in which it was assumed that systematic error at the .1 mm/sec

level would bias the data in such a way as to most inhibit the determination of each parameter separately. It can be shown that this assumption is equivalent to multiplying the final prediction of accuracy in the parameter estimate by $\sqrt{N}$, where N is the number of observations made.

Tracking times of 42 hours, 97 hours, and 136 hours were investigated leading to the following results:

A. Two-way data.

1. With J_2 and $\vec{J}$ as the only metric unknowns, accuracies of a part in 10^8 for J_2 and a fraction of the expected value for $\vec{J}$ were predicted.
2. When γ was added as a third unkown, the accuracies on J_2 and $\vec{J}$ were degraded somewhat while γ was determinable to a part in 10^3.

B. One-way data.

1. Predictions for the accuracy in the gravitational redshift were consistently about a part in 10^7 as opposed to parts in 10^2 for laboratory experiments and 10^4 for earth orbiting rocket experiments.
2. A prediction of marginal detectability of the second order redshift was also made.

In all these cases, assumptions of no data within various angular distances of the sun were made. The results did not vary appreciably.

The conclusions are that significant determinations of the solar quadrupole moment and of the gravitational redshift are possible with the Solar Probe, and that any measurement at the levels predicted must also allow for a measurement of $\vec{J}$. In fact, the effect of $\vec{J}$ would have to be modeled in order to produce the high accuracy measurements of J_2 and redshift.

*R. Hellings prepared this version of K. Nordtvedt's presentation from a tape of same.

N78

32969

UNCLAS

COMMENTS ON THE DRAG-FREE CONTROL OF A SOLAR PROBE RELATIVITY MISSION

C. W. F. Everitt and D. B. DeBra
W. W. Hansen Laboratories of Physics and
Department of Aeronautics and Astronautics
Stanford University, Stanford, California

A spacecraft going within 4 solar radii of the Sun experiences an acceleration up to 5×10^{-5} g from solar radiation pressure, and significant (though smaller) accelerations from the solar wind. To obtain satisfactory information about relativistic effects and the Sun's quadrupole mass coefficient J_2, these non-gravitational accelerations have to be either measured or compensated out, and much the best procedure for doing so is to make the spacecraft "drag free". A proof mass inside the spacecraft structure is shielded from the external forces, so that it follows a nearly ideal gravitational orbit, and a control system activates gas jets (or other translational forcers) to make the vehicle follow the mass. The problems are mechanizing the control laws and minimizing extraneous effects such as the self gravitational pull of the spacecraft. The extraneous forces can be averaged in one plane by having a spinning vehicle.

The lectures by R. D. Reasenberg, J. D. Anderson and K. Nordtvedt provided three complementary analyses of a Solar Probe relativity mission. All three authors emphasized that their analyses, though well advanced, depended on assumptions about drag-free performance. Reasenberg and Shapiro have concentrated so far on tracking requirements for the mission, assuming an ideally drag-free spacecraft. Anderson and Lau have assumed a drag-free system with no steady in-track bias, but with random noise in the residual acceleration treated as a parameter in experiment design. Nordtvedt investigated the effect of a steady in-track bias acceleration with negligible random noise.

Nordtvedt concluded that in order to measure the Sun's quadrupole mass coefficient J_2 to 10^{-8}, and calculate the Sun's angular momentum $\vec{J}$ to 20-30% from the Lense-Thirring drag, the in-track bias of the drag-free system should be below 2×10^{-12} g. Anderson's conclusion

was that random noise will not degrade the relativistic information obtainable from tracking measurements until it reaches a 1σ level of a few times 10^{-10} g. Both authors investigated the implications of their analyses for determining the combination of PPN relativistic parameters $(2 + 2\gamma - \beta)$, as also did Reasenberg and Shapiro. The precision of measuring the gravitational redshift in a Solar Probe mission was also investigated.

Reasenberg in the introduction to his lecture correctly pointed out that factor in drag-free performance ultimately limiting the relativity mission is neither the random noise nor the steady in-track bias, both of which can in some degree be separated from relativity terms in data processing, but rather the component of variation in the in-track acceleration having a time signature comparable with that of the relativistic terms, that is, the change in drag over the 19 hour duration of the fly-by of the Sun. In other words what counts are the components in the power spectral density of drag-free noise with frequencies comparable to the relativity or J_2 signals.

In this note we investigate the likely magnitude of the 19 hour period terms for spinning and non-spinning spacecraft. Our conclusion is that the variation in along-track acceleration for a non-spinning spacecraft from thermal warping of the spacecraft is likely to be of order 10^{-11} g and that electrical charging of the proof mass may easily give accelerations in excess of 10^{-10} g. The variation in along-track acceleration can be made significantly less in a spinning spacecraft, but the effectiveness of the reduction is somewhat of a subtle point which requires detailed investigation. We are not in a position to estimate the implication of our conclusion for the analyses of Anderson, Nordtvedt and Reasenberg; presumably there is some loss in accuracy of the relativity data.

Our starting point is the performance of the DISCOS drag-free controller for the TRIAD I satellite launched in July 1972 and operated successfully in space until the propellant was used up in November 1975.[1] The design was for a satellite free from extraneous disturbances down to the level of 10^{-11} g. This goal

was reached: in fact the measured in-flight performance after correcting for initial errors was 5×10^{-12} g.

At 10^{-11} g mass attraction effects on the drag-free proof mass are of dominant importance. To separate the proof mass as far as possible from most of the satellite mass, the satellite was formed in three bodies (a TRIAD) joined by extensible booms, with the DISCOS (DISturbance COmpensation System) as a separate central package. The top body contained the nuclear power supply for the satellite; the bottom one the transmitter, receiver, telemetry and antenna. The satellite weighed 86.6 kg. It was launched in a collapsed form; roughly a cylinder 1.6 m high, 0.75 m in diameter. In orbit both booms were extended 2.7 m.

The proof mass was 22 mm in diameter, placed in a 40 mm diameter cavity. It was made of a 70/30 gold-platinum alloy chosen for its high density and nearly zero magnetic susceptibility. The cavity contained three mutually perpendicular pairs of electrodes to sense the position of the proof mass by three capacitance jets: the three signals were used in a control loop to fire three corresponding pairs of gas jets. In orbit the proof mass floated in the cavity with a 9 mm clearance. It was permitted to move over a dead band of $\pm$ 0.9 mm before the valves were fired, which meant establishing a specification of 10^{-11} g/mm for the gradient of the disturbing forces. Note that for a Solar Probe subject to an acceleration of 5×10^{-5} g the valves would have to fire once every 2 sec to maintain an 0.9 mm deadband. The maximum number of firings in one axis during the 19 hour fly-by would be about 2×10^{4}, well within the reliability limits of the valve.

Figure 1 illustrates the DISCOS control system designed by the Stanford Guidance and Control Laboratory. All parts of the satellite not essential to the control system were located in the two bodies at the end of the extensible booms. In the final configuration the contribution to mass attraction uncertainty from the two end bodies was restricted to a few percent without unusual fabrication tolerance or accuracy in determining mass properties. Not surprisingly the mass distribution in parts closest to the proof mass required most attention. The beryllium oxide housing

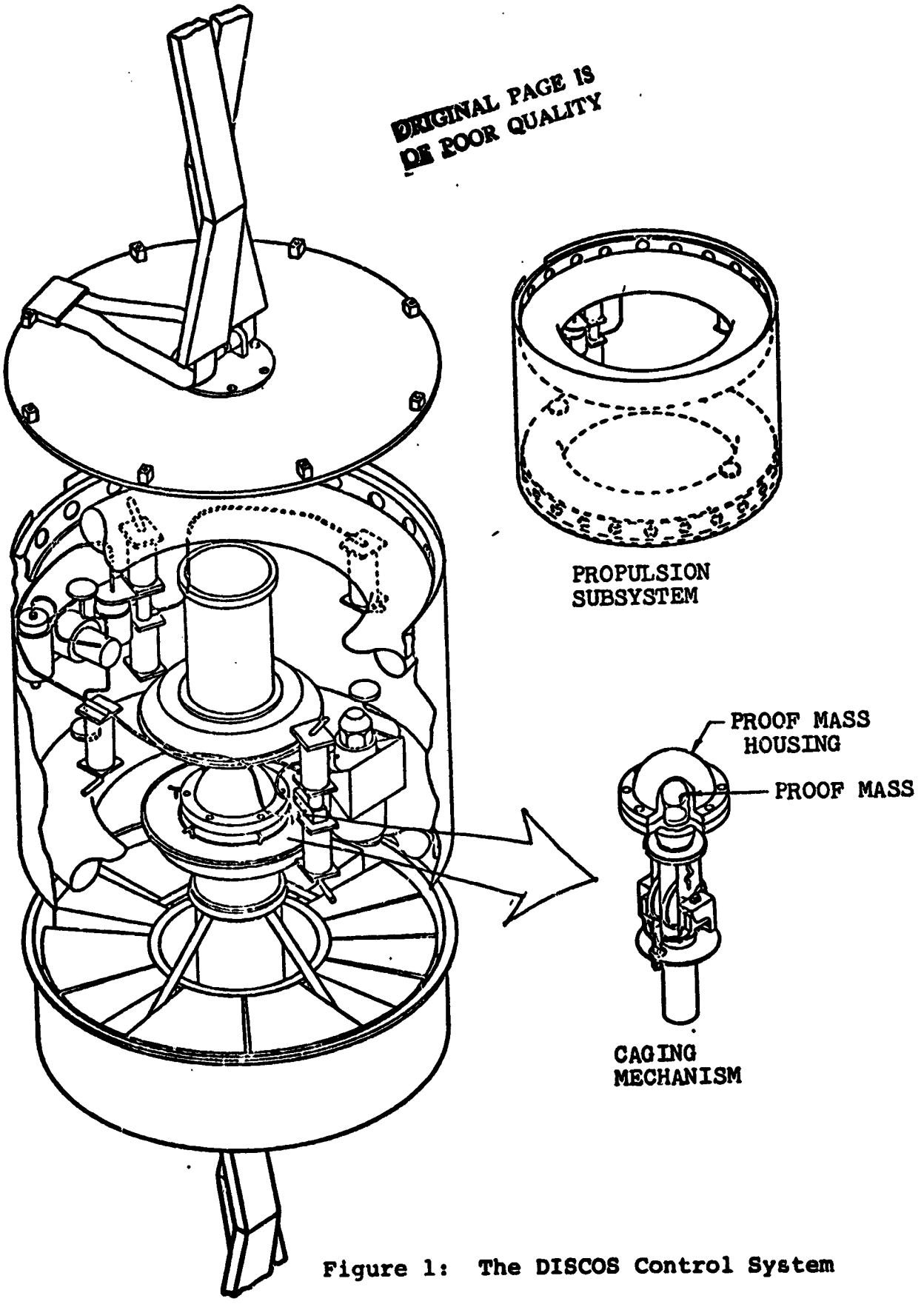


Figure 1: The DISCOS Control System

had to be allotted a share of the error budget five times larger than any other single component. Mass attractions were calculated with great precision. With the regulator, for example, terms up to the third moment of mass were included in an expansion about the regulator mass-center.[2] The first and second moments were found experimentally by measuring the mass and the six elements of moment of inertia. The third moment had to be calculated from drawings of the component parts and weights determined during assembly. Fabrication tolerances in the structure as tight as 0.5 mil were required to model the mass attraction adequately. Similarly fabrication tolerances of 0.1 mil were required in the pickoff housing and in thickness measurement of the propellant tanks. Finally a compensation mass was added immediately above the pickoff housing. The residual attraction before compensation was calculated as 83.6×10^{-11} g in the vertical direction and 0.8×10^{-11} g in the "along track" direction. With the compensating mass the maximum disturbances were calculated to be 2×10^{-11} g normal to the orbit plane and 0.7×10^{-11} g along track.

The propulsion subsystem was a conventional "cold-gas" propulsion system. However to satisfy the mass-attraction requirement, it was necessary to have two toroidal propellant tanks, placed equidistant on opposite sides of the proof mass in a configuration making the gravitational analog of a pair of Helmholtz coils. Without some such arrangement there would have been changes in the mass-center, center of attraction or mass attraction gradient of the propellant as the propellant mass decreased with time. Also the temperature differences across the propellant tanks had to be held below 2 ^{o}C to prevent gas redistribution from compromising the mass attraction specification.

Design considerations for a drag-free solar relativity mission were first discussed in 1970 in an admirable paper by Remy Juillerat.[3] Further analysis, drawing on the DISCOS flight experience, has been given in a Report "Drag-Free Control for a Solar Spacecraft" submitted to the Jet Propulsion Laboratory by D. B. DeBra in August 1976.[4] A general conclusion of DeBra's Report was that drag-free control of a non-spinning Solar Probe would be feasible at a d.c.

bias level of 10^{-10} g, but that levels much below 10^{-10} g would be difficult to reach since the constraints on spacecraft design are more stringent than those for TRIAD I. The statement is slightly obscured by typographical errors of 10^{-11} g for 10^{-10} g in a few places.

DeBra's Report discusses calibrating out the d.c. bias by in-flight checks of the drag-free system while the spacecraft is on the way to or from Jupiter. At Jupiter the radiation pressure from the Sun causes an acceleration on the spacecraft of order 10^{-10} g, so with appropriate tracking and maneuvers of the spacecraft an effective calibration is possible. A tradeoff exists between the expense of drag-free design and the expense of tracking and maneuvers. We should also emphasize that structural creep or other changes in the spacecraft during the interval between the encounters with Jupiter and the Sun may limit the effectiveness of the calibration.

The high frequency variations in disturbances due to control system activity coupling to the proof mass through disturbing force gradients may be largely averaged out. The main problem in a solar relativity mission, as already stated, is neither this nor the d.c. bias, but the magnitudes of any components of along-track acceleration with this signature comparable to that of the relativity signals during the 19 hour fly-by. Such effects might arise from any of the following causes:

(1) changing solar radiation pressure on the spacecraft;

(2) changes in the control gain or bias through changes in temperature of the electronics;

(3) thermal warping of the spacecraft causing a change in its self-gravitational attraction on the proof mass;

(4) changing electric charge on the proof mass.

The effect of solar radiation pressure increases by nearly three orders of magnitude from 10^{-7} g to 5×10^{-5} g in the first 9½ hours of fly-by and then dies away again. A variation in applied acceleration as large as this would cause trouble if the drag-free controller utilized a simple position sensor, owing to the gradients

in the disturbance forces on the proof mass. The displacement in average position of the proof mass needed to provide the control signal would result in a variation in residual radial acceleration of the spacecraft during the flyby, and hence a variation in the along-track acceleration with a 19-hour time signature. However, this term could be greatly reduced by using a drag-free system based on an integral control law, so the variation in external pressure is not a real problem.

With proper design, temperature changes in the electronics are also no problem. The DISCOS electronics boxes were passively controlled to levels at which the 5×10^{-12} g performance was achieved despite quite large swings in satellite temperature. For Solar Probe one could, if necessary, have active temperature controllers.

Warping of the spacecraft is a different story, however. The changes in mass distribution to be expected during fly-by are enough to cause changes of possibly 10^{-11} g in the self-gravitation attraction of a non spinning spacecraft on the proof mass from this cause alone.

Charging of the proof mass may be an even more serious problem. The discussion of this topic in Reference (4) may have been too optimistic. With DISCOS in an 800 km orbit there was no trouble, but charging rates of the French CACTUS accelerometer flown in an elliptical orbit with 1200 km apogee were such that the ball charged at a rate of 2×10^{-11} coulomb/day during periods of exposure to Van Allen belts.[5] The charging had a cyclic behavior associated with regression of the orbit: ten days of continuous charging would be followed by a quiet spell, repeating every 37½ days. The resultant accelerations were as high as 10^{-8} g, or in one instance nearly 10^{-7} g. What the Sun will do we would not presume to say, but large variations in charge during fly-by seem possible. It might be necessary to introduce a discharging mechanism, which would mean a significant increase in sophistication of the control system. Some possible techniques for discharging the ball are described by Juillerat in Reference (3).

Spinning the spacecraft averages the drag in the plane perpendicular to the spin axis probably by as much as two orders of

magnitude. Design and laboratory simulation of drag-free controllers for a spinning vehicle are described in References (6) through (9). However two cautionary remarks must be set down. First, as Juillerat has shown,[10] the thermal time constants of the structure may be such that significant errors occur even with a spinning spacecraft. Second, Solar Probe as presently configured must have one axis directed always to the center of the Sun. To spin about this axis requires great precessional torques, and the spin axis will only be perpendicular to the fly-by path at the instant of closest approach. This is the best place, of course, since it is the region of the orbit where the relativistic effects are largest, but in other regions there will be a component of along-track acceleration from the component of self gravitation of unknown magnitude parallel to the axis of spin. The principal effect will have a time signature with two peaks during the 19 hour period. This signature differs from at least some of the terms to be measured, so something might be done to separate it out in data analysis: how far one can go in this way we do not know.

On a more cheerful note we observe that the noise in a well-designed drag-free system should be appreciably less than the 10^{-10} g requirement determined by Anderson and Lau. The principal tasks for the control engineer therefore are to fight down effects of d.c. bias, thermal warping and proof mass charging. These are challenges enough.

References

(1) The Staff of the Space Department, Johns Hopkins Applied Physics Laboratory, and the Staff of the Guidance and Control Laboratory, Stanford University, "A Satellite Freed of All but Gravitational Forces: TRIAD I," J. Spacecraft & Rockets 11, (1974).

(2) A. W. Fleming and M. G. Tashker, "Final Report on Mass Attraction of TRIAD I/DISCOS" submitted to Johns Hopkins Applied Physics Laboratory under Contract NOW62-0604c, September 1972. Available from Stanford University Department of Aeronautics and Astronautics as SUDAAR Report No. 445.

(3) R. Juillerat, "Physical Limitations in Sensors for a Drag-Free Deep-Space Probe" in Proceedings of the Conference on Experimental Tests of Gravitation Theories, November 11-13, 1970, JPL Technical Memorandum 33-499, ed. R. W. Davies, pp. 197-212.

(4) D. B. DeBra, "Drag-Free Control for a Solar Spacecraft" Report submitted to Mr. Edward Mettler, Jet Propulsion Laboratory, PO CC-665329, August 1976.

(5) J. Bouttes, M. Delattre et R. Juillerat, "Qualification in Orbital Flight at the Cactus High Sensitivity Accelerometer" ONERA TP No. 1976-66E.

(6) J. D. Powell, "Control of a Spinning Drag-Free Satellite with an Application of Estimator Theory," Ph.D. dissertation Stanford University, Department of Aeronautics and Astronautics, May 1970 SUDAAR No. 402.

(7) M. G. Tashker, "Integral Control of a Spinning Drag-Free Spacecraft," Ph.D. dissertation Stanford University, Department of Aeronautics and Astronautics, April 1974 SUDAAR No. 472.

(8) S. Sanz Fernandez de Cordoba, "Orientation and Three Dimensional Mass-Center Estimation in a Rotating Drag-Free Spacecraft," Ph.D. dissertation, Stanford University, Department of Aeronautics and Astronautics, August 1975 SUDDAAR No. 496.

(9) R. L. DeHoff, "Minimum Thruster Control of a Spinning Drag-Free Satellite, including Design of a Large Cavity Sensor," Ph.D. dissertation, Stanford University, Department of Aeronautics and Astronautics, December 1975 SUDAAR No. 497.

(10) Reference (3) p. 200.

MEASUREMENT OF SOLAR GRAVITATIONAL OBLATENESS WITH GRAVITY GRADIOMETERS

Robert L. Forward

Hughes Research Laboratories
Malibu, California 90265

ABSTRACT

We propose to measure the gravitational oblateness of the Sun during a Solar Probe Mission by means of a very high sensitivity rotating gravity gradiometer onboard the spacecraft. The gravity gradiometer would be a self-contained structure in a thermal-vacuum-magnetic shield in the shape of a cross, with a mass of 20 kg and arm length of 1.0 m and thickness of 0.1 m. The sensor inside would have a resonant frequency of 1/30 Hz, a mechanical Q of 10^6 and would use gravitational radiation antenna technology to achieve a sensitivity of 6×10^{-8} Eotvos, which would provide a measurement of the solar oblateness to an accuracy of $1.5\text{-}6 \times 10^{-8}$. The gravity gradiometer will require a spinning spacecraft, so that it will not sense the spacecraft gravity fields, but the gradiometer does not need to be at the spacecraft center of mass, or even on the spacecraft spin axis. Major problem areas to be addressed are demonstration of the instrument sensitivity prior to flight (a shuttle test flight would be suitable), and the measurement and compensation for any residual spacecraft angular rates.

INTRODUCTION

A Solar Probe mission is part of the 5-year plan of NASA's Solar Terrestrial Office. The trajectory of the Solar Probe is a highly elliptical polar orbit with a perihelion near 4 solar radii (see Figure 1). At 4 solar radii the gradients in the gravitational field of the sun are substantial, and it is possible to consider the measurement of the oblateness of the gravitational field using onboard gravity gradiometer instruments.

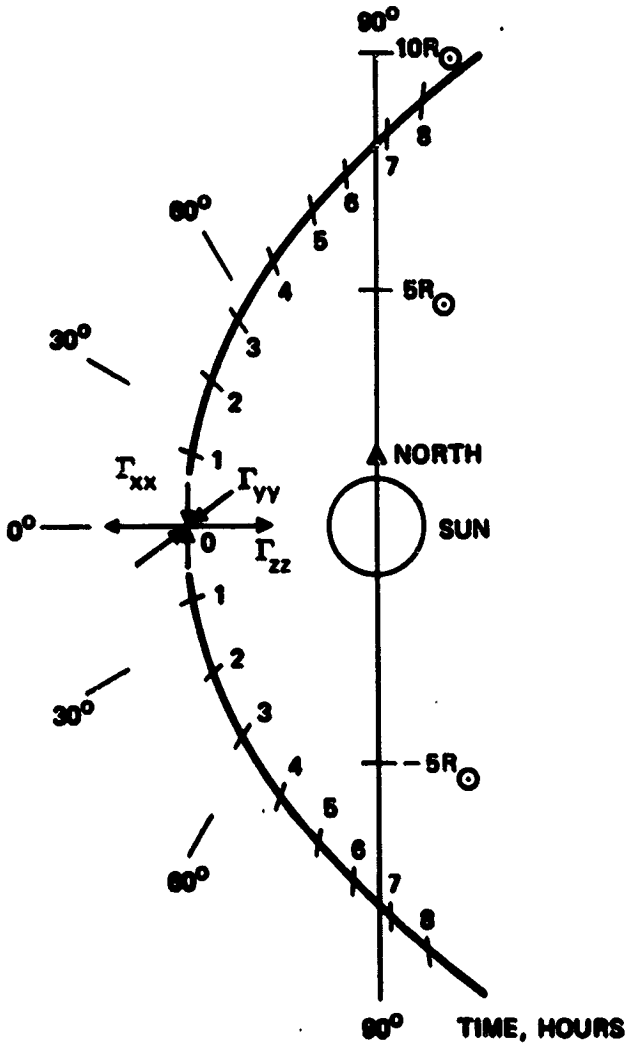


Figure 1. Gravity gradients for near-sun trajectory 4 $R_{\odot}$ perihelion.

GRAVITY GRADIENTS OF SOLAR OBLATENESS

We assume that the gravitational potential of the sun has the form

$$U = \frac{GM}{R}\left[1 - J_2 P_2 (\cos \psi)\left(\frac{A}{R}\right)^2 + \text{higher order terms}\right]$$

where J_2 is the solar oblateness coefficient, M and A are the solar mass and radius, and R is the distance from the gradiometer to the center of the sun. At the point of closest approach of the solar probe to the sun, there will be three principal gradients at the position of the probe. There will be the gradient in the polar direction

$$\Gamma_{xx} = \frac{GM}{R^3}\left[-1 + \frac{9}{2} J_2 \left(\frac{A}{R}\right)^2\right]$$

the gradient along the equator

$$\Gamma_{yy} = \frac{GM}{R^3}\left[-1 + \frac{3}{2} J_2 \left(\frac{A}{R}\right)^2\right]$$

and the vertical (radial) gradient

$$\Gamma_{zz} = \frac{GM}{R^3}\left[2 - 6 J_2 \left(\frac{A}{R}\right)^2\right] .$$

The rotating gravity gradiometer measures the differences in the two gradients in its plane of rotation. Therefore, the gradiometer output depends on the direction of its spin axis. The output for a vertical (solar oriented) spin axis is

$$\Gamma_{xx} - \Gamma_{yy} = \frac{GM}{R^3}\left[0 + 3 J_2 \left(\frac{A}{R}\right)^2\right] .$$

For a spin axis along the polar direction (along the trajectory for a polar orbit) the output is:

$$\Gamma_{yy} - \Gamma_{zz} = \frac{GM}{R^3}\left[-3 + 7.5 J_2 \left(\frac{A}{R}\right)^2\right]$$

and for a spin axis orthogonal to the orbital plane (spacecraft rolling along its trajectory) the output is:

$$\Gamma_{zz} - \Gamma_{xx} = \frac{GM}{R^3}\left[3 - 10.5\ J_2 \left(\frac{A}{R}\right)^2\right] \quad .$$

For the assumed parameters of the Solar Probe mission

$$GM = 1.3 \times 10^{20}\ m^3/sec^2$$

$$A = 7 \times 10^8\ m$$

$$R = 4A = 2.8 \times 10^9\ m$$

we see that the gradients of the sun monopole moment will be of the order of

$$\Gamma = \frac{GM}{R^3} \approx 6 \times 10^{-9}\ sec^{-2} = 6\ E \quad .$$

$$(1\ E = 1\ Eotvos = 10^{-9}\ sec^{-2} \approx 10^{-10}\ G/m) \quad .$$

The contribution due to the J_2 solar oblateness term depends on the orientation of the gradiometer and the type of output and can vary from

$$\Gamma_{J_2} = (3 \rightarrow 10.5)\ J_2 \left(\frac{A}{R}\right)^2 \frac{GM}{R^3}$$

or

$$= (1.1 \rightarrow 4)\ J_2\ (Eotvos) \quad .$$

Since the estimates of J_2 range from $\sim 10^{-8}$ to $\sim 10^{-5}$, it would be desirable to measure J_2 to an accuracy of the order of 10^{-8}.

Since the only operating gravity gradiometer instruments are prototype airborne survey models with a sensitivity of 1 E at integration times of 10 sec, it would seem that the goal for a solar oblateness measurement is too far from present technology. However, the Solar Probe mission requirements are significantly different from the airborne survey mission requirements and it is possible to consider gradiometers for the mission. The free fall of space provides a benign acceleration environment for an instrument, allowing the design to increase in length and mass to increase its capture of gravitational energy. The length of the 6 hr close encounter combined with the fact that all the data collected during the encounter can be used to obtain a single best fit to determine the desired parameter J_2 will increase the accuracy with which we can measure J_2. Also, recent advances in our understanding and operation of highly sensitive detectors of gravitation, such as resonant gravity gradiometers and gravitational radiation antennas, gives us reason to believe that the desired sensitivities can be achieved.

GRAVITY GRADIOMETER INSTRUMENTATION

The sensing portion of the rotating gravity gradiometer consists of two orthogonal arms with weights at the ends. These arms are coupled to each other and to an outer rotor case by stiff torsional pivots plus a pair of piezoelectric bender transducers. The inertia of the arms and the spring constant of the pivots and transducers are adjusted to provide a high-Q, mechanically resonant structure. Present models are resonant at 35 Hz,[1] but larger models resonant at lower frequencies have been constructed.[2,3] When the rotor case is closed, the resultant mechanical structure is very rugged and can withstand multi-gee shocks without requiring caging. The rotor is usually placed on self-pumping oil-lubricated journal bearings and spun inside a cylindrical stator at exactly one-half the sensor resonant frequency. However, if the sensor is attached to a spinning spacecraft, the spacecraft can provide the desired rotation.[2-4] The gravity gradient field of the mass being sensed produces a difference in acceleration across the arms of the sensor and generates a differential torque between the two arms that is

proportional to the gravity gradient (see Figure 2). This differential torque is modulated by the rotation of the sensor to produce a sinusoidal driving torque at twice the sensor spin speed or exactly at the sensor mechanical resonant frequency. Because of the resonant response, the structure will provide a large amplitude response to even very small gravity acceleration differences. These oscillations of the sensor arms are then detected by the piezoelectric bender transducers coupled between the two arms. This is the same transducer technique used in detectors for gravitational radiation.

The sinusoidal signal from the transducer is amplified, filtered, and converted into an FM signal by internal electronics. The resultant data consists of a measurement of the amplitude and direction of the difference between the two components of the gravity gradient field in the plane of rotation of the sensor.

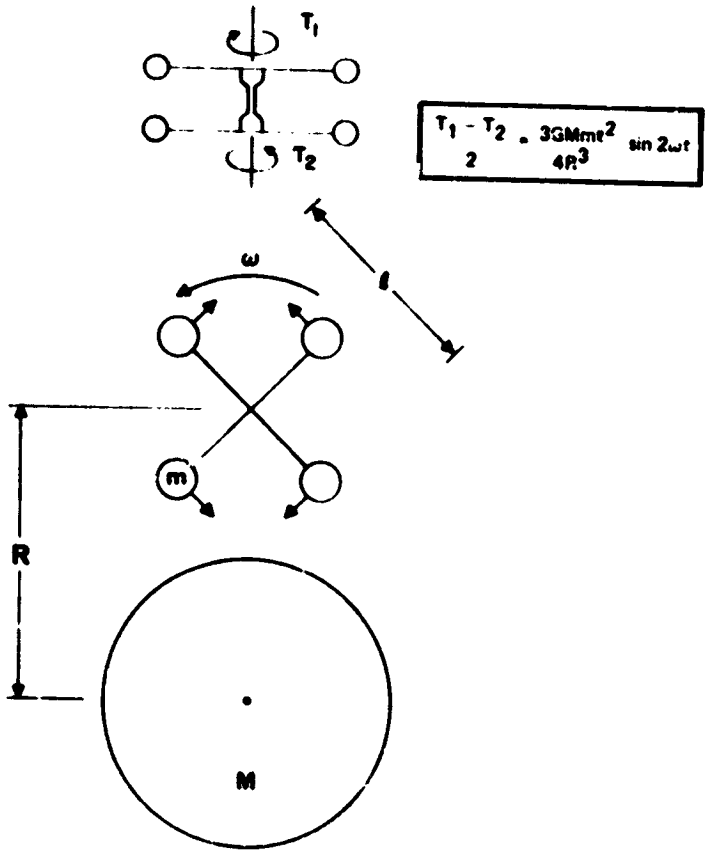


Figure 2. Method of operation of rotating gravity gradiometer.

GRADIOMETER SENSITIVITY

The ultimate sensitivity of a resonant gravity gradiometer is given by

$$\Gamma_{noise} = \left(\frac{2kT}{t_m \, t_s \, m\ell^2} \right)^{1/2}$$

where kT is the thermal noise introduced into the structure by the inherent mechanical damping, $m\ell^2$ is the effective inertia of the sensor arms, t_m is the measurement time, and t_s is the 1/e decay time of the sensor as determined by the level of mechanical damping.

We will assume that every effort will be made to make the initial mechanical damping in the sensor as small as possible. Since a mechanical Q of 2×10^4 has been seen in gradiometers where initial damping was not of concern, we expect to be able to obtain initial mechanical Q's of 10^6 or better. With a Q of 10^6 and a frequency of f = 1/30 Hz, the decay time of the instrument would be

$$t_s = \frac{2Q}{\omega} = \frac{Q}{\pi f} = 10^7 \text{ sec}$$

or approximately 1/3 of a year.

The decay time of the sensor t_s is also the time that it takes the state of excitation of the sensor to change by an amount kT due to random fluctuations introduced by the exchange of energy with the rest of the thermodynamic environment through the damping. This means that during the 6 hr (2×10^4 sec) measurement time duration of the encounter, the change of excitation of the instrument would be almost totally that due to the driving forces of the gravity gradient field acting on the sensor arms, while only a small fraction is random excitation due to thermal energy slowly leaking in through the damping mechanisms.

For a sensor mass of m = 10 kg, length ℓ = 1 m, sensor time constant $t_s = 10^7$ sec, measurement time $t_m = 2 \times 10^4$ sec, and temperature of T = 300°K, the equivalent noise level is

$$\Gamma_{noise} = \left(\frac{2kT}{t_s \, t_m \, m\ell^2} \right)^{1/2} = 6 \times 10^{-17} \text{ sec}^{-2} = 6 \times 10^{-8} \text{ E}$$

where $k = 1.38 \times 10^{-23}$ J/°K. This would allow for a measurement of J_2 to an accuracy of $1.5\text{-}6 \times 10^{-8}$. For comparison, we have constructed[3] a brassboard prototype of a spinning spacecraft gravity gradiometer with $m = 24$ kg, $\ell = 0.75$ m, $f = 8$ Hz, $Q = 900$, and both τ_s and $\tau_m = 35$ sec (see Figure 3). This instrument had a design sensitivity of 7×10^{-3} E.

It will not be easy to build a gravity gradiometer with the required sensitivity, but it should be considerably cheaper than a drag-free spacecraft design. In addition, the instrument package can be checked out in the laboratory and further tested in earth orbit prior to the Solar Probe mission.

As presently envisaged, the gradiometer instrument will be in the shape of a self-contained X shaped thermal-vacuum jacket with arms from 0.50 to 2 m in length, rigidly attached at any convenient place on the spacecraft (it does not have to be at the center of mass and does not

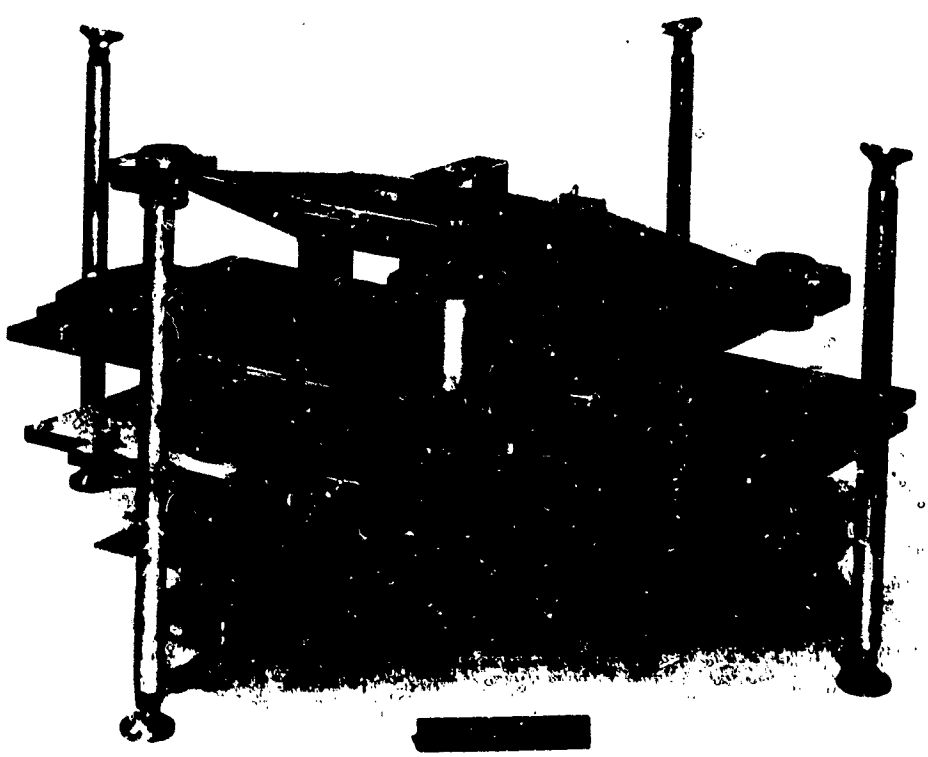


Figure 3. Brassboard prototype of 0.75 m diameter sensor.

have to be on the spacecraft rotation axis). For optimum operation, the sensor should be attached firmly to the spacecraft and the whole spacecraft rotated at 1 rpm (approximately 1/2 the sensor resonant frequency).[4] In this mode of operation, the gravity gradiometer does not see the gravity gradient fields of the spacecraft (not even the gravity changes of a moving platform or fuel motion) only the gravity gradient fields of the sun.

ERROR SENSITIVITIES

Although one would think that the 1/3 year decay time would make the response time of the instrument too long to use in a 6 hr encounter, this is not a real problem. The amplitude of the sensor response is merely measured periodically and the change in the response is the amount of excitation inserted by the gravity gradient forces since the previous measurement of the state of excitation of the sensor. The only problems that arise from the use of this technique come when the state of excitation becomes so high that nonlinearities in the detection system start to become significant. There are a number of techniques to cope with this problem. One is to use electronic "clamping" to stop any residual motion of the sensor prior to the encounter or whenever the excitation level of the sensor becomes too high.

The preferred technique is to use electronic feedback damping using "electrically cold" circuits to lower the response time without adding an equivalent amount of noise (we have demonstrated this on a resonant gravity gradiometer).[5] The electronic damping technique will produce a low-noise sensor with an appreciable bandwidth. This means that less care need be taken to match the satellite spin speed to one-half the sensor resonant frequency.

The major source of error in the experiment will be angular rates of the spacecraft at right angles to the spacecraft spin axis. An angular rate has a rotational rate gradient that is indistinguishable from a gravity gradient. The value of the rotational rate gradient is zero along the axis of the angular rate and is $\Gamma_{\Omega} = \Omega^2$ in the two directions orthogonal to the rate axis.

To achieve a measurement of the gravity gradient to an accuracy of 6×10^{-8} E will require the control or measurement of the spacecraft angular rates (except about the spin axis) to

$$\Omega = 3 \times 10^{-9} \frac{\text{rad}}{\text{sec}}$$

$$= 0.001 \text{ °/hr}$$

To measure this would require a high quality rate gyro.

If a gravity gradiometer was to be considered for the Solar Probe mission, it would be essential that R&D work on a high-Q, low-frequency prototype model be started at once. The desired Q of 10^6 should not be hard to obtain, but the low frequency means that R&D work needs to be done on the transducer-preamplifier configuration to obtain sufficient voltage output from the transducer at the 10^{-8} E level to overcome the 1/f noise of the preamplifier, without using so much transducer that the mechanical Q is affected.

REFERENCES

1. R.L. Forward and R.M. Lemmen, "Rotating Gravity Gradiometer Study," Final Report, JPL Contract 954309, Hughes Research Laboratories, Malibu, CA (30 June 1976).

2. R.L. Forward, Lunar Orbiter Selenodesy Gravity Gradient Study, Final Report NAS 8-24788, Hughes Research Laboratories, Malibu, CA (July 1971).

3. R.L. Forward, "Development of a Rotating Gravity Gradiometer for Earth Orbit Applications," (AAFE) Final Report NAS 1-10945, Hughes Research Laboratories, Malibu, CA (January 1973).

4. J.A. Gardner and E.J. Sherry, Earth Physics Satellite Gravity Gradiometer Study 760-70, Jet Propulsion Laboratories, Pasadena, CA (May 1972).

5. R.L. Forward, Electronic Cooling of Resonant Gravity Gradiometers, Research Report 516, Hughes Research Laboratories, Malibu, CA (March 1978).

SECTION III.

THE SOLAR SURFACE

N78-32971

New Perspectives for Solar Observations

H.S. Hudson
Department of Physics
University of California, San Diego

I. Introduction

The Solar Probe offers us a chance to approach our nearest star more closely than about one AU, with the obvious advantage of permitting more detailed observations of its surface than we can obtain from Earth. In these comments I will discuss the merits and demerits of a very close approach, while the subsequent contributions will describe some specific ideas in greater detail. In addition to the advantage of proximity, any deep-space mission that carries instruments for observation of the solar surface also views the Sun from a perspective different from that of Earth. This is not merely a trivial improvement over Earth-based observations, because the inhomogeneities of the solar atmosphere actually prevent a full understanding of its nature from observations at one point in space.

Let R represent the heliocentric distance of a point of observation. Table 1 then gives the R dependence of various observing parameters: the spatial resolution scales as 1/R, for example. The total flux from an unresolved point source scales as $1/R^2$. This benefits gamma-ray spectroscopy among others, where the limiting source of noise may come from photon statistics in the signal itself. The surface brightness (flux per resolution element) does not vary with R for a resolved source. The time of flight from a solar source to the point

[redacted] INTENTIONALLY BLANK

Table 1

Radial Dependences

Distance to Sun	R
Spatial Resolution	R^{-1}
Flux from Unresolved Source	R^{-2}
Flux from Resolution Element	—
Time of Flight from Source	R
Naive Equilibrium Temperature	$R^{-1/2}$

of observation scales as R; smaller times of flight are good for observation of neutrons and charged particles. Finally, one must point out that the equilibrium temperature (the simple blackbody temperature of the observatory unmodified by active or even passive controls) will increase according to the law $R^{-1/2}$ as one draws nearer to the Sun. This is the basic price one pays for getting better observing conditions.

In numerical terms, the spatial resolution of a telescope will be given by

$$d \cong 75 \left(\frac{1.25\ m}{D}\right) \left(\frac{R}{1\ AU}\right) km \qquad (1)$$

where we scale to a diffraction-limited aperture D that will give 0.1 arc sec resolution for D = 1.25 m at 1 AU. By comparison, the best ground-based resolution of 1/3 arc sec corresponds to the diffraction limit of a telescope with ~ 40 cm aperture. The "helio-synchronous" orbit occurs at R ≅ 0.17 AU, as an interesting special case of Equation (1); a satellite in circular orbit here would continuously view the same solar hemisphere (ignoring differential rotation). A 25 cm telescope at this location would give results far better than the best ground-based imaging.

II. Solar Observations from Earth and from Near-Earth Orbit.

A. Optical Wavelengths

Observations in visible light have depended mainly upon ground-based instrumentation, with slow improvement in telescope optics over the last century. The best resolution quoted above occurs only

rarely even in the new vacuum telescopes that have partly overcome seeing problems. A Shuttle-borne solar optical telescope (SOT) is now planned for high resolution ($\sim$ 0.1 arc sec, or 70 km) solar observations. It will obtain individual pictures with excellent "seeing". Even more importantly it will obtain long sequences of such pictures, something beyond the capability of any ground-based telescope. These SOT observations will achieve complete resolution of optically thick solar surface structure with horizontal size scales at the limit determined by the hydrostatic scale height of the atmosphere. This corresponds to a typical photon mean free path in normal photospheric layers. The SOT has many additional conveniences such as interchangeability of focal-plane instruments, abundant telemetry (or the use of film), capability for real-time control, and others. All of these considerations confirm SOT's definitive role in solar optical observational astronomy in the next decade.

Nevertheless important smaller-scale structure will certainly exist. One solar example is the transition region, where simple models require very abrupt temperature gradients that define formation regions much thinner than a hydrostatic scale height. The Earth's atmosphere provides many nearby examples---clouds, lightning, airplanes---of structures smaller than a scale height that have important influences on atmospheric processes. The photon-mean-free-path limit on interesting size scales in any case does not apply to optically thin regions, and size scales smaller than $\sim$70 km will assuredly be of interest for the explanation of many solar phenomena, especially those in the upper atmospheric layers.

B. Ultraviolet and X-ray Wavelengths

At shorter wavelengths all observations to date have of course been made from space platforms---notably OSO satellites. These wavelengths relate especially well to problems of the chromosphere, transition region, and corona, including dramatic forms of activity such as solar flares. Angular resolutions have not exceeded a few arc sec, and severe technical problems stand in the way of achieving high (sub-arc-sec) resolutions at the shorter wavelengths, where normal-incidence reflecting optics fail to work. Here also flux limitations may become predominant even for broad spectral bands.

C. Other

There have been virtually no imaging observations of hard (~2 keV) X-rays and energetic particles, including neutrons---"imaging" of energetic particles here means identifying and mapping their sources in detail. Hard X-rays can be imaged with suitable telescopes from 1 AU, but these observations are subject to limitation of signal-to-noise ratio by photon statistics in the weak solar signal. The energetic particles and neutrons suffer from peculiar propagation effects in their transit from the Sun to the Earth, and mapping their sources from 1 AU is virtually impossible.

III. Solar Observations from the Solar Probe

A. Proximity

Getting close to the Sun basically enlarges the apparent solar structure and makes telescope requirements easier to meet; the disadvantages include severe thermal loading, very limited spacecraft resources and short observation time. For comparable observational objectives at the level of the Shuttle solar telescope (SOT), i.e. angular

resolution 0.1 arc sec or more, a telescope at 1 AU is clearly the preferred choice. For ultra-high-resolution observations, however, it seems clear that a closer approach to the Sun will be helpful. In the extreme application of equation (1), $R = 4\ R_\odot \cong .02$ AU, a 25-cm telescope would resolve 7 km structural features---it would therefore rival a diffraction-limited 13m telescope at one AU! Given any substantial interest in these size scales, and many theories do require filamentation this small, a Solar Probe or equivalent is the only feasible choice for an observing platform.

B. Perspective

A new perspective, giving two lines of sight for solar viewing, will dramatically increase our knowledge of the three-dimensional structure of the solar atmosphere.

We know little about the horizontal inhomogeneities. The shadowing caused by such unknown structures greatly affects the limb-darkening profile in infrared and radio continuum , corresponding to heights of formation near the temperature minimum. In general three-dimensional effects must be strongest just at the most interesting parts of the atmosphere---the sites of the stronger magnetic fields---because the field can aid in establishing the inhomogeneous structures there.

The normal perspective (from Earth) of these structures does not suffice to untangle them by the traditional center-to-limb statistical approach. Compounding the problem are (i) the $\sec\theta$ dependence of the height of formation of radiation, where θ is the angle of incidence relative to the local solar normal (radiation of the same wavelength

comes from a different layer on the limb) and (ii) the greater line-of-sight averaging of inhomogeneous structures toward the limb.

It is clear that viewing from a second simultaneous perspective will give a qualitatively new and important means of studying solar atmospheric structure.

(1) For optically thick sources, two telescopes observing the same region can help sort out shadowing effects.

(2) For optically thin sources, the two lines of sight begin to make possible a full three-dimensional reconstruction, by analogy with medical X-ray tomography practice.

(3) The vector magnetic field can be more easily determined

(4) One telescope will view objects physically occulted by the solar limb for the other telescope. This extends continuous coverage and has other benefits, such as early warning of the arrival of solar activity at the limb.

These advantages are not unique to the Solar Probe Mission, since they would be available from any deep-space solar observatory. Only the Solar Probe could however completely accomplish item (2), namely the synthesis of true three-dimensional structures for stable optically-thin sources, during the rapid transit of perihelion. For example a soft X-ray or ultraviolet telescope could determine the complete luminosity distribution for a quiescent active-region loop, and of course it would simultaneously resolve it in great detail because of the close range at Solar Probe perihelion.

IV. Conclusion

High-resolution imaging of the solar surface has been and continues to be crucially important for solar astronomy. At present, we plan to obtain a high resolution ($\sim$ 0.1 arc sec) corresponding to a physical limit ---the hydrostatic scale height---with the Shuttle-borne solar optical telescope (SOT). At this resolution we may expect conclusive observations relating to the structure of the quiet solar atmosphere, sunspots, spicules, oscillations and many other current problems of solar astrophysics.

Beyond this limit important unresolved structure will exist, especially in optically-thin regions or in regions with strong magnetic fields. Furthermore, we expect ambiguity to remain in our solar imagery, because a single line of sight cannot suffice completely to untangle the vertical dimension from the two horizontal dimensions. To remedy these deficiencies, the observatory must move close to the Sun. A Solar Probe with a complement of solar telescopes would be very desirable, even given the severe limitation on weight and telemetry and special design constraints necessary to get around the intense solar heat input. Even a few images with ultra-high resolution would be immensely rewarding scientifically and would have an impact probably comparable to that of the first pictures of Martian craters or of the back side of the Moon.

C-2

OPTICAL OBSERVATIONS OF THE SUN FROM A SOLAR PROBE

H. Zirin
California Institute of Technology

Optical observations of the sun from a solar probe offer a unique opportunity to make a great step forward in solar physics by pushing spatial resolution of solar surface structure to ultimate limits with simple instrumentation. The logic is simple: a 10 cm aperture telescope at 4 solar radii resolves the equivalent of a five meter telescope in earth orbit. Given the ease of producing a diffraction limited 10 cm telescope and the difficulty of fabricating a diffraction limited 5 meter telescope, the performance of the small telescope could well be better.

No one knows the ultimate limit of size in structures on the sun. There are some who argue it is limited by the mean free path of radiation, perhaps 100 km, or by the scale height, also 100 km. There are others who believe it may be limited only by the radius of gyration in the magnetic fields, a few meters. And there are those like myself, who having found that most of the physical processes we study on the sun appear to be governed by phenomena on a scale below our resolution, simply feel that, without prejudgement, we should find out what is going on at the smallest possible scale.

The arithmetic on resolution is straightforward. The modulation transfer function (MTF) of a diffraction-limited system is

$$MTF(v) = \frac{2}{\pi} (\varphi - \cos\varphi \sin\varphi) (\cos\theta)^k$$

$$\varphi = \cos^{-1} \left(\frac{\lambda v}{2NA}\right)$$

where

v = spatial frequency (cycles/mm)
λ = wavelength (mm)
NA = numerical aperture (2/f number)
θ = half-field angle
k = 1 for radial, k = 3 for tangential.

The MTF is the ratio of the intensity modulation of the source to that of the image. The last term in brackets is of order unity. The limiting resolution, ν_o is given by:

$$\nu_o = \frac{1}{\lambda\ (f/\#)} \cdot$$

Since at this resolution the modulation is zero, it is not too useful. The modulation in the final image should be such as to produce a reasonable signal to noise ratio in the counted photons per detector element. It is reasonable to use about .05 as a limiting MTF.

Since one can practically expect to use detectors of about 20 microns size, this has an MTF of .05 for f/27; using a 10 cm aperture, this gives 1.5 arc sec per detector element, or about 20 km at the sun. Of course other ratios might be adopted according to data rates, etc. Exposures would be quite short even for monochromatic observations.

Because data rates are limited, it would be wise to only look at limited areas where resolution and not the general picture, are important. For example, a granule would cover 2500 elements of the above detector; since the granule lives 8 minutes, one would want perhaps 5 frames or 12500 data points, each with many gray levels (say 6 bits) or 72000 bits/8 minutes. Since more than one granule should be studied at once (say 50), one comes to 7200 bits per second for 20 minutes. Such a sequence would define rather well the structure and evolution of individual granules. In the case of Hα chromospheric structure, the structures are much more varied (if you saw one granule you saw them all, within limits). On the other hand, one can settle for individual frames which reveal structure. In the case of spicules, one requires again sequences of a limited number of objects which may be assumed to be a homogeneous class.

The list of phenomena for which very high resolution images will provide new understanding is long: sunspots,

flare kernels, the extreme limb of the sun (for vertical structure and possible oscillations). In every case one feels that a really good look, a really good measure of emission as a function of space and time, will clear up some of our most pressing problems. If there are any very small sunspots associated with the filigree field (the strong quiet sun fields) we should see them. Of course, providing more sophisticated secondary instrumentation, such as magnetograph capability, will be even better; but if weight and power constraints keep the instrument to a simple white light telescope or maybe an Hα filter, the achievement will be great.

It should be emphasized that this instrument complements planned solar optical telescopes in earth orbit. The time and data limitation mean it can only go after one-shot problems. On the other hand, there is no obvious way that this resolution could be obtained from earth orbit in the foreseeable future. The telescope discussed would only need a few arc seconds of pointing - the same resolution at the earth would require .02 arc sec pointing. The difficulties of figuring and aligning a great telescope in earth orbit to this capability are well-known. True, we will only get a few glimpses, but they will guide our thinking for years to come, as did the first closeup frames of Mars.

I have not mentioned stereo viewing; what actually can be done depends greatly on details of the instrument and time limitations. Three-dimensional studies of sunspots and filaments are particularly appealing. It is doubtful that any of the fine structures will last for the interval needed for a real 3D effect, but particularly the Wilson effect of the depression of sunspots will be important to study.

SOLAR PROBE STUDIES OF THE SOLAR CONVECTION ZONE

R. K. Ulrich

Department of Astronomy

University of California, Los Angeles

A Solar Probe mission would provide a unique opportunity to study the dynamics of the solar convection zone. The critical portion of this zone occurs at or just below the solar photosphere. The important size scale of 500 to 1000 km for the energy-carrying eddies is just at the limit of somewhat better than 1" resolution when the Sun is viewed from 1 AU. This quality of resolution has been attained with the largest practical telescope size of about 24". From a distance of 3 $R_\odot$ as may be attainable with a Solar Probe, resolution of 5" such as would be possible from a small telescope will allow observations of features as small as 50 km. Since the Solar Probe will be as close to the Sun as 0.014 AU, the effective resolution is increased a factor of 70. A 3" telescope on the Solar Probe will have resolution equivalent to a 200" telescope on Earth. Thus observations could be carried into the size scale which presumably is responsible for the turbulent viscosity. A test of theoretical dynamical models will then be possible. Full three-dimensional models are now under development at UCLA by Dr. Undo Uus on leave from the Tartu Astrophysical Observatory in Estonia. These models will be available well before the mission to provide a framework for the interpretation of high resolution observations.

The preferred instrument is a magnetograph operated in a Doppler mode. A Fabrey-Perot etalon can provide the spectral discrimination necessary for the measurement of velocities. An optical pointing system will be required which has the capability of tracking a fixed portion of the solar surface with an accuracy of 3". Thirty minutes of observation

at a rate of one 128 x 128 frame every 15 s will permit a detailed study of the evolution of the granulation through three lifetimes.

This instrument can also provide long time base observations of the solar p-mode oscillations before and after the solar passage. These observations, when combined with similar ground-based observations, will permit a very accurate determination of both the horizontal wavelength and the frequency. This combination will allow a determination of the rate of solar rotation at a depth 25% below the solar surface to an accuracy of better than 0.5 km/s. Flare-generated oscillations will be observable and will permit a detailed test of the average envelope structure below the solar photosphere. The instrument operated in a magnetograph mode can be used to study the filamentary magnetic structure.

This research was supported in part by NASA:NSG 7407 and NSF:AST76-20260

THE DETERMINATION OF THE STRUCTURE AND HEATING MECHANISMS OF CORONAL LOOPS FROM SOFT X-RAY OBSERVATIONS WITH THE SOLAR PROBE

John M. Davis and Allen S. Krieger
American Science and Engineering, Inc.
955 Massachusetts Avenue
Cambridge, Massachusetts 02139

Abstract The question of whether high resolution soft x-ray imaging from the Solar Probe can be justified in terms of the expected scientific return is answered in the affirmative. A specific objective namely, the determination of the temperature and density structure of a coronal loop, is defined. Its understanding is shown to be crucial to the more general problem of coronal heating. An instrument and its effect on the design of the Solar Probe spacecraft is discussed.

1.0 INTRODUCTION

The Solar Probe Mission is designed to make a single very close approach to the sun which will allow in-situ measurements of the acceleration region of the solar wind and high energy particles and a determination of the solar quadrupole moment J_2. In addition it may be possible to use the probe as a platform for taking close-up pictures of different regions of the solar atmosphere. However, because of the fly-by nature of the mission, the time spent close to the sun is short and the number of images that can be obtained will be small. Therefore we have taken a critical look at the uniqueness of the scientific results which we could expect and weighed them against the technical difficulties which must be overcome to reduce the severe environmental constraints imposed by the close solar approach.

On balance we believe that a strong case can be made for imaging experiments in several wavebands; however we have restricted our presentation to soft x-ray wavelengths which provide images of the solar corona. For imaging experiments in general, the advantage of the Solar Probe is that it will provide higher spatial resolution measurements than can be obtained currently from Earth orbit. In the soft x-ray band the statement will be true for the next decade and probably for the remainder of the century. Therefore the perihelion passage of the Solar Probe provides a singularly unique opportunity. To take full advantage of it, we have developed an observational program which, we believe, could have considerable impact on the future development of coronal and space plasma physics.

The measurements we would perform are crucial to the general problem of coronal heating. This problem has been central to the study of solar physics for over 30 years and there exists no concensus between theory and observation except in the most general terms.

No immediate change in this unsatisfactory situation is likely because the differences among the various theoretical descriptions can be tested only at a level of resolution far superior to that currently available. Consequently, there exists little incentive for theorists to develop precise models which are suitable for observational test. We believe that there are specific measurements which could be made from the Solar Probe which would change this situation. By providing information on the fine structure of the corona, the measurements would force a more critical review of the competing theories of coronal heating to be made.

Space observations of the sun at x-ray and EUV wavelengths have shown that the material in the inner corona is, to a large extent, concentrated in loops which surround the lines of the solar magnetic field (Figure 1). The loops are at coronal temperatures (above 2×10^6 K) and have lifetimes that are very long compared to radiative or conductive cooling. Thus the observations imply that the loops are undergoing continuous heating. However, neither the power source nor its transfer mechanism have been observationally identified. This is because the effects of different heating mechanisms on the observable parameters, plasma temperature, density and pressure, will be most noticeable in the direction perpendicular to the field lines and, in this direction, the heating region is predicted to be at most 100 km thick. This is well below the resolution of current instruments in Earth orbit although not below the resolution of the same instruments if they are carried close to the sun. Thus the objective for a soft x-ray imaging experiment on the Solar Probe can be stated as follows: to obtain stereoscopic high resolution soft x-ray images of a coronal loop from which electron temperature, density and pressure maps can be constructed with a spatial resolution of 20 km. The maps would

against which the competing mechanisms of loop heating can be tested.

In the following sections we will briefly review the structural differences of coronal loops that might be expected on the basis of current theoretical ideas of loop heating, review the capabilities provided by the Solar Probe that are applicable to imaging experiments, and finally outline a possible instrument and its consequences for the design and operation of the Solar Probe.

2.0 LOOP HEATING MODELS

High resolution soft x-ray imaging has demonstrated that the inner corona is a complex structured region in which the material is concentrated in loops of varying height, length, temperature and density (Figure 1). The material in the loops is at coronal temperatures ($> 2 \times 10^6$ K). The observational evidence suggests that they undergo continuous heating. Thus the classical problem of heating a homogeneous corona has been reduced to that of heating individual loops. It is conceivable that the different classes of loops are heated by different mechanisms but, since no one has proven observationally how any loops are heated, we would postpone discussion of this question. The loops we would observe are those which occur near the centers of active regions. Present observations have sufficient resolution to resolve these loops but they are unable to unravel the structure of individual loops or to specify the radial variation of the crucial plasma parameters. Measurement of the variation in the direction perpendicular to the field lines is particularly important for in this direction thermal conduction and mass flow are inhibited, allowing different regimes to be maintained. Conversely, along the field lines where conduction and flow are allowed they will act to smooth out differences.

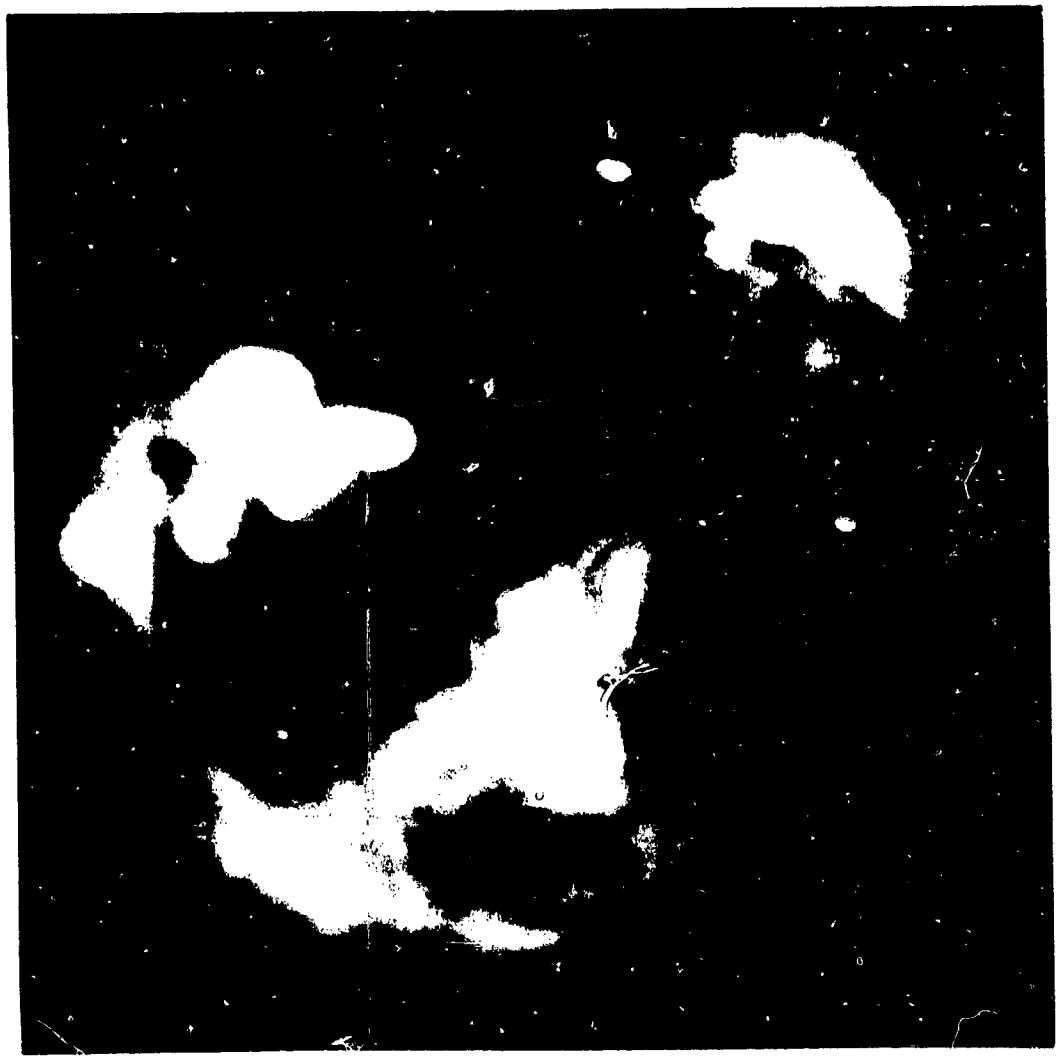


Figure 1. A Soft X-Ray Photograph of Solar Corona Taken on January 31, 1978 and Showing the Predominance of Loop Structures.

Many models of coronal heating rely on wave dissipation as the mechanism which transfers energy from the convective zone to the corona. Others rely on the dissipation of magnetic flux generated by subsurface convective motions. A more important distinction, observationally, exists between those models in which loops are heated in shells at their outer surface and those in which loops are heated in individual strands more or less uniformly distributed in radius throughout the presently resolvable loop bundle. Of the many different classes of loop heating mechanisms we have chosen two conflicting models to exemplify the disparities in our understanding of coronal heating.

The two possible internal structures are represented schematically in Figure 2. The first is based on the model of Ionson (1978) in which heating occurs in a narrow sheath surrounding a much cooler and larger core. The heating is a result of the resonant absorption of kinetic Alfven waves which are generated by Alfvenic surface waves. Here the term surface refers to a non-uniformity in the Alfven speed. The sheath in which the irreversible heating takes place is only a few kilometers thick for typical coronal plasma parameters. However it is separated from the cool core by a boundary layer which is approximately 100 km thick and in which the temperature gradients are large. The transverse dimension of the sheath and boundary layer together are unlikely to exceed the resolution element of any instrument projected for use in Earth orbit; however, if they exist, they will be resolvable from the Solar Probe. One would anticipate that its characteristic signature would appear as a difference in the magnitude of the apparent temperature gradient, resulting from a line of sight average, at the physical boundary of the loop. A schematic diagram of what we would expect to see for a pure surface heating model is shown in the upper right hand panel of Figure 2.

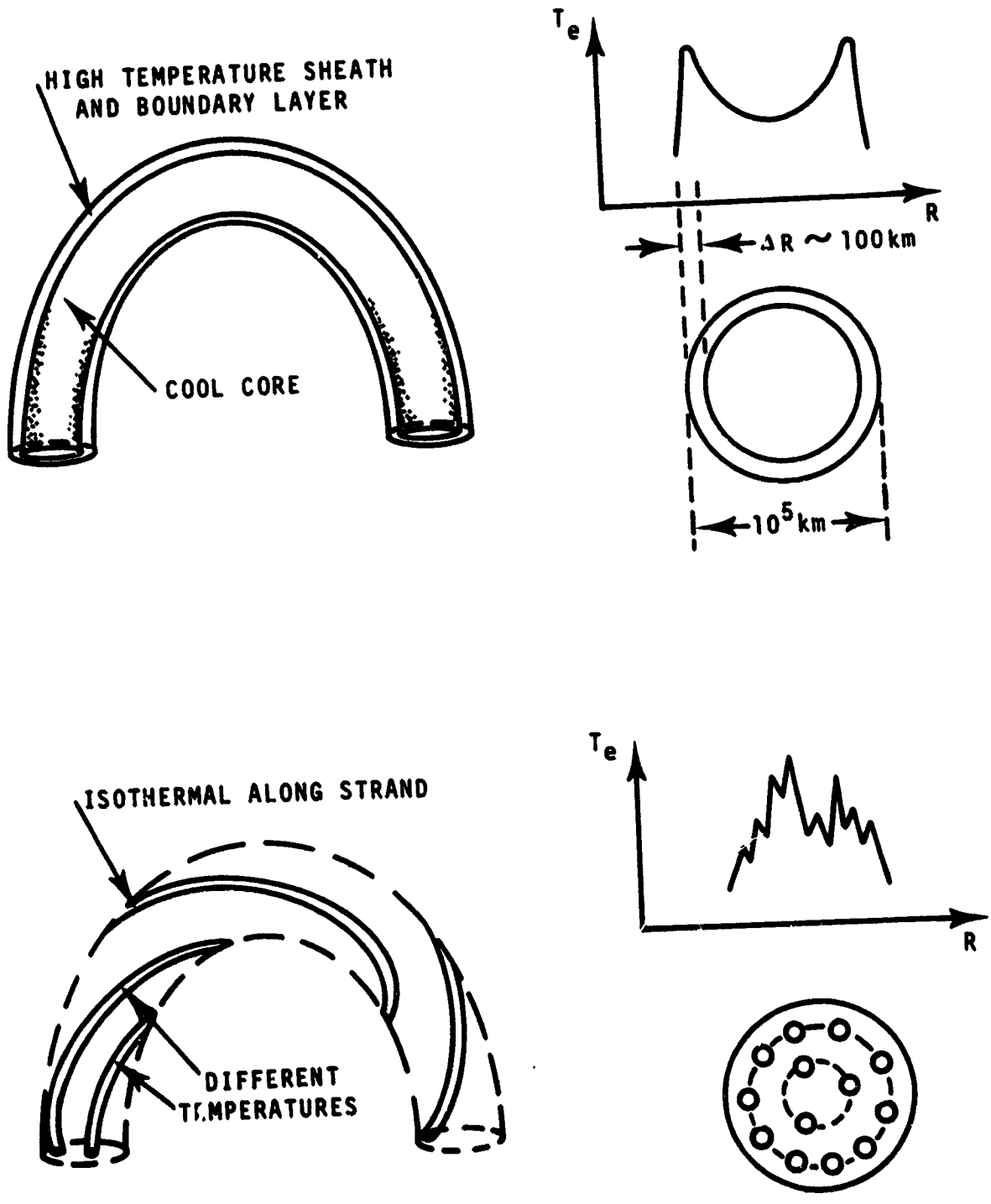


Figure 2. **Schematic Representations of Two Conflicting Loop Models and the Corresponding Radial Temperature Distributions which might be observed.**

In a contrasting model the loop is thought to consist of many separate strands representing the plasma distribution around individual field lines. The strands may or may not be twisted as shown in the bottom half of Figure 2. The temperature distribution along each strand is probably isothermal owing to the high conductivity along the field although different strands may be at different temperatures. In a variant of this model proposed by Rosner et al. (1978) the strands result from current filamentation and the heating results from anamalous current dissipation. The strands might be located in sheaths concentric with the loop axis. The current sheaths are expected to be quite narrow although presumably they will be surrounded by transition layers. If this mechanism is correct the line of sight temperature distribution across the loop would be unsymmetrical and would vary almost randomly at high enough resolution. Also in this case it should be possible to trace the longitudinal paths of the individual strands of the loop.

In this very brief review we have discussed the effect on the internal temperature distributions of two typical models of loop heating which could be uniquely distinguished from the Solar Probe and probably by no other instrument. In light of the complete absence of observations at this level of resolution the results from Solar Probe would be of major importance. By establishing limits on the scale size of plasma inhomogeneities, on the presence and widths of heating zones and boundary layers, or of other non-uniform geometries or topologies the Solar Probe would provide new observational constraints for theoretical models together with an increased incentive to place the models on a more practical basis.

3.0 HIGH RESOLUTION IMAGING FROM SOLAR PROBE

In this section we review the technical benefits and constraints presented by the Solar Probe mission for high resolution imaging and compare and contrast them with the expected capabilities of proposed Earth-orbiting instruments. The parameters so derived can then be used as a basis for specifying the uniqueness of the science that can be performed with this mission.

The major concerns to be satisfied in the design of any space imaging experiment are to produce a match between the resolution of the optics and the detector and to balance this against the capability of the transmission system to record the data. The present generation of x-ray imaging telescopes have limiting resolutions on the order of one arc second. This angle corresponds to a linear distance on the sun of 700 km, when viewed from Earth orbit. The optics for the Soft X-ray/XUV Facility telescope which will probably be developed in the next few years for flight on the space Shuttle will have a design goal of 1 to 2 tenths of an arc second corresponding to 70 to 140 km on the sun. The fabrication of such an instrument will be a major undertaking. It is by no means certain that the design goal can be met or, if it is, that the pointing stability of the spacecraft would enable its full potential to be realized. However even if the design goal is achieved, the spatial resolution is still likely to be insufficient to reveal fine enough detail to differentiate between the various models of loop heating. On the other hand a present day mirror with 1 arc second resolution would have an "effective" resolution of 2/100s of an arc second and would be able to resolve 14 km on the sun if it was operated at 4 $R_\odot$, the perihelion distance of the Solar Probe. This is a factor of 5 better than the most optimistic estimate for resolution from Earth orbit and the scale is now at a level where

meaningful tests of current theories of coronal heating could be made.

To achieve this goal in a practical system it is necessary to match the resolution of the optics with that of the detector which we assume will be an improved back-illuminated CCD. CCD's are projected to be available soon in 800 x 800 arrays of 15 micron elements. Coupled with a telescope of 2m focal length this results in a pixel size of 21 km at closest approach and a corresponding field of view of 23 x 23 arc seconds as seen from Earth orbit. This appears to be a reasonable compromise. Reducing the pixel size any further would increase the information content of each image to a level which would severely impact the transmission capability of the spacecraft. Thus, an 800 x 800 array contains 6.4×10^5 pixels, if each pixel requires 8 bits for intensity identification, a single image contains $\sim 5 \times 10^6$ bits. With the proposed data transmission rate of 5×10^2 bits/sec, for the imaging experiment, a single image would take almost 3 hours to transmit and two images would fill the proposed bubble memory. However since close to a factor of 2 reduction can be gained using reasonable data compression schemes, a program in which 6 images are recorded during the perihelion passage is possible with present technology. These images would be sufficient to allow stereoscopic reconstruction of the target loop and to map the plasma parameters.

The alternative approach to improving the resolution, increasing the image size by increasing the effective focal length of the telescope, would result in a corresponding decrease in the field of view. This would make it increasingly difficult to locate a suitable target. In Figure 3 we show a different exposure of the sun, taken shortly after Figure 1, which enhances the active region loops. Two possible targets are identified. The 23 x 23 arc second field of view is sufficient to contain a single loop. But, obviously, any further reduction in

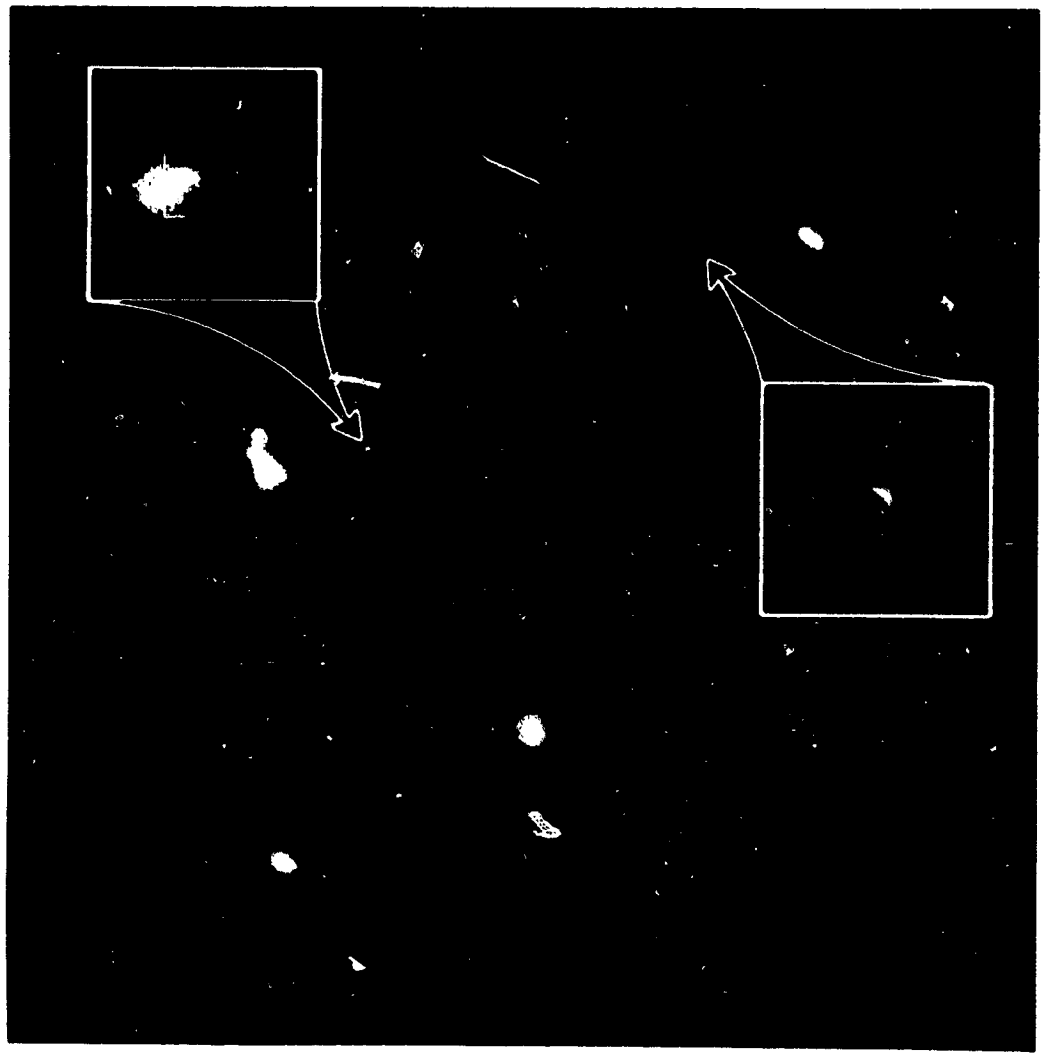


Figure 3. The Location of a Loop within the Proposed Field of View of the Imaging Experiment. The Smaller Box has Sides of 23 arc seconds.

the field of view would severely reduce the probability of locating a suitable target. This image (Figure 3), and our previous experience, indicate that several suitable targets will be available provided encounter does not occur at sunspot minimum. The loops we have selected are associated with bright x-ray features. This is done deliberately as we believe that the fine pointing of the spacecraft would have to be done on-board by a servo system driven from the imaging display. Small bright features will be easier to lock-on to.

Thus our preliminary analysis indicates that it is feasible to obtain the 6 images which would be necessary to analyze the three-dimensional structure of a coronal loop. Pairs of observations would be taken at three points in the orbit, including perihelion, with a spatial resolution of close to 20 km. This level of resolution is sufficient to meet the scientific objectives described earlier.

4.0 INSTRUMENTATION AND OPERATIONS

The two techniques which have been applied to soft x-ray imaging are the coded mask or multiple pin-hole camera and the grazing incidence telescope. Because the major objective of the Solar Probe instrument will be to obtain the highest resolution images possible, we suggest that the superficially simpler coded mask is less suitable than a grazing incidence telescope, as the latter instrument is capable of much higher intrinsic resolution. A further disadvantage of the coded mask is that the image is contaminated by side lobes which have to be removed during ground processing. However their removal requires that two pictures with complementary coded masks be taken, thus doubling the number of data bits to be

stored and transmitted. As we showed in Section 3.0 such a doubling could severely compromise the usefulness of the experiment by placing severe demands on the spacecraft telemetry system.

The instrument package we suggest would consist of a grazing incidence mirror, a thermal pre-filter, a three position filter wheel and a focal plane detector baselined as an 800 x 800 back-illuminated CCD. The structural assembly together with the data processing and control electronics would draw heavily on the designs being developed for the Solar Polar Mission. A layout of the instrument is shown in Figure 4.

Two major constraints that the experiment design will impose upon the spacecraft are:

1. That the mirror must look directly at the sun through the heat shield, and
2. that the small field of view requires that the instrument be pointed precisely (~ 1 - 2 arc second) and that the location of the pointing axis on the solar feature be maintained throughout the encounter so that stereoscopic images of a single target can be obtained.

The first constraint results from the fact that, unlike visible optics, a 45° mirror cannot be used to reflect the x-rays into the experiment bay from outside the heat shield. Although it might be possible to use either Bragg reflectors or a sector of a grazing incidence optic mounted outside the heat shield but whose focal point falls within the shadow of the shield, we have at this stage taken the direct approach of assuming that the instrument will look out through the heat shields. Although this approach may appear to be the most radical, we believe that the constraints it places on the spacecraft are not unreasonable. If the primary heat shield is composed of molybdenum it will not transmit soft x-rays and it will be necessary

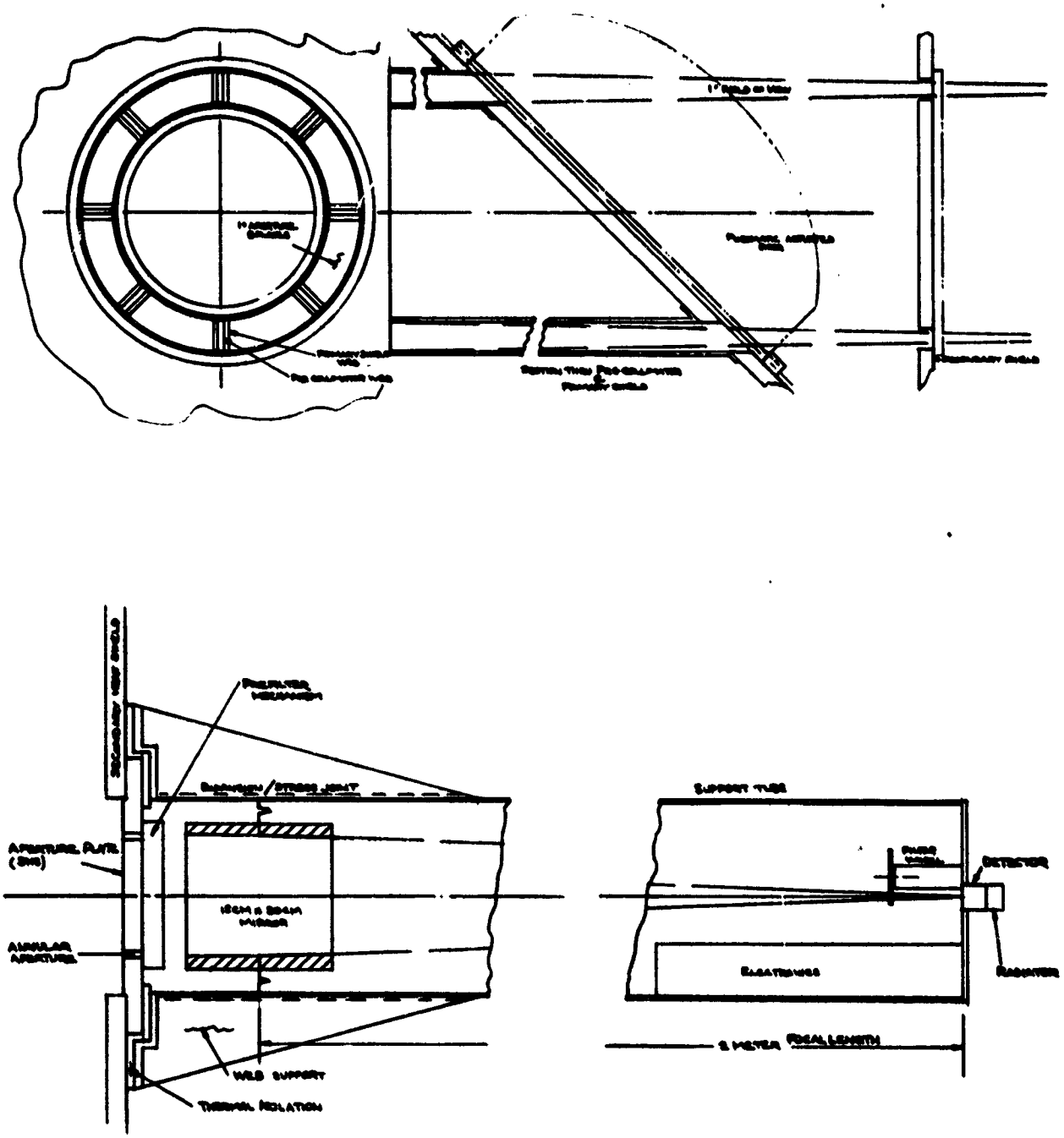


Figure 4. **A Preliminary Layout of a Soft X-Ray Grazing Incidence Telescope for the Solar Probe. The Upper Section shows the Heat Shield and Pre-Collimator and the Lower Section the Telescope.**

energy passing through this aperture, we suggest that it be provided with both a door and a pre-collimator. If the pre-collimator is designed to limit the field of view to 1^{o}, at 4 $R_{\odot}$ only one thousandth of the solar surface would be viewed. Since the solar energy flux is proportional to the inverse distance squared its total magnitude will be 3000 times greater at 4 $R_{\odot}$ than at 1AU. However the effect of limiting the field of view will be to reduce the energy entering the instrument to only 3 times the value in Earth orbit. Although the calculation is over simplified, for instance it neglects the direct radiation from the pre-collimator which will be at the temperature of the primary shield, it does show that the perhaps alarming suggestion of providing an open aperture in the shield is not totally unrealistic.

Infra-red and visible radiation would be prevented from reaching the interior of the experiment package by aluminum pre-filters, mounted either on the front or back of the secondary heat shield, depending on the expected internal temperature distribution.

During the twelve hour period centered on perihelion, where the thermal loads are greatest, the instrument will record only 6 images with a total exposure time of no more than 1 - 2 minutes. Therefore the aperture in the primary heat shield could be fitted with a door to minimize the energy entering the instrument. Although the design of the door mechanism will present an interesting challenge, because of the high temperature at which it will be required to operate, it is within the state of the art.

The second constraint concerns the pointing of the instrument. Controlled pointing will be required because the number of suitable targets will be small and they will probably occupy less than 1% of the area of the solar disk. The pointing axis of the instrument will have to follow the target as the spacecraft orbits the sun from North to South. As it is unlikely that the sub-satellite point will

be able to offset point from the purely radial direction so that it can look at a target within a band of solar longitudes ± 20 degrees wide. It must also be able to offset in an orthogonal direction in order to maintain the target with the field of view during the 12 hour encounter period. Both these requirements translate into orthogonal motions of less than ± 10 degrees about the satellite-sun center line. Although this requirement does not appear unreasonable, the sizing of the heat shield, during the spacecraft design, should take these maneuvers into account.

Finally our decision to observe a loop was based in part on the fact that they are relatively long lived features. This would allow target selection to be made from Earth orbit 24 - 36 hours prior to encounter. The spacecraft orientation could be calculated and transmitted to the Probe prior to the close encounter phase. However this requirement, makes it imperative that the trajectory of the Probe is to the west of the sun, as seen from the Earth, so that the targets can be observed and studied prior to encounter.

5.0 CONCLUSIONS

The Solar Probe Mission will provide a unique opportunity to obtain high resolution soft x-ray images of the solar corona. The level of detail in the images is likely to be an order of magnitude finer than can be expected from other space observatories before the end of the present century. The disadvantage of the mission is that, as a fly-by, the number of high resolution images that can be obtained, and the area of the solar corona that they will cover, is small. Therefore we have defined a single objective, the determination of the temperature and density structure of a coronal loop from stereoscopic x-ray images. The observations would, for the first time, be sufficiently detailed to distinguish between the competing

has been one of the central problems of solar physics for the last thirty years observations leading to its solution would be extremely worthwhile and would more than justify the inclusion of x-ray imaging experiments on the mission.

We believe that the images should be obtained by using a grazing incidence telescope. We have analyzed the requirements that this would place on the spacecraft and have found them to be not unreasonable. However they are fairly stringent and should be kept in mind by the designers in the development of the spacecraft.

REFERENCES

Ionson, J. A., Resonant Absorption of Alfvenic Surface Waves and the Heating of Solar Coronal Loops, Astrophys. J. (to be published), November 1978.

Rosner, R., L. Golub, B. Coppi and G. S. Vaiana, Heating of Coronal Plasma by Anamalous Current Dissipation, Astrophys. J. (to be published), 1978.

D11

N78-32975

High Resolution Stereoscopic X-Ray Imaging of Coronal Features during a Solar Probe Flyby

Dayton W. Datlowe

Lockheed Missiles & Space, Inc.
3251 Hanover Street
Palo Alto, California 94304

The possible role of a Solar Probe mission in answering fundamental questions about the structure and heating of coronal loops is examined. The experimental technique consists of imaging 1-10 keV X-rays to give accurate temperature profiles of hot active regions and post-flare loops. A limitation on the interpretation of such pictures is that stereoscopic reconstruction of the 3-dimensional arches requires many lines of sight. This kind of information can be provided only by a rapid solar flyby. In addition, the proximity to the sun will provide useful spatial resolution with compact instrumentation. The pictures thus obtained will provide crucial tests of theoretical models of coronal arches.

Recent studies of the dynamics of the solar corona have emphasized the importance of loops in active region and flares. In flares, all of the high temperature plasma, all of the short wavelength thermal X-ray emission, and perhaps all of the energy release occur in loop structures. Skylab soft X-ray images [1,2] have shown clearly the X-ray emitting arches which are at the center of active regions. Thus the physics of arch structures is vital to an understanding of the solar coronal dynamics.

Although the morphology of coronal arches is well defined, at least in general outline, the temperature structure of these arches is not at all well known. The principal problem is that earth-bound measurements have a line of sight which includes gases at many temperatures. Since the X-ray emission from arches is optically thin, the contributions from a wide range of temperatures must be unfolded. Several lines-of-sight are required to find the temperature distribution. In addition, observations made in only one projection leave an ambiguity about the inclination of the arch to the line of sight and to the Sun.

Active regions and post-flare loops have lifetimes from hours to days. Significant evolution during the time scale of the Solar Probe flyby is not expected, except during impulsive flares. Thus the solar probe is a unique opportunity to image X-ray arches from many perspectives. In this

regard the Solar Probe is better than simultaneous observations by a few widely separated interplanetary probes.

The X-ray energy range best suited to temperature measurements of coronal plasma is 1 to 10 keV. Since the photon energy is greater than the mean thermal energy of the plasma, a small change in temperature makes a large change in the X-ray flux. This wavelength range is about one order of magnitude higher in energy than the Skylab broadband X-ray images, which were weakly sensitive to temperature variations in hot 3-20 x 10^6 K plasmas. The long wavelengths are much better suited to morphological studies than to temperature profiling.

X-ray measurements in the 1-10 keV range can also be used to study flares in conjunction with a hard X-ray imaging experiment. These measurements will define the location and role of the energetic hard X-ray emitting electrons in making the 10-20 x 10^6 K flare plasma. This can be a crucial test of flare models. However, because of the large amount of information in an X-ray image and limited telemetry, special modes would be required to follow phenomena which change as rapidly as impulsive flares. Impulsive hard X-ray bursts are produced by subflares and have been observed at earth [3] to occur at the rate of about once per day. Without accounting for the increase in sensitivity produced by the Solar Probe's proximity to the Sun, there is a 50% chance of observing a hard X-ray burst during flyby. Since flares occur in the same active regions which are the targets of coronal arch studies, there is a good probability that when a hard X-ray burst occurs, it will be in the field of view.

In summary, pictures of the solar corona taken in the 1-10 keV energy range will see the hot plasma in arches, and the large number of perspectives which a solar probe flying will provide can be used to unfold their temperature structure uniquely.

The theoretical study of arch structures [1,4,5] indicates that the loops should be hot at the top and conductively transport energy to cooler transition zone layers. For a static arch to be stable, it must be hottest at the top [1]. Thus there is good theoretical justification for expecting temperature gradients along the loop length. However, Skylab observations in monochromatic lines (assumed to be formed at the single temperature of maximum emissivity) show approximately constant loop cross-sections at a given temperature. The best estimate for the temperature structure shows [6,7] that the principal temperature gradient is radially across the loop magnetic field, being coolest at the center. The gradient along a particular magnetic field line is very small. Such temperature profiles have led to the invocation of esoteric plasma processes [8,9] to salvage static arch structures. Clearly there is a dichotomy between the most complete models and the best available data on the structure of coronal arches. High resolution 3-dimensional temperature profiles will go a long way toward achieving a good understanding of the physics of these structures.

In the energy range 1-10 keV, hot arch structures are expected to give a bright image on a dark background. The relatively diffuse structures seen in the longest Skylab soft X-ray images will not be detectable; the images should look more like the shortest X-ray telescope exposures. However, the relatively few regions which have hot plasma will be intense emitters.

Thus reasonable exposures can be obtained in a few minutes with very small apertures. The choice of an aperture involves a trade-off between angular resolution and sensitivity. In the present case apertures of 20-100 μ (.02 - .1 mm) will be sufficient. Although such apertures are negligible in size compared to the Skylab X-ray telescopes, they are adequate because

1) photo-electric detection is more efficient than film and
2) only structures which are very bright by Skylab standards are to be studied.

X-rays in the energy range 1-10 keV cannot be focussed. The simplest suitable imaging device is a pinhole camera. Since the required collecting area is so small, such a device can give useful angular resolution in an experiment of practical size. Of course, the X-rays cannot be reflected with a mirror, so that the pinhole must be placed in the front heat shield of the spacecraft. However, the heat leak thus created should be small compared to the intensity reflected onto the experiment package by an external mirror. With appropriate collimation the heat leakage through the second and third shields can be kept low.

The angular resolution of a pinhole X-ray camera would be 1' to 15", depending on the photoelectric detector type used. Assuming CCD's, this would give a pixel size of 500 km on the Sun at 10 $R_\odot$ and 200 km at closest approach. While this resolution is very modest compared to what focussing instruments could achieve at longer wavelengths, it should be more than adequate to determine the temperature gradients in arches which are 10^4 km long.

Determining the temperature at each point in the picture means that every pixel is a pulse-height distribution. For a 256 x 256 array with 6 pulse-height levels, this is a very large number of bits. However, there is a large amount of redundant information which can easily be squeezed out before transmission. Since we are looking at a bright image on a dark background, only about 10% of the image will be brighter than threshold. Also since the lowest energy photons are much more probable, the amount of pulse-height information per pixel is less than six bits. The amount of non-redundant information is about 3×10^5 bits/picture, which would take 30 minutes to read out at 200 bits/sec. For the coronal structures which are the objectives of this experiment, one picture every half hour is sufficient, thus during the flyby about two dozen pictures would be obtained. Many more synoptic, low resolution pictures could be obtained during the approach phase.

In conclusion, the Solar Probe will provide an opportunity to make a fundamental contribution to our understanding of the structure of coronal loops. Since a proper unfolding of the temperature profile requires many lines of sight, the measurements can only be made from a rapid solar flyby. The closeness of the approach makes high linear resolution measurements feasible with very simple X-ray instrument, a pinhole camera. With such a mission we would for the first time be able to construct accurate temperature profiles of hot coronal loops, and thus provide significant input to the theoretical understanding of how the solar corona is heated.

REFERENCES

1. Rosner, Tucker & Vaiana, Ap.J. 220, 643 (1978).
2. Pallinicini, Serio & Vaiana, Ap.J. 216, 108 (1977).
3. Datlowe, O'Dell, Peterson & Elcan, Ap.J. 212, 516 (1977).
4. Culhane, Vesecky & Phillips, Solar Physics,11, 117 (1970).
5. Antiochos & Sturrock, Ap.J., 220, 1137 (1978).
6. Foukal, Ap.J., 210, 575 (1976).
7. Dere, Horan & Kreplin, Ap.J., 217, 976 (1977).
8. Rosner, Golub, Coppi & Vaiana, Ap.J., 222, 1317 (1978).
9. Ionson, Proc. of the November 7-10, 1977 OSO-8 Workshop, p134.

N78-32510

THE ANALYSIS OF HIGH SPATIAL RESOLUTION UV AND X-RAY IMAGES BY COMPUTATIONAL MODELING

J. F. Vesecky, S. K. Antiochos and J. H. Underwood

Stanford University, Stanford, CA 94305

ABSTRACT

High-temperature loop structures observed at UV and X-ray wavelengths are the dominant features in the solar corona above solar active regions. The Solar Probe Mission offers the possibility of obtaining very high resolution (~20 km) images of these structures including steroscopic views. This possibility raises the question of how to exploit this very detailed information in terms of better understanding of physical processes in the corona. After describing an existing computational model and displaying some sample results, we suggest that detailed computational modeling of coronal structures can indeed take advantage of very high resolution images. The sample results include the run of temperature T and number density n along a typical loop and the variation of the differential emission measure ξ with T. Noting that the integration of $\xi(T)$ along a column commensurate with an instrument's spatial resolution is the relevant parameter obtained from UV and X-ray observations, we show the effects of loop geometry and energy input on $\xi(T)$.

I. INTRODUCTION

The Solar Probe mission offers the possibility of obtaining very high spatial resolution UV and X-ray images of coronal structures. Skylab/ATM UV and X-ray images resolved coronal structures on a scale of $\sim$ 1500 km from Earth orbit. The same imaging optics carried on Solar Probe could resolve coronal structures to $\sim$ 20 km. The Solar Probe also offers the advantage of viewing the same object from a variety of aspects in a short period of time thus obtaining three-dimensional (stereoscopic) information on quasi-static coronal structures. Can we take advantage of such observations in terms of better understanding the physical conditions and processes in coronal structures? This paper suggests an affirmative answer to the question by briefly describing a new computational model of coronal loop structures. This model is based on current understanding of the physics of these structures and reveals structural details on spatial scales below the resolution of Skylab/ATM UV and X-ray images. The model described here is, of course, only a forerunner of more sophisticated models which will presumably be created to interpret future, higher resolution coronal imaging experiments. Such experiments could be carried out from Earth orbit with large aperture optics as well as from near the Sun as with Solar Probe.

Much of the plasma in closed-field regions of the solar corona is confined to more-or-less discrete loops or arch-like structures with enhanced plasma density which are brightest and most striking over active regions. For a large part of their lifetimes these loops are in a quasi-static state with little change in structure or brightness (e.g., see Rosner, Tucker and Vaiana, 1978), which implies that no mass flow is associated with the loop. Since the lifetimes of loops are frequently greater than typical time scales for radiative or conductive energy losses,

a continuous energy input must be assumed. Hence we are led to a steady state model in which the energy deposited (by mechanical waves or other means) in each elemental volume is balanced by radiative losses from the element or conductive transport of energy into or out of the element. Because conductive transport is severely inhibited across magnetic field lines, we consider only conduction along magnetic field lines, i.e., along the loop. The particular form of the magnetic field used for calculations thus far is a line dipole. Although other forms such as a point dipole could easily be implemented, the line dipole is consistent with the geometry of many observed loop structures such as arcades of loops.

It should be remembered that the steady state discussed here does not last forever and that the stage in loop evolution described by quasi-static models must be preceeded and followed by out-of-equilibrium, transient states during which material flows along magnetic field lines, in or out of the loop. Pye et al. (1978) present an interesting qualitative model for the transient states.

Rosner, Tucker and Viana (1978) and Craig, McClymont and Underwood (1978) have independently investigated quasi-static loop models using an analytic approach. Our treatment differs from these principally in that it is computational and can therefore include several effects which were neglected in the analytic treatments in order to make the equations mathematically tractable. These effects include variation of the loop cross sectional area, energy deposition along the loop, a more realistic radiative loss function, and solar gravity. A computational model allows detailed comparisons between model loops and observed loops on spatial scales well below the $\sim$ 1500 km resolution of Skylab/ATM UV and X-ray images.

This and future models will allow interpretation of coronal images and other solar data in terms of such questions as: What is the form of the energy input along a loop? How does the energy input vary with magnetic field strength, temperature, density, etc. and with what heating mechanisms are these variations consistent? Is the X-ray and UV emission from an active region well described by a model consisting of an ensemble of loops? Does the energy input vary significantly from loop to loop within an active region? Can we explain bright knots or other anomalous emission along a loop simply by variations in energy deposition and, if so, what deposition mechanisms are consistent with the plasma parameters indicated by the model. Are twisted flux ropes necessary in modelling a coronal loop? High resolution images of coronal structures, such as could be obtained from a Solar Probe mission, coupled with detailed interpretation techniques, such as described here, could provide answers to these questions. Such answers have implications regarding the physics of the convection zone via the distribution of magnetic fields and mechanical energy emerging from it.

II. A COMPUTATIONAL MODEL FOR QUASI-STATIC CORONAL LOOPS

Loop Geometry: Our quasi-static model is a loop of high temperature plasma, much denser than the surrounding corona, confined to a magnetic flux tube. The flux tube geometry (see Fig. 1) is that of a line dipole magnetic field with its origin at a depth d below the base of the loop, e.g., the T = 30,000 K level in the solar atmosphere. The plane of the field line is assumed to be vertical with the apex of the loop a distance h vertically above the dipole origin. Line dipole field lines are circles and the parameters d and h completely determine the field configuration. A generalized coordinate s specifies the distance up the field line from the origin. The area factor Γ, defined as the area at the apex of the

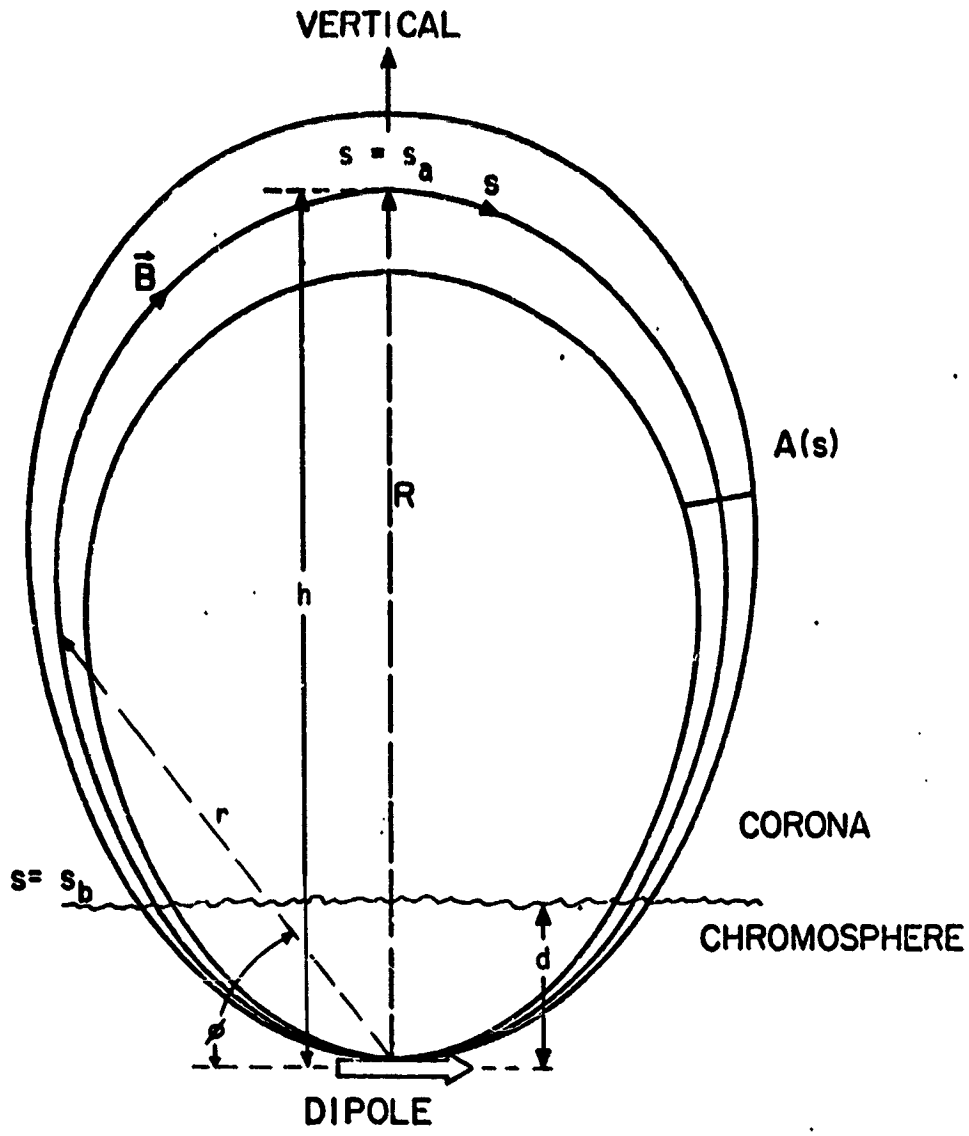


Fig. 1. Schematic diagram of line dipole magnetic field geometry. The coordinate s runs along the magnetic field line beginning at the origin, located at a depth d below the T = 30,000 K level in the solar atmosphere. The base of the loop is at $s = s_b$ and the apex at $s = s_a$. A(s) is the cross-sectional area of the loop at point s.

loop A_a divided by the area A_b at the base level, which is selected to be some distance $s = s_b$ up the loop, is given by

$$\Gamma = h/d, \tag{1}$$

the loop length by

$$L = 2h \cos^{-1}(\Gamma^{-1/2}) \tag{2}$$

Governing Equations: The equations governing the plasma loop are the momemtum and energy equations for a single fluid plasma (c.f., Boyd and Sanderson, 1969). The momentum equation describes the hydrostatic equilibrium along the loop where pressure gradient forces balance solar gravity:

$$\nabla_s P = \rho \vec{g} \cdot d\vec{s} \tag{3}$$

where p is pressure, $\vec{g}$ solar gravitational acceleration and ρ the mass density. The energy equation describes the energy balance discussed above where the energy input per unit volume ε is balanced by conduction and radiation:

$$\varepsilon = \frac{1}{A}\frac{d}{ds}(A \alpha T^{5/2} \frac{dT}{ds}) + n^2 P_{rad}(T) \tag{4}$$

where A(s) is loop cross sectional area, $\alpha T^{5/2}$ is the thermal conductivity coefficient along the magnetic field line, n is the plasma number density and $P_{rad}(T)$ is the radiative loss function (approximated by power laws over eight temperature ranges). The derivatives are ordinary rather than partial since there are no time variations and all spatial variations are along s.

Boundary Conditions: If we assume the loop to be symmetrical about the apex, then the temperature gradient at the apex must vanish to prevent conduction from one half to the other. The flexibility of a numerical model also allows one to consider non-symmetrical cases. At the loop base we assume that there is no conductive flux (dT/ds = 0) into the chromosphere at the T = 30,000 K level. If there were significant conductive flux into the chromosphere, evaporation of plasma into the loop would alter the assumed

steady state. Here again the numerical model allows one to assess the effect of conduction into the chromosphere so far as a steady state is concerned simply by changing that boundary condition.

Computational Method: To compute solutions for our loop model we consider an initial value problem governed by the momentum and energy equations (3 & 4) and subject to the above boundary conditions at the base of the loop ($s = s_b$) and a desired condition at the apex ($s = s_a$). A shooting technique is then employed to adjust the plasma number density at the base $n(s = s_b)$ such that the temperature gradient vanishes at the apex of the loop. The numerical integration of the governing equations is accomplished by reformulating eqs (3) and (4) as three first order, ordinary differential equations and numerically integrating them upward along the loop using a computer code devised by Shampine and Gordon (1975). Details are given by Vesecky, Antiochos and Underwood (1978).

The code is arranged so that any of the quantities in the governing equations (or functions of them) can be output at a prescribed set of locations along the loop. A typical set of output parameters at each location along the loop includes fractional distance along the loop, T, dT/ds, n, conductive flux F, conductive flux density, emission measure $\xi[T(s)]$ and power radiated per unit volume. The differential emission measure $\xi(T)$ is computed according to the definition of Craig and Brown (1976). For present purposes this definition yields:

$$\xi[T(s)] = A(s)[n(s)]^2/|dT/ds|. \qquad (5)$$

IV. SAMPLE RESULTS FROM THE COMPUTATIONAL MODEL

In this section we display a few sample results from the computational model described above. In Figs. 2 and 3 we consider a loop of length $L = 9.4 \times 10^9$ cm and an area factor $\Gamma = 6.7$. From eqs. (1) and (2) we see that such a loop requires a line dipole field line of height $h = 4 \times 10^9$ cm and

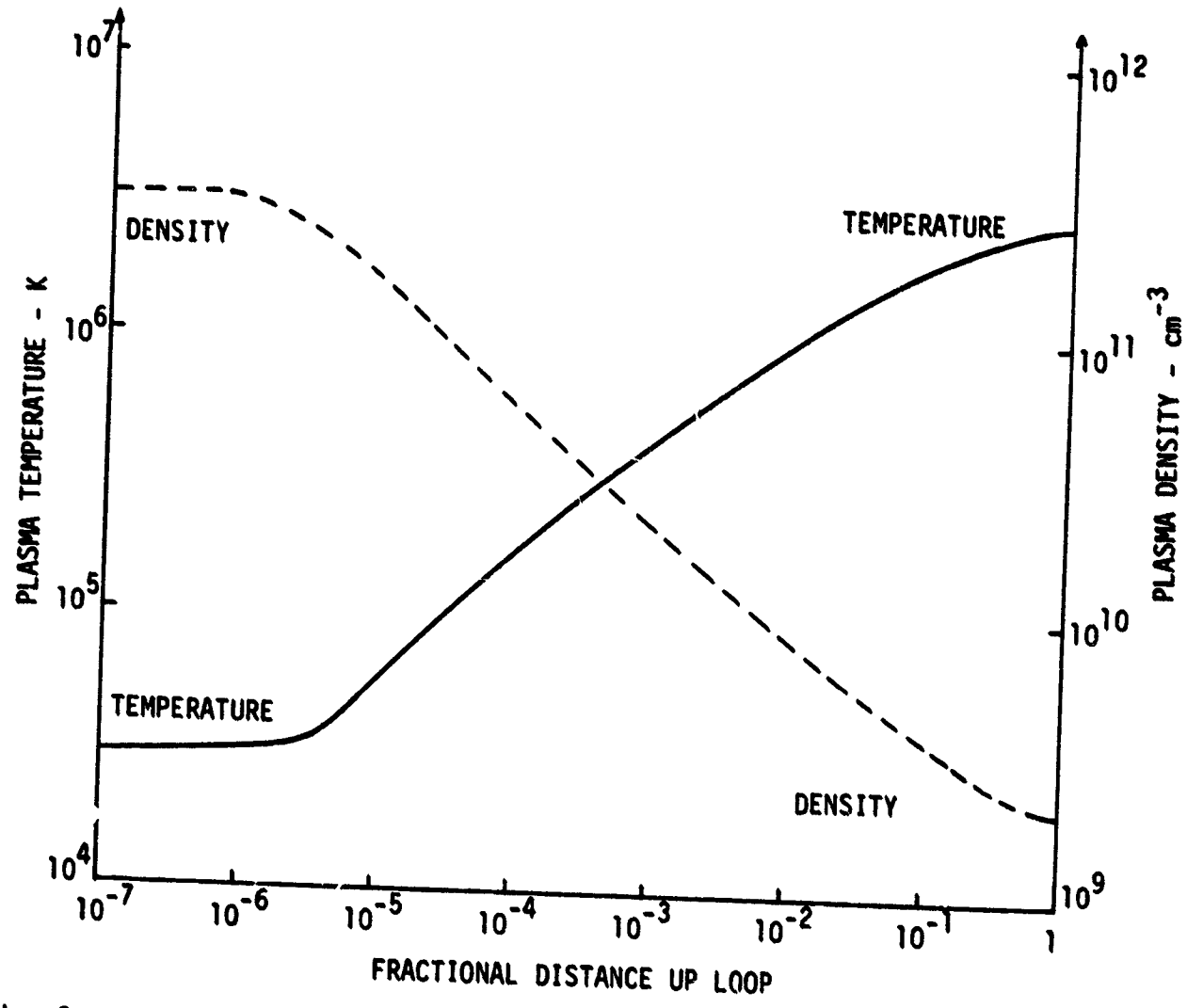


Fig. 2. Run of temperature and density vs. distance along the loop. The loop length L is 9.4×10^9 cm with an area factor $\Gamma = 6.7$. With a cross sectional area of 10^{18} cm^2 at the loop apex the emitting volume is 3.1×10^{27} cm^3. The energy input ϵ is uniform at 10^{-3} ergs s^{-1} cm^{-3}. The corresponding run of differential emission measure vs. temperature $\xi(T)$ is given in Fig. 4.

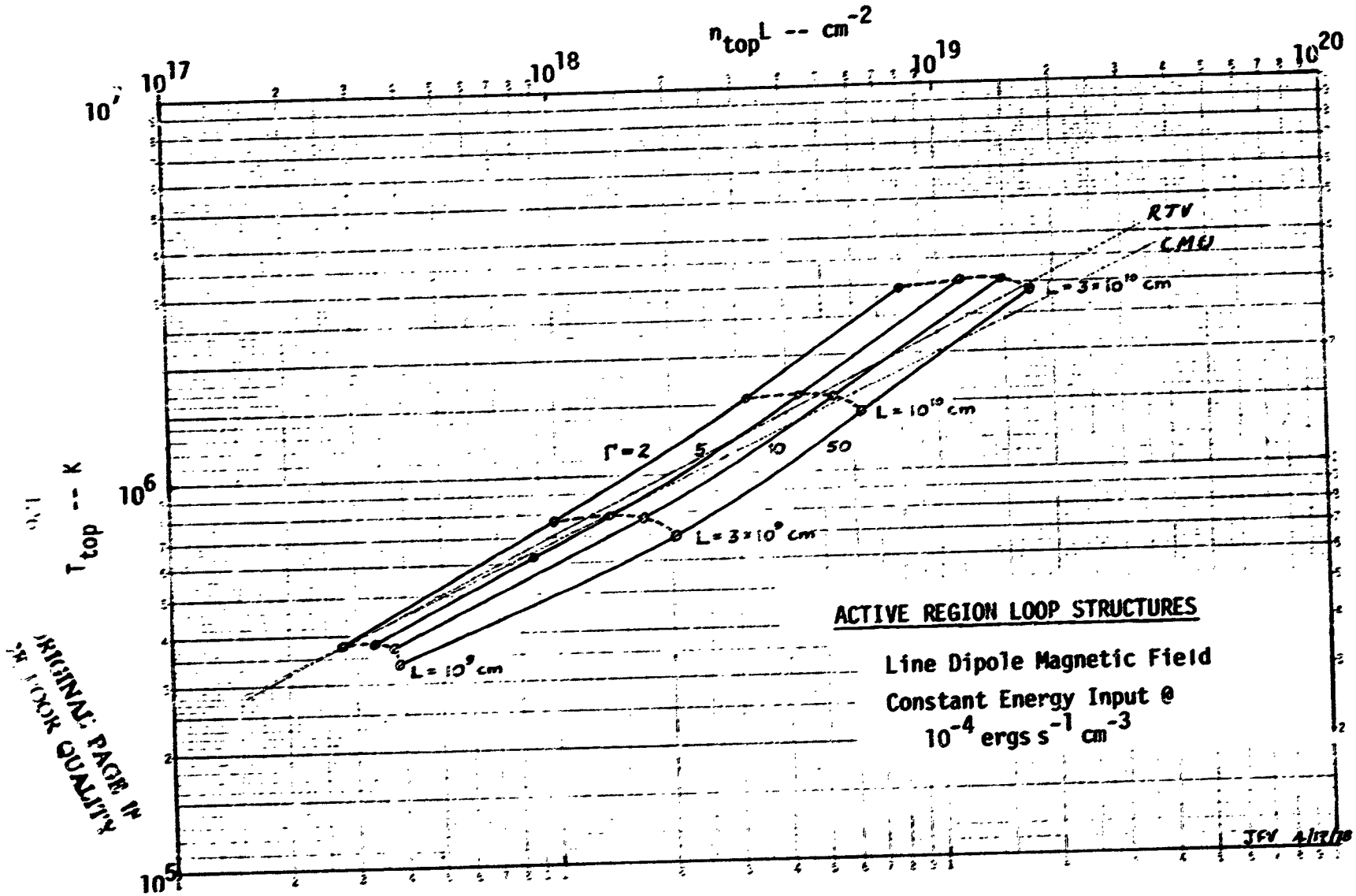


Fig. 3. Variation of loop apex temperature T_{top} with density at apex times length ($n_{top}L$) for varying values of L and area factor Γ. Also shown for comparison are analytically derived power law expressions - RTV refers to the work of Rosner, Tucker and Vaiana (1978) and CMU to Craig, McClymont and Underwood (1978).

depth d of the dipole origin 6×10^8 cm below the chromospheric 30,000 K level. Such a loop has an overall diameter of about 50 arc seconds. For an apex area $A_a = 10^{18}$ cm^2, the cross-section of the loop at its apex is about 8 arc seconds. These dimensions are typical of loops seen in the Skylab/ATM UV and X-ray images. The total volume of the loop is 3.1×10^{27} cm^3. We must also specify the form and strength of the energy imput ε. For the case considered here, we take ε to be uniform within the loop at 10^{-3} ergs s^{-1} cm^{-3}. In Figures 2 and 4 we also show cases of uniform energy input which differ by an order of magnitude.

In Figure 2 we see the variation of temperature and density with distance up the loop. We note that pressure $p = nkT$ is approximately constant along the loop with the pressure at the bottom 1.6 times greater than at the top. For the case $\varepsilon = 10^{-2}$ ergs s^{-1} cm^{-3} the pressure is more constant being only 1.3 times greater at the bottom then at the top. Comparing Figure 2 with a model for temperature and density in the quiet chromosphere and corona (Withbroe and Noyes, 1977, Fig. 3), we note that although the temperature gradient in Figure 2 reaches large values (on the order of one K cm^{-1}), the gradient is not nearly so large as that shown by Withbroe and Noyes. In addition, the curve of Figure 2 is more smooth. We also note from Figure 2 that the temperature gradient is very large at temperatures near a few 10^5 K. The temperature increases from 2 to 3 $\times 10^5$ K over a distance of $\sim 10^2$km.

Figure 3 is a convenient plot on which to illustrate some physical effects as well as make comparisons with analytic models. Here we consider a group of loops each having a uniform energy input of $\varepsilon = 10^{-4}$ ergs s^{-1} cm^{-3}. We vary the length L and area factor Γ and note the effects on the temperature and density at the apex of the loop. Thus we can assess the effects of

changing the magnetic field structure (area factor) of the loop. We see from Figure 3 that varying Γ from 2 to 50 increases the apex density by about a factor of 2 and that this change is more pronounced at small L. This effect is apparently due to material being preferentially squeezed up toward the top of the loop and hence away from where T = 1 or 2 x 10^5 K and radiation is more efficient. Hence the loop requires a denser plasma to radiate away the input energy. Interestingly, increasing Γ tends to lower the temperature at the loop apex rather than raising it. One can also compare the expressions for the variation of $n_a L$ vs T_{top} predicted by analytical models with the numerical results. Craig, McClymont and Underwood (1978) find that $T_{top} \propto (n_a L)^{0.44}$ while Rosner, Tucker and Vaiana (1978) find that $T_{top} \propto (n_a L)^{0.5}$, Both these functions are shown in Figure 3 where they are normalized at $\Gamma = 2$, $L = 10^9$ cm. From Figure 3 we find that the exponent for $(n_a L)$ varies from about 0.4 to 0.7 over the ranges $10^9 \leq L \leq 3 \times 10^{10}$ cm, $2 \leq \Gamma \leq 50$. The differences between the analytical and numerical results apparently lie in the rather severe approximations to $P_{rad}(T)$ necessary for an analytic treatment.

Underwood and McKenzie (1977) and Craig and Brown (1977) have shown that the information contained in UV and X-ray spectral measurements is related to the physical characteristics of the (optically thin) emitting plasma through the differential emission measure $\xi(T)$. We have computed $\xi(T)$ according to the definition of Craig and Brown and the results for our sample loop are shown in Figure 4. The two curves for $\xi(T)$ illustrate the effect of an increase in energy input for a given loop. It is interesting to compare the results shown in Figure 4 with analytic models. Craig, McClymont and Underwood (1978) estimate that $\xi(T)$ should vary approximately as T^{α} with $\alpha \lesssim 0.5$. We note from Figure 4 that there are

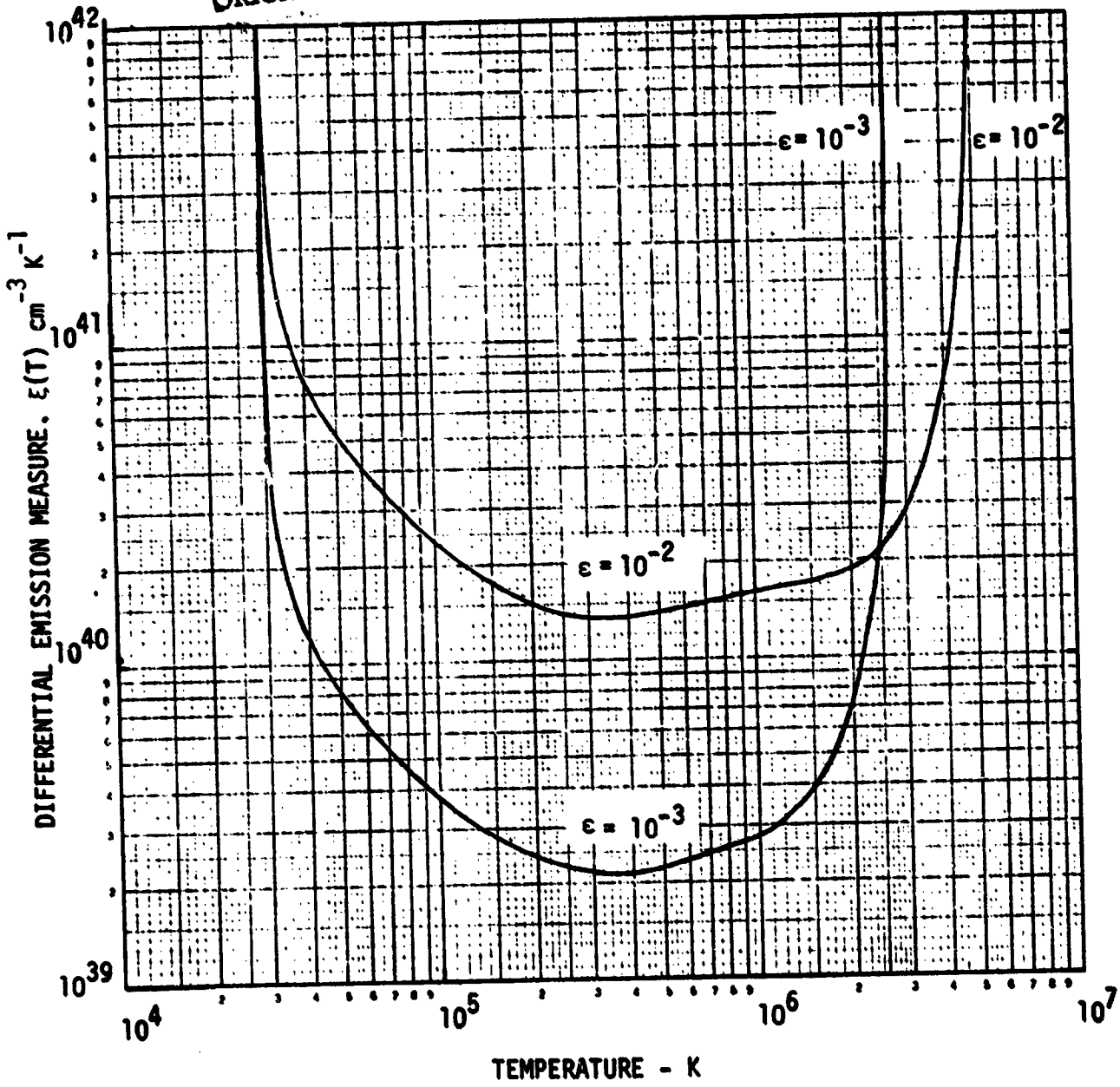


Fig. 4. Differential emission measure with respect to temperature $\xi(T)$ for loop of length 9.4×10^9 cm (see Fig. 2 for other information). Note the approximately straight (power law) portions of the curves at $4 \times 10^5 \leq T \leq 2 \times 10^6$. The curves given are for uniform energy input ε of 10^{-3} and 10^{-2} ergs s^{-1} cm^{-3}.

straight portions of the $\xi(T)$ curves running from approximately 4×10^5 to 10^6 and 4×10^5 to 2×10^6K for $\varepsilon = 10^{-3}$ and 10^{-2} respectively. The power law coefficients for these straight portions are $\alpha \approx 0.3$ and 0.14 for $\varepsilon = 10^{-3}$ and $\varepsilon = 10^{-2}$ respectively. As in our previous comparison of analytical with numerical modeling, the analytical model shows the general form of the solution, but it takes a computational model to bring out important details such as the value of α.

Since Figure 4 shows that, above $\sim 2 \times 10^5$ K, $\xi(T)$ is a monotonically increasing function of T for a single loop, we conclude that observational determinations of $\xi(T)$ which show $\xi(T)$ decreasing with T for $T \geq 10^5$K (e.g., Pye et al., 1978) are unlikely to be explained in terms of a single loop. Rather, the fall of ξ with T indicates the observation of an ensemble of loop structures within the field of view.

It is evident from Figure 4 as well as Equation 5 that $\xi(T) \to \infty$ at the base and the apex of the loop as well as at any other point where the temperature gradient vanishes. This situation is simply a result of the definition of ξ and does not correspond to an infinite energy flow. Wherever the conductive energy transport is very small, all the energy deposited is radiated away.

IV. CONCLUSIONS

The Solar Probe Mission will make it possible to obtain UV and X-ray images of the solar corona with resolution approaching $\sim$ 20 km. It is important to ask how we are to exploit such observations to better understand the coronal structures. In this paper we suggest that detailed computational modeling of coronal structures can indeed take advantage of high resolution imaging information. Further, we discuss the questions which such models can help to answer when coupled with high resolution coronal images. The answers to questions such as the nature of the heating and magnetic field geometry of coronal loops have implications

Hard X-Ray Imaging From The Solar Probe

H.S. Hudson
Physics Department, UCSD

Abstract

The Solar Probe offers a platform with several unique advantages for studying solar non-thermal plasma processes via the observation of hard X-radiation from energetic electrons in the chromosphere and corona. These include (i) high sensitivity, (ii) a second line of sight (in addition to the Earth's) that can aid in 3-dimensional reconstruction of the source distribution, and most importantly (iii) the possibility of correlation with direct measurements of the non-thermal particles from the Probe itself.

I. SOLAR X-RADIATION

Solar X-ray emission comes from many different sources in the solar atmosphere. Soft ($\lesssim$ 10 keV) X-rays constitute the natural thermal emission from the solar corona, both in active and quiet regions. Hard ($\gtrsim$ 10 keV) X-ray emission occurs during flares and their coronal consequences. We discuss the hard X-ray phenomena briefly here.

A. Impulsive Phase

The impulsive phase of a solar flare (e.g., Kane, 1974) marks an explosive energy release detectable in a variety of wavelengths, especially clearly in hard X-radiation. A burst of short duration (typically $\sim$50 sec), possibly with rapid fluctuations (as fast as $\sim$1 sec) occurs in the 10-100 keV energy range; simultaneous microwave and EUV flashes also appear. Figure 1 shows a typical example of such a burst from OSO-7 data (Datlowe et al., 1978) The hard X-ray burst appears to be the derivative of the time profile of the accompanying soft X-ray burst, also shown in Fig. 1. The gradual phase roughly matches the time profiles of the Hα flare and the gradual microwave burst.

Based upon reasonable theoretical modeling of the physical processes, since we have virtually no direct imaging observations to date, the impulsive phase probably occurs on compact closed loops. Knowledge of its association with the gradual phase, whose geometry does consist of such loops, may help to define the basic causes of the flare, or at least the nature of the non-thermal behavior involved in the particle acceleration.

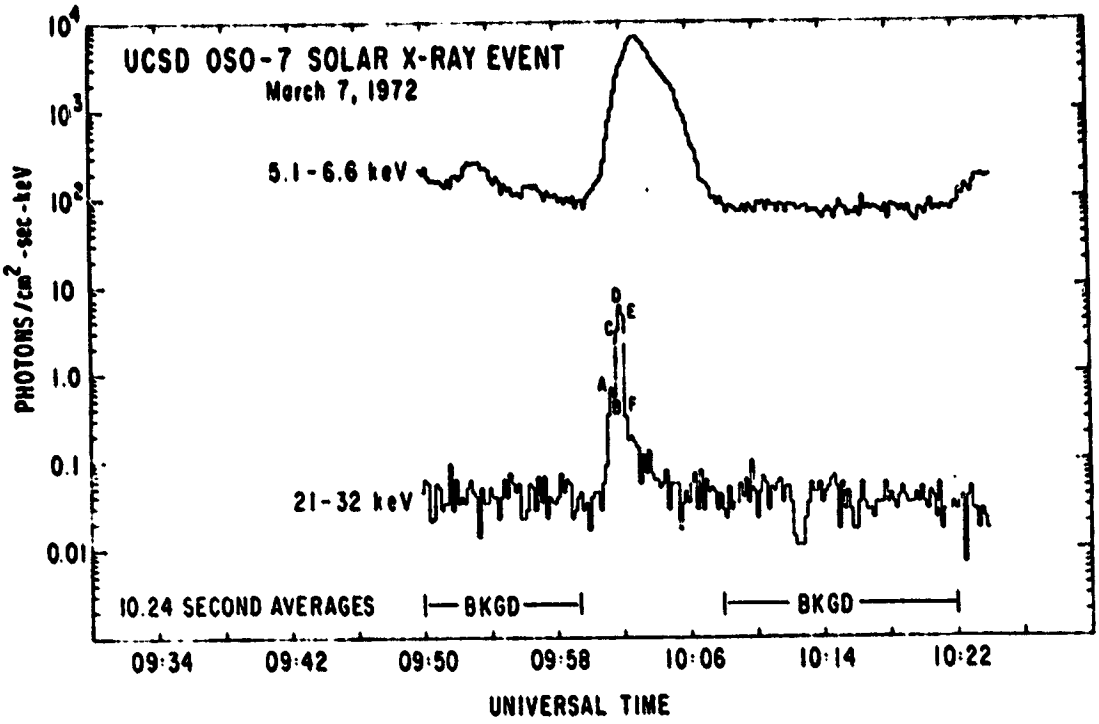


Figure 1. Typical X-ray burst observed by OSO-7 (Datlowe et al., 1978), showing the gradual time profile of soft (5.1-6.6 keV) X-radiation and impulsive time profile of hard (20-30 keV) X-radiation. This particular soft X-ray burst had a rapid rise time and a relatively strong hard X-ray burst.

B. The Coronal Phase

The meter-wave radio phenomena (e.g., Wild and Smerd, 1972) require a wide variety of populations of non-thermal elections in various configurations: unconfined streams (Type III), shock fronts (Type II), trapping regions (Type IV) at 10-100 keV energies, and Type IV and "flare continuum" bursts in several possible geometries. In all of these sources the details of the magnetic field and plasma density distributions strongly affect the propagation of the radio waves as well as their emission. To help untangle these complex sources and take substantial quantitative steps in interpreting the plasma physics, we need further diagnostic information. The Solar Probe can make these possible both by direct sampling of the distribution functions, and by sensitive imaging of the bremsstrahlung hard X-radiation from the source.

At present we have only extremely limited observational information about X-ray emission from these coronal phenomena, primarily because the more intense radiation from the chromosphere tends to swamp the weaker coronal phenomena. However, several observations of over-the-limb flares (Frost and Dennis, 1971; Hudson, 1978; Datlowe et al., 1978) have excluded these brighter sources by limb occultation. Figure 2 shows an example. The resulting purely coronal emission has several interesting properties: The hard X-ray emission has very long duration, a smooth time profile, and very flat spectra. Figure 3 plots the spectra of the three clearest examples of such events; all permit fits to $(h\nu)^{-2}$ spectra over the 10-250 keV spectral region.

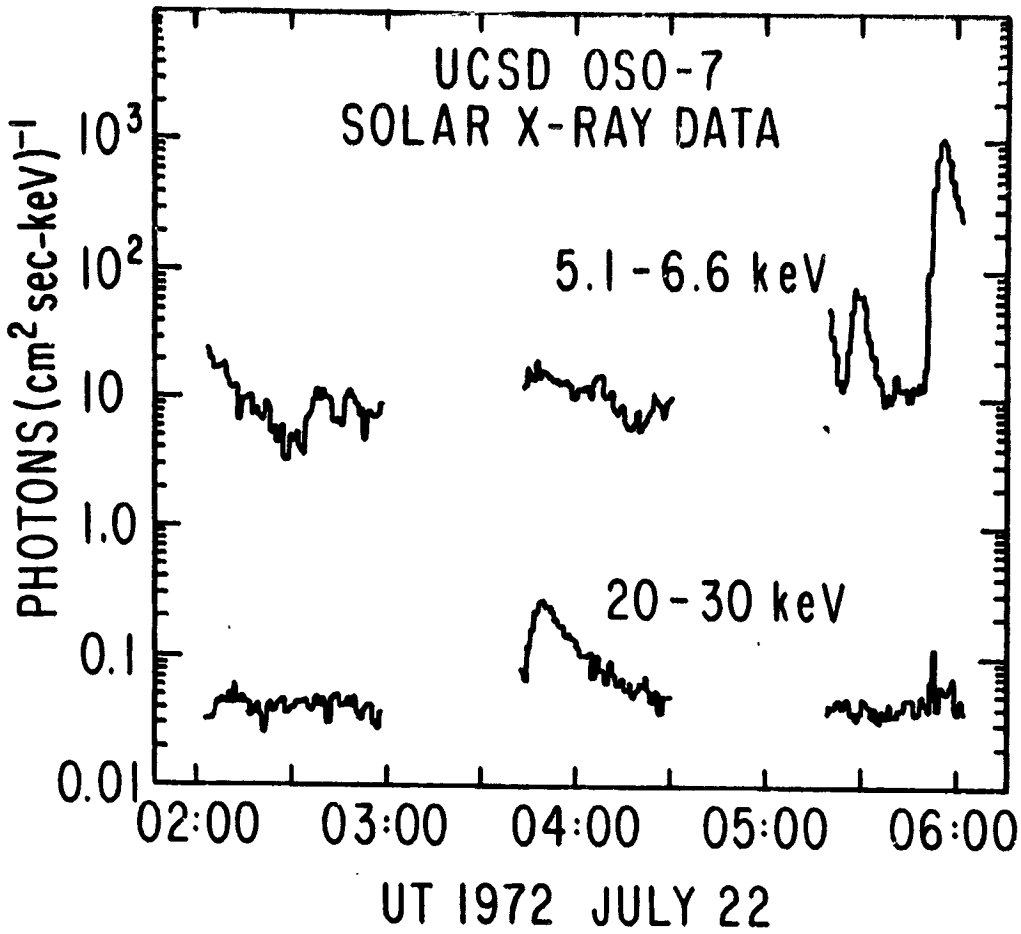


Figure 2. Purely coronal hard X-ray emission from a great flare on 1972 July 22, identified (van Hollebeke et al., 1976) with an active region ~90° beyond the solar limb. During the same time interval a more typical impulsive burst (~0540 UT) came from an active region on the visible hemisphere. The coronal event included Type II, Type III and flare continuum meter-wave emission on the appropriate limb area.

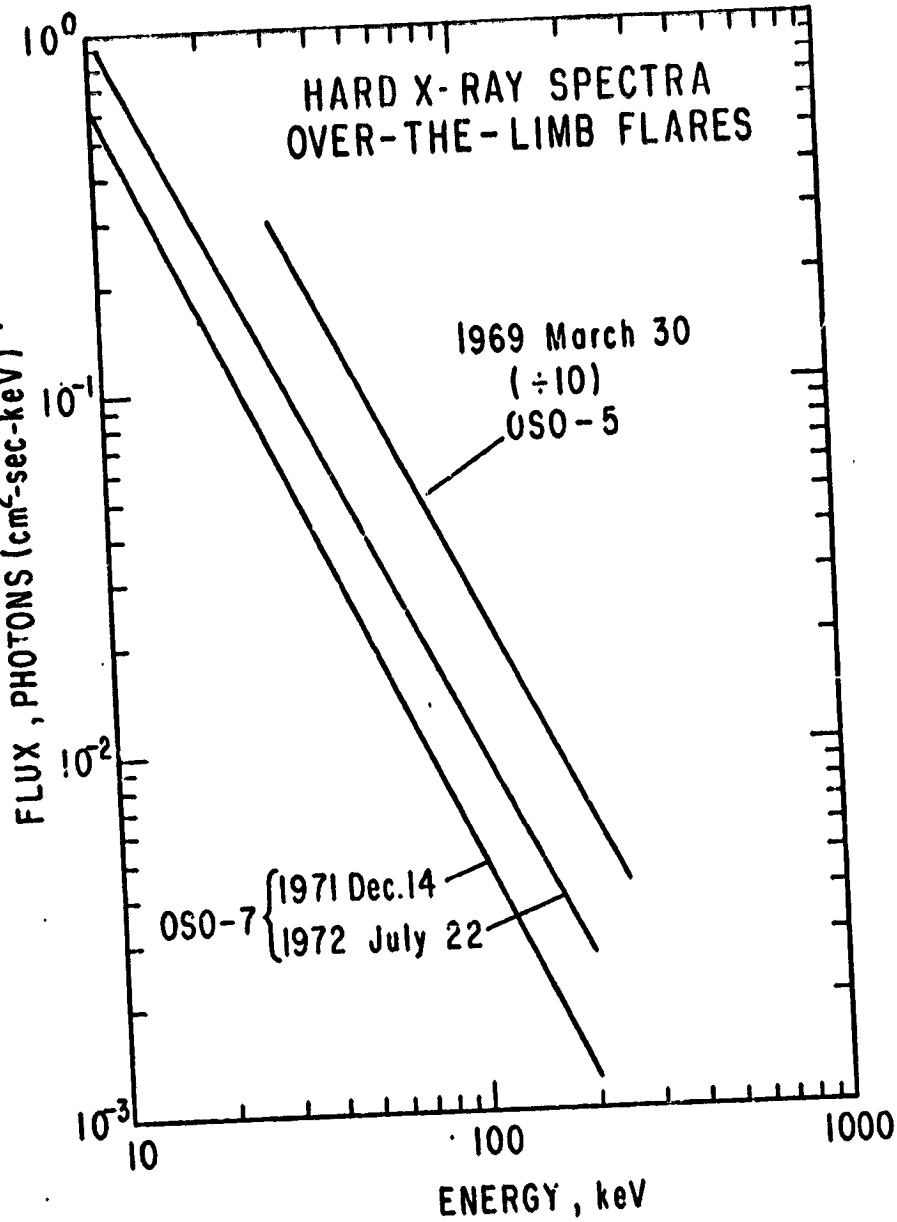


Figure 3. Hard X-ray spectra of the solar X-ray burst shown in Figure 2, along with two other members of this class of event: 1969 March 30 (Frost and Dennis, 1971) and 1971 Dec. 14 (Hudson, 1978). All had photon number spectra consistent with power-law index - 2.

The most extreme example (shown in Figure 2) occurred on 1972 July 22, from a parent flare apparently $\sim 180^{o}$ in longitude from the Earth's position (van Hollebeke et al., 1976). Confirmation of this identification comes from the Culgoora 80 MHz radioheliograph images (Stewart, personal communication), which show type II and "flare continuum" radiation across a wide sector of the 80 MHz limb. The time profile of the hard X-rays matches that of the flare continuum best. No type IV burst occurred in this event, but type II bursts occurred in each of the three examples shown in Figure 3. In any case naive estimates for the spatial extent of this hard X-ray emission suggest $\sim$ one $R_{\odot}$, simply because the 1972 July 22 event must have begun in a flare almost diametrically opposite the Earth in solar longitude.

We conclude that coronal phenomena long known from their radio-wave characteristics will also emit detectable X-ray flares. X-radiation has totally different emission mechanisms and propagation conditions. Thus sensitive imaging hard X-ray observations can greatly improve our understanding of the plasma physics in the interesting region of the solar atmosphere between the transition region and $\sim 10\ R_{\odot}$. A spacecraft close to the sun could allow us to observe this region effectively in X-radiation as well as to sample its particle populations directly.

II. Unique Advantages of the Solar Probe

A. High Sensitivity

A closer approach to the sun will generally make possible more sensitive observations, owing to the inverse-square law. This appears to

represent an extremely interesting advantage for X-ray observations, because the transient nature of the phenomena tends to limit the signal-to-noise ratio from photon counting statistics in the signal itself. The ultimate limitation, the diffuse cosmic X-ray background, lies far below the surface brightnesses of the coronal X-ray sources (Hudson, 1978).

B. Stereoscopic Viewing

The solar X-ray sources must have small optical depths. Stereoscopic viewing can therefore in principle lead to a full three-dimensional reconstruction of the source configuration. In practice, of course, this would require a large number of view directions (Bracewell and Riddle, 1969). Nevertheless two views of a complex three-dimensional source will give an enormous improvement over one perspective. The three-dimensional analysis should permit modeling the magnetic-field structures that underlie the coronal non-thermal phenomena, and their motions and growth. The second perspective in stereoscopic viewing of solar coronal phenomena will then come from observations from near-Earth spacecraft.

C. Correlated Particle Observations

Direct sampling clearly provides the greatest potential advantage of a spacecraft such as the Solar Probe for understanding coronal non-thermal particle populations. X-ray observations of the same source would greatly amplify this advantage in several ways. Firstly, the particle observations sample only along a single track, while X-ray imaging would show the

larger-scale structure of the sources. This would also aid in untangling spatial from temporal variations. The X-ray and direct particle observations can complement each other in particle species and energy range as well. Finally, the X-ray observations can provide information about the location of the flare that originates a solar-particle event. The combination of such detailed knowledge with the proximity of the particle detector to the source will greatly improve our understanding of particle propagation in the lower corona.

III. Instrumentation

A simple hard X-ray telescope with high resolution, high sensitivity and a large field of view can fit easily into the tentative configuration of the Solar Probe. It should consist of a high-Z mask (presumably tungsten) behind the first heat shield, with a position-sensitive X-ray detector near the main instrumentation. The combination of a single mask and a position-sensitive X-ray detector makes an X-ray telescope of the type conceived by Mertz (1965), Ables (1968) and Dicke (1968). A pinhole camera represents the simplest example of this type of telescope. Recent developments (Gunson and Polychronopulos, 1976) have made it possible for such telescopes to have very large effective collecting areas - the most important requirement in a signal-limited situation such as the rapidly variable solar phenomena provide.

Table 1 summarizes potential X-ray telescope properties. Absorption by the heat shield will probably limit the response to $\geq$ 30 keV. The telescope can incidentally act as a non-imaging monitor with high time

resolution for unresolved sources, and in this mode can contribute to the accurate location of both solar and cosmic burst sources by precise timing.

Table 1 roughly describes the parameters of such a system. It would have a field of view covering the whole coronal region of interest except at $\leq$ 0.1 AU near perihelion. It would contribute extremely interesting information during the whole orbit.

TABLE 1

PROPERTIES OF A HARD X-RAY TELESCOPE

FOR SOLAR PROBE

Angular resolution	3 are min
Field of view	20 degrees
Mask material	tungsten, 3 gm cm^{-2}
Mask-detector separation	1 meter
Detector resolution	1 mm
Detector area	25 cm^2
Detector type	to be selected
Total weight	5 kg

REFERENCES

Ables, J., 1968, Proc. Ast. Soc. Australia, 1, 172.

Bracewell, R., and Riddle, A.T., 1968, Astrophys. J., 150, 427.

Datlowe, D.W., Elcan, M.J., Hudson, H.S., and Peterson, L.E., 1978, in preparation

Dicke, R.H., 1968, Astrophys. J. (Lett.), 153, L101.

Frost, K.J., and Dennis, B.R., 1971, Astrophys. J., 165, 655.

Gunson, J., and Polychronopulos, B., 1976, Mon. Not. R. Astr. Soc., 177, 485

Hudson, H.S., 1978, Astrophys. J. (to be published)

Kane, S.R., 1974, in G. Newkirk (ed.), Proc. IAU Symp. 57, "Coronal Disturbances," p. 105.

Mertz, L., 1965, Transformations in Optics (New York: Wiley).

van Hollebeke, M.A.I., MaSung, L.S., and McDonald, F.B., 1975, Solar Physics, 41, 189.

Wild, J.P., and Smerd, S.F., 1972, Ann. Revs. Astron. Astrophys., 10, 159.

MULTICOLOR IMAGING AND PHOTOPOLARIMETRY: DISK FEATURES, K AND F CORONAS, ZODIACAL LIGHT

J. L. Weinberg* and R. M. MacQueen**

Abstract

The unique vantage points and viewing geometries afforded by a close-in solar probe are briefly examined with regard to line and continuum measurements of features on the solar disk, the K and F coronas, and the zodiacal light. Common instrument and observing requirements are identified, suggesting that a single instrument could provide much of the necessary observational data on these phenomena.

*Space Astronomy Laboratory
State University of New York at Albany
Executive Park East
Albany, NY 12203

**National Center for Atmospheric Research
High Altitude Observatory
P.O. Box 3000
Boulder, CO 80307

Multicolor Imaging and Photopolarimetry: Disk Features, K and F Coronas, Zodiacal Light

Introduction

Scientific interrelationships, overlapping instrument and observing requirements, and limitations on total weight and power suggest the use of multipurpose instruments, where possible, for measurements from a close-in solar probe. One such set, interrelated scientifically and/or by measurement, includes but is not limited to the following:

- Physical conditions in active and quiet regions on the sun's disk - e.g., morphology, history, and interrelationships of events on the disk, magnetic field strengths and directions in sunspots and in quiet regions, energetic particles and their propagation through the complex, changing fields of the solar atmosphere;
- Detailed 3-dimensional structure and time history of the corona - e.g., the distributions of ions and electrons, magnetic field configuration and strength, temperatures;
- Spatial distribution, optical properties, and composition of interplanetary dust - especially dust near the sun.

A brief discussion is given of some of the observations and instrument requirements associated with this set.

Observations

1. Measurements of the sun's disk.

The need for continuous high resolution imaging of the sun's disk in selected lines is well known. Of particular interest is the comparison of images obtained from 1 AU with simultaneous images obtained over a range of heliocentric distances, with emphasis near perihelion. This would make it possible to determine to what extent structure and time changes as seen in low resolution images from 1 AU reflect "real" conditions.

Broad and narrow band measurements of line and continuum polarization on the sun's disk have been discussed by Leroy (1972), Mickey and Orrall (1974),

and others. More recently, Baur and House (1977) have developed a sensitive, high spectral resolution polarimeter capable of measuring all four Stokes parameters (I,Q,U,V) for any line from 3900Å to 7000Å. Figure 1 shows a computer plot of vector magnetic fields derived from a grid of such spectral scans in the vicinity of a sunspot group. The information content in broad and narrow band polarimetry and in spectropolarimetry is significantly enhanced by "close-in" measurements. The Solar Probe could, for example, obtain high spatial resolution maps of vector magnetic fields in active regions, on the quiet disk, and in prominences - all with the 3-dimensional aspect afforded by the rapidly changing viewing geometry near perihelion. Such data are needed for a wide range of problems in solar physics.

2. K and F coronas and zodiacal light.

The K corona arises from Thomson scattering of photospheric radiation by free electrons close to the sun, and the F corona arises from scattering of photospheric radiation by dust particles close to the sun. The polarized brightness (Stokes parameter Q) of the K corona is proportional to the coronal electron density. Thus, measurements of the K corona over a range of heliocentric distances near perihelion of a close-in solar probe would make it possible to obtain 3-dimensional, high spatial resolution maps of coronal electron density. Figure 2 depicts the location in space of the particles which contribute to the K and F coronas and to the zodiacal light.

As seen from the upper portion of Figure 2, the brightness of the F corona and of the zodiacal light at angles closer than 30° to the sun comes primarily from dust particles near the sun. Thus, in principle, measurements of brightness, polarization, and color in this region can be deconvolved to <u>define</u> the spatial distribution and optical properties of dust near the sun. Unfortunately, there are relatively few observations of the F corona or of the zodiacal light closer than 30° to the sun. From the ground this effective "limit" on zodiacal light

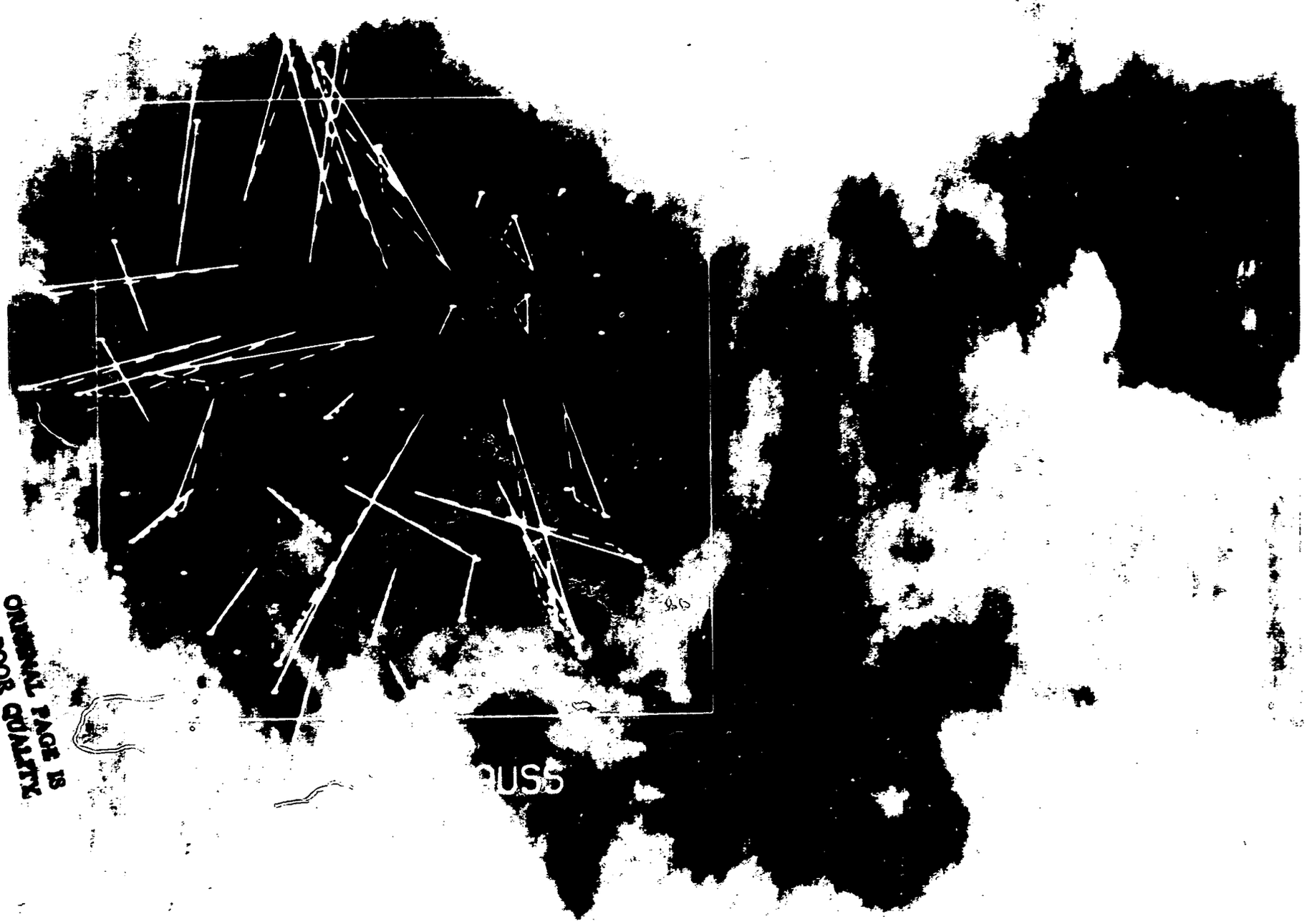


Figure 1. A computer plot of vector magnetic fields over a sunspot. The vectors have been measured from Stokes scans made at the locations of the footpoints of the vectors. (Baur and House, 1977)

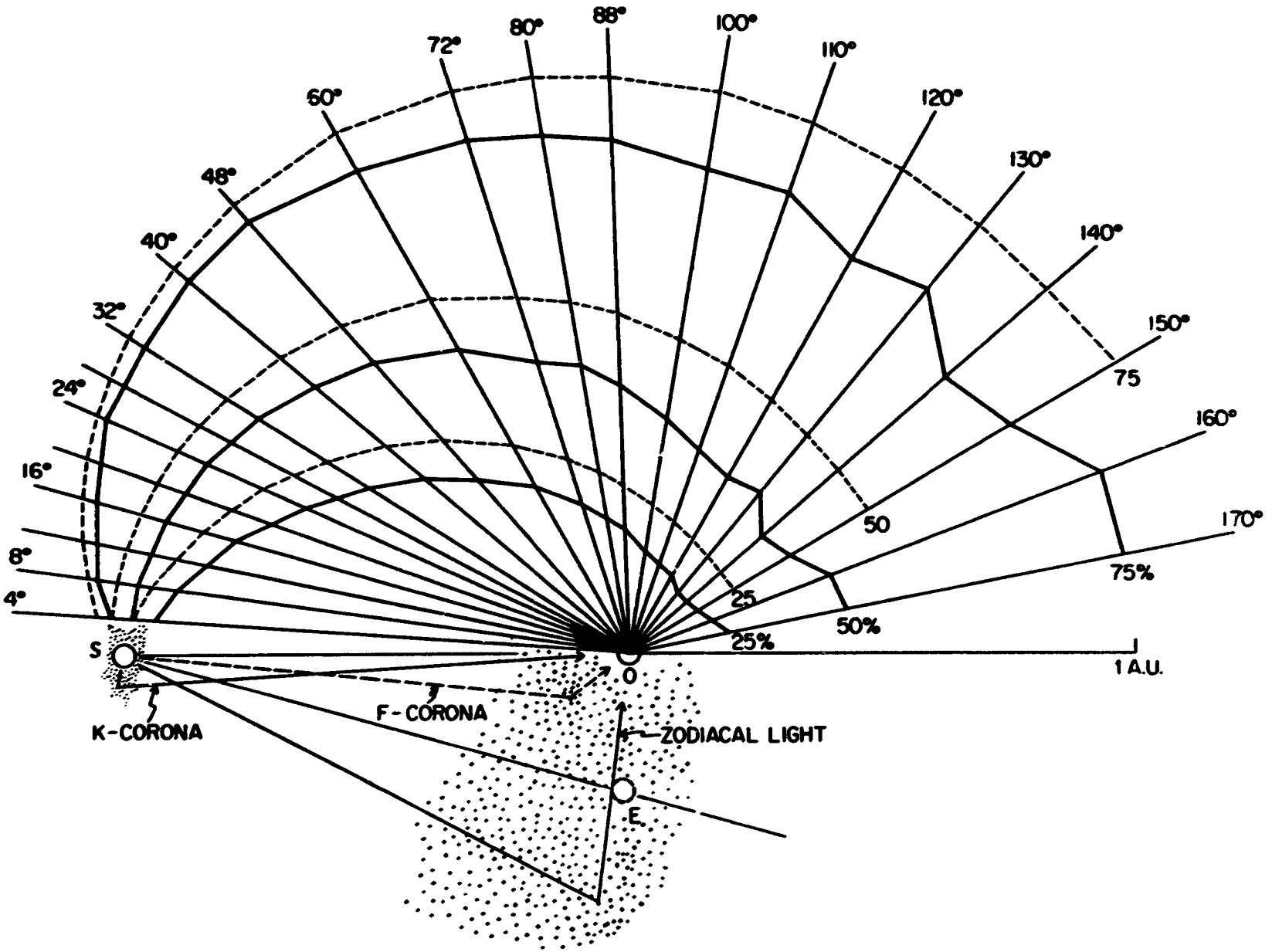


Fig. 2. The location in space of the particles which contribute to the K (electron) and F (dust) coronas and to the zodiacal light. S = sun, E = Earth, O = observer. The upper portion of the figure shows the brightness contribution along various lines of sight for a large dielectric particle model. The solid curves represent the points where the cumulative brightness is 25, 50, and 75 percent of the total brightness for an observer at 1 AU from the sun in the plane of the ecliptic (Misconi, 1977). The dashed curves are the corresponding cumulative brightnesses for an observer (spacecraft) at 1.03 AU from the sun, in a 15° inclination orbit. The lower portion (adapted from Blackwell, 1960) shows the classical explanation for the F-corona and the zodiacal light. The F-corona was formerly thought to arise from scattering by particles near the observer, and the inner zodiacal light was thought to arise from particles closer to the observer than to the sun.

observations is imposed by the need to observe without twilight interference and by uncertainties in airglow and in the corrections for atmospheric extinction and scattering near the horizon. Although there is generally no such limitation for observations from a spacecraft, other difficulties have limited most space observations to relatively large angles from the sun (Weinberg, 1976). At 30° from the sun, in observations obtained at 1 AU, there is no information about regions closer than 0.5 AU to the sun. Only observations from a close-in solar probe with a properly shielded instrument can provide the kinds of data needed to fully characterize the scattering processes and to define the spatial and size distributions and the albedo and other optical properties of dust near the sun.

The effect of the orbital inclination of the Solar Probe can be seen from the dashed curves in Figure 2 which show the location of the particles which contribute most of the brightness seen by an observer at 1 AU from the sun in a 15° inclination orbit. For this sample solar probe inclination, the Earth is shown as a dashed circle below the observer. Comparison with the 1 AU, in-ecliptic calculation (solid curves) shows that the region which contributes most of the brightness at small angles moves closer to the sun when viewed from a 15° inclination orbit. At any inclination, the general characteristics of the brightness curves are maintained for spacecraft positions at constant ecliptic latitude. Measurements from a moderate (30° to 45°) inclination orbit are, in fact, best suited to providing information on dust in the vicinity of the sun. Higher inclinations are less suitable since they provide a shorter total observing time near the sun.

Detailed measurements of both the K and F coronas near perihelion would also help to remove much of the current difficulty in separating the K and F components in observations of the solar corona during eclipse.

Other aspects associated with dust in the inner solar system:

Thermal emission from interplanetary dust has been detected by Peterson (1963, 1967), MacQueen (1968), and Léna, et al. (1974) at various positions beyond 3 $R_\odot$. Based on observations from Helios 1, no evidence was found for a decrease in dust concentration at heliocentric distances $R \geq 0.09$ AU (19 $R_\odot$) (Leinert, et al., 1978). Since the dust vaporization zone is expected to be close to 3 $R_\odot$, it should be able to be observed from a close-in (4 $R_\odot$ perihelion) solar probe if observations are made closer than 90° to the sun. The composition and size of dust particles spiraling in to the sun due to the Poynting-Robertson effect may result in the boundary of the vaporization zone being patchy.

Sky mapping observations from different heliocentric distances would make it possible to derive the spatial distribution of dust inside 1 AU and to determine possible solar wind electromagnetic effects on the orbital elements of the dust. These observations would also provide the additional data needed to examine further the possible concentration of dust toward the orbital planes of the inner planets (Misconi and Weinberg, 1978). By inverting the zodiacal light brightness integral (Schuerman, 1977; Dumont, et al., 1978), it would also be possible to search for changes in the properties or character of dust particles (size, shape, albedo) as a function of heliocentric distance, especially near the sun.

Solar storm effects (Misconi, 1976) might be seen directly in terms of changes in the brightness of the F corona/inner zodiacal light and in terms of electron depletion on the trailing edge of high speed streams (and, perhaps, in terms of dust depletion).

A comparison of close-in measurements of Fraunhofer line profiles on the disk and in the F corona/inner zodiacal light, especially east and west of the sun, could provide additional information on electrons (shallow line cores) and on dust (Doppler shifts, changes in line shape). The relative distribution of prograde and retrograde orbits obtained from measurements of Doppler shifts provides information on the cometary and interstellar contributions to the dust complex. The uncertainty in Doppler measurements made from 1 AU limits conclusions to the statement that the dust is probably in prograde orbits around the sun (James and Smeethe, 1970). Misconi (1976) has shown that ion drag resulting from impinging solar wind protons and heavy ions is much stronger for solar storms than the Poynting-Robertson drag at heliocentric distances less than 0.1 AU. This would alter the size distribution and, perhaps, the spatial density distribution of the dust, resulting in a physical separation of small and large particles and possible color effects in the brightness and polarization of the zodiacal light. Krishna Swamy (1978) suggests that a separation of small and large particles in comet tails due to radiation pressure effects could be responsible for the polarization reversal observed in the tail of Comet Ikeya Seki (1965 VIII) (Weinberg and Beeson, 1976). This same polarization reversal might be present in the F corona at heliocentric distances less than 0.1 AU and at scattering angles between 115° and 135°, given the proper observing geometry (Misconi, 1978).

Other aspects associated with the solar corona:

The Stokes parameters I and Q of coronal emission lines such as 10747Å and 10798Å of Fe XIII contain information on magnetic field directions, ion

Combining these data with observations of the K corona and with disk and limb measurements would give a more complete picture of the structure of the corona and of the interrelationships with phenomena at lower heights in the solar atmosphere.

Thoughts on instrumentation

1. Parameters to be measured.

Table 1 identifies the Stokes parameters that are generally used for the various disk and sky measurements discussed earlier.

Table 1

	Stokes parameter			
	I	Q	U	V
disk measurements				
imaging	X			
polarimetry	X	X	X	X
vector magnetic fields	X	X	X	X
coronal emission lines	X	X	X	
K corona		X		
F corona	X	X	X	
zodiacal light	X	X	X	

The Stokes parameters I, Q, U, and V are proportional to the total brightness, the polarized brightness, the orientation of the plane of polarization, and the ellipticity, respectively. In many cases, it is convenient to use the degree of polarization, p, which is defined as that fraction of the total brightness, I, which is polarized:

$$p = \frac{(Q^2 + U^2 + V^2)^{1/2}}{I} .$$

Although not often done, measurements of V would be useful on the sun's disk and in the F corona and inner zodiacal light.

2. Wavelengths.

Since color effects may be present in F corona and zodiacal light measurements at small heliocentric distances, bandwidths should be kept relatively narrow (10's to 100's of Å's, not thousands). Several wavelengths are needed here, relatively far apart, such that 1/wavelength changes appreciably. Short wavelengths are important as a probe of the size of the particles, especially near the vaporization zone, and for possible detection of a component of interstellar grains. Long wavelengths are important for detection of thermal emission, including possible signatures characteristic of particle composition. Several narrow wavelength bands are also required for studies of line emission on the disk or in the corona. For vector magnetic fields it is important to be able to scan over a broad range of wavelength so as to measure selected lines.

3. Pointing.

All aspects of the measurement set discussed here require some form of pointing. An articulated mirror system is suitable for disk and limb measurements and could perhaps be coupled to the sky measurement modes. One way to obtain the necessary sky coverage for the F corona and zodiacal light so as to be able to invert these data is to observe the antisolar hemisphere in a Pioneer 10/11-type, "spin-scan" configuration (Weinberg, et al., 1974) as the spacecraft "backs" toward the sun and to subtract hemispheres to get the "local" contribution. Stray light is minimal for this mode if observations are restricted to angles from the sun (elongations) greater than 90°. Discrete and background starlight become less and less important as the spacecraft nears perihelion, and the total signal is essentially just zodiacal light and F corona. To extend this to smaller elongations would require careful stray light rejection and automatic gain changing, but it would allow the K and F coronas to be observed from "close-up". To get the spatial resolution needed at both large and small heliocentric distances, large and small fields of view and a corresponding change in scanning step size would be needed. For large heliocentric distance R, a field of view of 2° or 3° (diameter) is needed; at small R, a field of view smaller than 1° is needed. For disk, limb, and K coronal observations, variable fields of view would be required.

Experience with optical instrumentation developed or in process for other probes such as Pioneer 10/11, Voyager 1/2, Helios 1/2, and the Solar Polar Mission suggests that an articulated instrument suitable for such multipurpose measurements might require a mass of 7 to 12 kg, a power of 4 to 8 watts, and a variable data rate ranging from 5 to several hundred bits per second.

Concluding Remarks

The measurements discussed here contain data on interrelationships and detailed morphology of active regions and surface fine structure, photospheric and coronal magnetic fields, changes in the number and distribution of coronal electrons, and the spatial density distribution, optical properties and composition of interplanetary dust, especially dust near the sun. The number of common instrument and observing requirements suggests that a single instrument could be developed that would provide much of the necessary data on all of these phenomena.

Acknowledgements

We thank our colleagues D. W. Schuerman and H. Tanabe for comments concerning observations from a close-in solar probe. N. Y. Misconi kindly provided model calculations of the brightness contributions of zodiacal light expected for a typical solar probe trajectory and suggested a number of areas of special interest for such a mission.

References

Baur, T. G. and House, L. L. (1977). A spectrum scanning Stokes polarimeter, SPIE, 112, Optical Polarimetry, 209-217.

Blackwell, D. E. (1960). The zodiacal light, Sci. Am., 203, 54-63.

Dumont, R., Rapaport, M., Schuerman, D. W., and Levasseur-Regourd, A. C. (1978). Inversion of the zodiacal brightness integral for an out-of-ecliptic photometer, pres. at XXIst meeting of COSPAR, Innsbruck, June.

James, J. F. and Smeethe, M. J. (1970). Motion of the interplanetary dust cloud, Nature, 227, 588-589.

Krishna Swamy, K. S. (1978). On the observed polarization of Comet Ikeya-Seki (1965 VIII), Astron. and Astrophys., in press.

Leinert, C., Hanner, M., Link, H., and Pitz, E. (1978). Search for a dust free zone around the sun from the Helios 1 solar probe, Astron. and Astrophys., in press.

Léna, P., Viala, Y., Hall, D., and Soufflot, A. (1974). The thermal emission of the dust corona during the eclipse of June 30, 1973, Astron. and Astrophys., 37, 81-86.

Leroy, J.-L. (1972). Nouvelles mesures de la polarisation de la lumière au bord du disque solaire, Astron. and Astrophys., 19, 287-292.

MacQueen, R. M. (1968). Infrared observations of the outer solar corona, Astrophys. J., 154, 1059-1076.

Mickey, D. L. and Orrall, F. Q. (1974). Polarization measurements on the sun's disk, in Planets, Stars and Nebulae studied with photopolarimetry (T. Gehrels, ed.), 686-694, Tucson: University of Arizona Press.

Misconi, N. Y. (1976). Solar flare effects on the zodiacal light?, Astron. and Astrophys., 51, 357-365.

Misconi, N. Y. (1977). On the photometric axis of the zodiacal light, Astron. and Astrophys., 61, 497-504.

Misconi, N. Y. (1978). Personal communication.

Misconi, N. Y. and Weinberg, J. L. (1978). Is Venus concentrating interplanetary dust toward its orbital plane? Science, in press.

Peterson, A. W. (1963). Thermal radiation from interplanetary dust, Astrophys. J., 138, 1218-1230.

References, continued

Peterson, A. W. (1967). Experimental detection of thermal radiation from interplanetary dust, Astrophys. J., 148, L37-L39.

Querfeld, C. W. (1977). A near-infrared coronal emission-line polarimeter, SPIE, 112, Optical Polarimetry, 200-208.

Schuerman, D. W. (1977). The general inversion of the zodiacal light brightness integral, Bull. AAS, 9, 564-565.

Weinberg, J. L. (1976). Space observations of the zodiacal light, in Proc. IAU Colloq. 31, Interplanetary Dust and Zodiacal Light, Lecture Notes in Physics 48 (H. Elsässer and H. Fechtig, eds.), 41-44, Heidelberg: Springer-Verlag.

Weinberg, J. L. and Beeson, D. E. (1976). Polarization reversal in the tail of Comet Ikeya-Seki (1965 VIII), Astron. and Astrophys., 48, 151-153.

Weinberg, J. L., Hanner, M. S., Beeson, D. E., DeShields, L. M., and Green, B. A. (1974). Background starlight observed from Pioneer 10, J. Geophys. Res., 79, 3665-3670.

D15

N78-32979

MAGNETIC FIELD MEASUREMENTS FROM THE SOLAR PROBE

A. M. Title
Lockheed Palo Alto Research Laboratory

During the past ten years a growing body of indirect evidence has indicated that most of the magnetic flux on the solar surface is surprisingly strong. Recent direct evidence (measurement of the Zeeman splitting) has indicated that virtually all the measured fields have a strength of 1200 ± 100 Gauss. At present it appears, at least to me, that there are two critical questions about the solar magnetic field that can only be answered from the Solar Probe. The first is, does weak field exist? The second is, what is the detailed structure of the prototypical strong field flux tube?

To answer these questions a great deal of work has been done on the ground; however, it is now clear that because of its displacement of the image and its degradation of the resolution, atmospheric seeing has placed a limit on progress. In order to get better data NASA is supporting a magnetograph system with 1/2 arc second resolution on Spacelab 2, and work is at least in the preliminary stage for a Solar Optical Telescope (SOT) Facility magnetograph with 0.1 arc second resolution. In earth orbit, Spacelab 2 and SOT telescopes can resolve 200 km and 80 km magnetic fields, respectively.

It is clear that the Solar Probe has the potential for substantial improvements in spatial resolution. A modest 12.5 cm telescope on the Solar Probe will resolve 40 km at 10 solar radii and 16 km at 4 radii.

The question then arises is this factor of 2 to 5 in resolution necessary? First, it is important to point out that when a telescope

builder describes his or her telescope the resolution quoted is for detection of 100% contrast resolution targets. Thus, if we desire to measure the vector field in an elementary magnetic structure that has a spatial scale of 70 to 100 km and is semi transparent because the mean free path for photons is on the order of the scale structure, a telescope resolution of 16 to 40 km is necessary.

It appears then that Solar Probe offers the opportunity to really measure the magnetic structure of the elementary flux tube which we now have reason to believe has a scale on the order of 100 km. Further, the rapidly changing aspect of the probe allows us to obtain the vector field geometry. It also offers the possibility of looking for "low contrast" weak field structure in this size range. If the orbit of the probe can be made to pass over the pole, such data can be gained as a function of latitude, and the structure of the polar field can be measured directly for the first time.

Finally, can we build a magnetograph that can carry out these measurements? I believe the answer is yes. Shown in Figure 1 is a drawing of a telescope that obtains line profiles with 50 ma resolution in circular and linear polarization. It can operate in as many lines as it has pre-filters. In addition it may obtain velocity information on the way into the Sun to gain data on the solar oscillations.

Figure 2 shows the detector plane of the system. Falling on the detector is a telecentric image of the objective. Since there is an aperture in the focal plane that can both move and change size, polarized line profiles can be obtained at arbitrary points with variable spatial resolution. I believe such a system can be made smart, so that the telescope microprocessor conducts the search for the magnetic elements and then constructs the vector field.

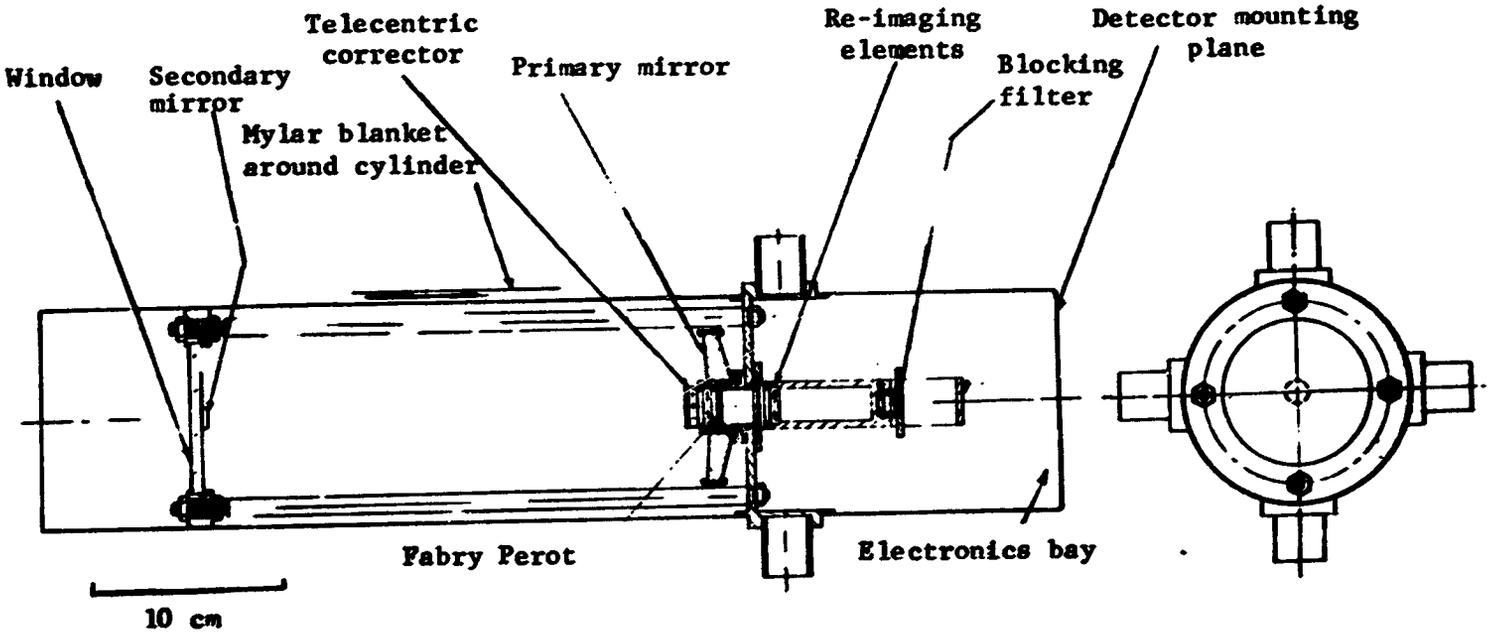


Figure 1. Layout of the optical system.

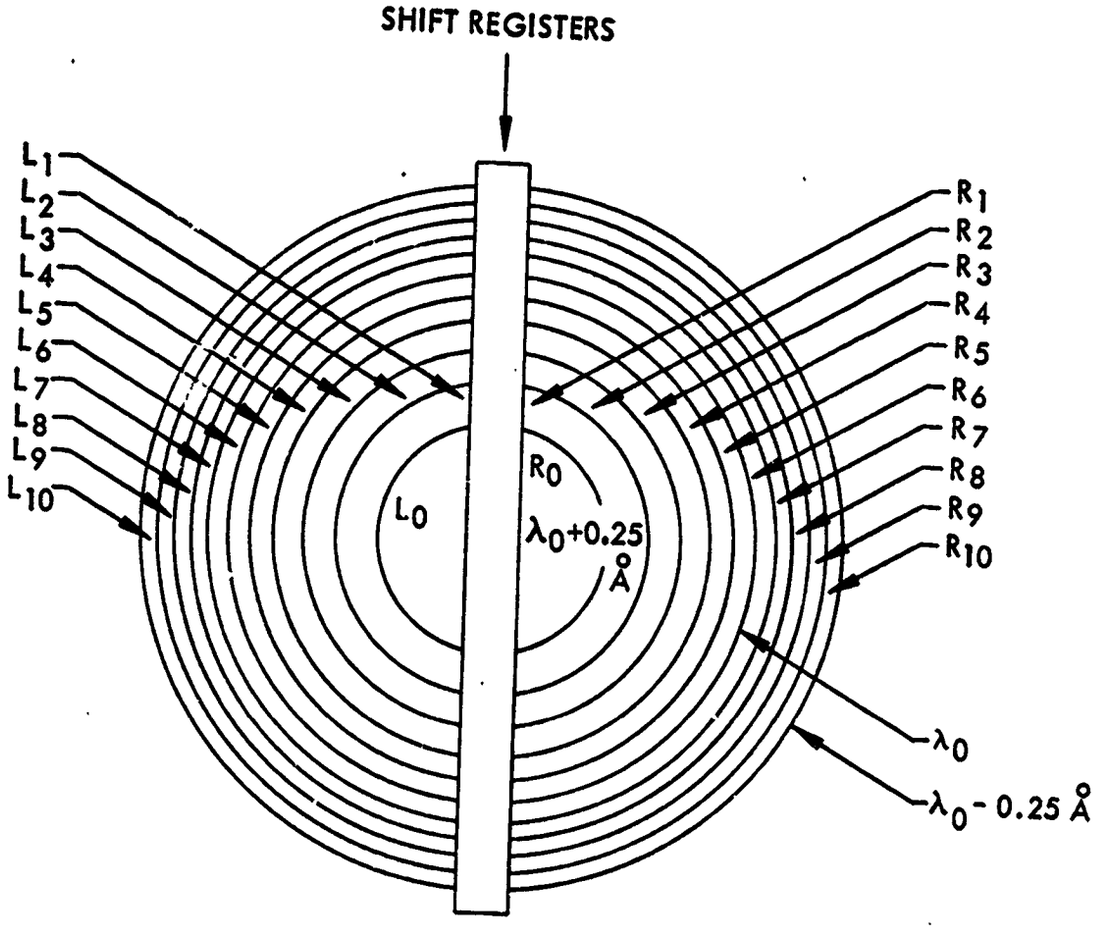


Figure 2. Detector plane, ring L1 to L10 and R1 to R10 are covered respectively by left and right circular polarizers. Rings R0 and L0 are not illuminated.

A NEW USE OF HIGH RESOLUTION MAGNETOGRAMS

P. J. Baum and A. Bratenahl

Institute of Geophysics and Planetary Physics
University of California, Riverside, California 92521

Abstract

Ground-based solar magnetograms are frequently in error by as much as twenty percent. This imprecision is part of the reason for the poor correlation between magnetic changes and solar flares since flare related changes can be much less than twenty percent. High resolution measurement of the magnetic field component $B_{\parallel}$ which is normal to the photosphere and measured at photospheric height can be used to construct a magnetic flux partition function F. We show that dF/dt is an EMF which drives atmospheric currents in reconnecting solar active regions. With a high quality magnetograph the solar probe can be used to obtain good estimates of F and dF/dt and thereby the energy stored as induced solar atmospheric currents during quiescent interflare periods. Should a flare occur during a favorable observing period, the present method of analysis should show characteristic signatures in F, dF/dt, and especially, in the stored flux $\Delta\phi$ computed from dF/dt.

1. INTRODUCTION

It has been widely conjectured that solar flares are energized by the magnetic energy stored in complex active regions. Paradoxically, however, in attempting to show that magnetic changes cause or characterize flares, solar magnetic observations have produced equivocal results. In previous attempts at resolving the paradox, it has been contended that magnetic measurements are imprecise or that magnetic theories of flares are incorrect. We present an alternative explanation: the present use of magnetograms to examine active region structure through numerical integration of miscellaneous field lines (under various force-free assumptions) provides qualitative information only and does not utilize the quantitative information available. Therefore, we propose a new approach to the analysis of magnetograms which is illustrated with a highly symmetrized example that permits integration in closed form. The proposed approach exploits the cellular structure of the flux of field lines present in a complex active region. The various topological connectivities distinguish parent and daughter flux cells. A function F is developed expressing the flux partitioned into the daughter cell of interconnected field lines in a potential field. This F is a function of the location, strength, and relative motions of the photospheric sources. Then dF/dt is used as an EMF in the direct calculation of the stored magnetic energy available for flare production. In carrying out this program the flux partitioning surface (separatrix) is calculated along with its line of self-intersection (separator). The separator is the location of the principal energy release site.

In this presentation we will briefly review the relation of flares to magnetic fields, describe how solar magnetic fields have been interpreted, and discuss the relation between the topology of individual field lines and the cellular topology of flux surfaces. With these rudiments in mind, we shall proceed to calculate the cellular fluxes of a complex bipolar spot pair as a function of spotgroup separation. The calculation is done for the limiting case where the spot pairs are colinear and will make use of the magnetic point charge model. Having demonstrated exact solutions for the cellular fluxes in this special case, it is suggested that solar magnetic field codes be further developed to compute the cellular fluxes in more general arrangements. With such a code available and if a flare should be observed by the solar probe,

we expect that significant conclusions could be drawn from measurement of even the line of sight photospheric magnetic field ($B_{||}$).

1.1 Flares and Magnetic Energy Changes

There are a number of observations indicating a decrease of magnetic energy during large flares (Evans, 1959; Howard and Severny, 1963; Gopasyuk et al., 1963; Zvereva and Severny, 1970) and a later recovery to near its previous interflare value, but Rust (1976) cites many examples where photospheric fields seem not to have changed when flares occurred. He cautions, however, that most sunspot field measurements are typically in error by 10-20% or 500 Gauss which could easily mask expected changes. Thus, it appears that some advances are required in observational procedures and/or interpretation before unequivocal flare related magnetic field changes can be firmly established.

1.2 Magnetic Topology and Magnetic Flux

There exists considerable evidence that major flares tend to occur in magnetically complex regions (Svestka, 1976) similar to those fields generated by groups of magnetically bipolar pairs. Since the motion of individual field lines is usually not defined uniquely, it seems more natural to study the flux of field lines (ϕ) since its time changes ($d\phi/dt$) are simply voltages which are unique and, in principle, measureable. Accordingly, we have studied the field of bipolar spot pairs not only from the point of view of discrete field lines, but have shown the cellular nature of flux in this arrangement. A perspective view of the three-dimensional flux cells is drawn in Figure 1 to show the flux topology of two bipolar spot pairs in the plane of the photosphere. Cells 1 and 2 are termed parent cells while 3 and 3' are the two parts of the daughter cell. The intersection of the separatrix with the photosphere is shown as the darkest pair of closed oval lines. The separatrix is the surface separating the parent and daughter flux. The separatrix intersects itself along an arched curve above the photosphere called the separator (not shown) which is the field line joining the points "a" and "b". The points "a" and "b" are true x-type neutral points on the photosphere but there is a non-zero field component everywhere else along the separator. The separator is a generalization of what is a neutral line ($\vec{B} \equiv 0$) in two dimensions. The lighter lines are individual photospheric field lines linking the north (N) and south (S) spots.

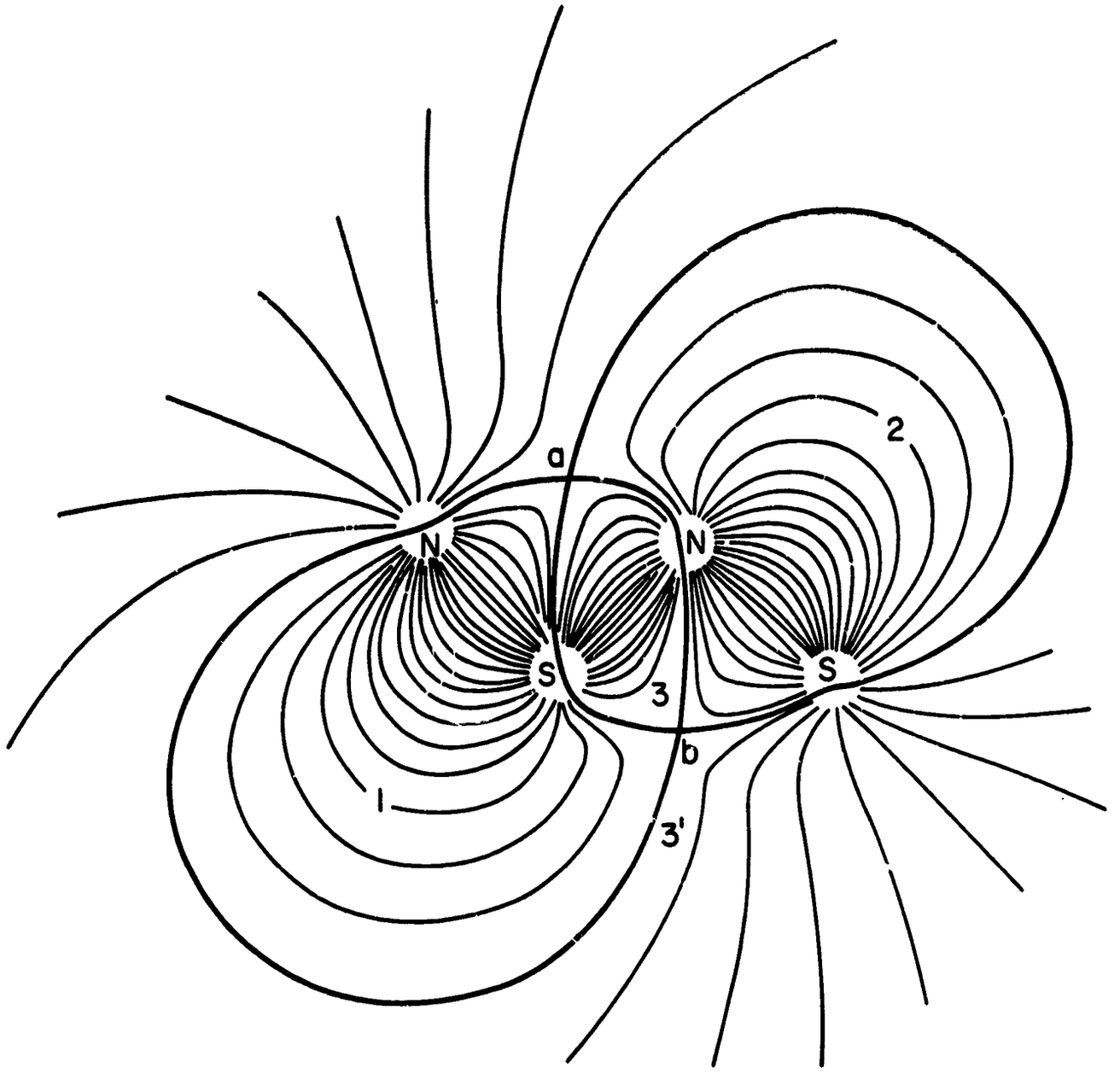


Figure 1. The magnetic topology of two bipolar spot pairs arbitrarily arranged. The darkest curves are the separatrix at the photosphere and the lighter lines are discrete field lines. The separatrix divides the flux into cells 1, 2, 3, 3' in which lines have differing connectivity. Points "a" and "b" are x-type neutral points ($\vec{B} \equiv 0$) and are the foot points of the separator field line.

From Faraday's law, the induced voltage V_S along the separator measures the rate of flux transfer from parent to daughter cells.

With solar magnetic codes (e.g. Schmidt, 1964), it is common practice to compute the total flux and flux disbalance of a solar region and to develop contour maps of the line of sight magnetic field ($B_{\parallel}$). However, these analyses do not compute the flux within the various cells defined by the separatrix since the separatrix itself is not computed. It should be apparent from the separatrix shown in Figure 1 as well as from the pioneering work by Sweet (1958) that the separatrix passes through sunspots so that the flux of any given spot involves more than one connectivity. Hence, in determining flux transfers during flares, it is necessary to measure separately those fluxes which reside in parent and daughter cells. Further, it seems entirely reasonable as a first step to know the cell flux content in the potential case which represents the energy ground state of the spot fields. Therefore, in this paper we will derive fields, fluxes, and the separator height for colinear spot pairs using the point charge analogy. While spot pairs are rarely arranged in a colinear fashion, it has occurred, for example, during a flare of September 7, 1973 (Wu and Smith, 1977). Further, preliminary observations at this laboratory (Baum et al., 1976) indicate that magnetoplasma processes above non-colinear bipolar solenoid pairs are quite similar to those above colinear pairs. Hence, it is appropriate to study the colinear problem as a guide for further interpretation and experimentation.

2. APPROACH

Figure 2 shows the magnetic topology of the system under study. We consider two bipolar pairs of sunspots arranged in the order NSNS along the Z-axis. The outermost spot pairs are separated by the distance Z_0. The separatrix, shown as the heavy solid curves in Figure 2, appears to divide the flux into four cells labelled 1-4. If the spots are also labelled 1-4 as in Figure 2, then cell 1 contains flux linking spots 1 and 2, cell-2 flux links spots 3 and 4, cell-3 flux links spots 2 and 3, and cell-4 flux links spots 1 and 4. Note that the flux in cells 3 and 4 represents interconnected (perhaps reconnected) flux if spots 1-2 and 3-4 are "parental pairs" (Bratenahl and Baum, 1976).

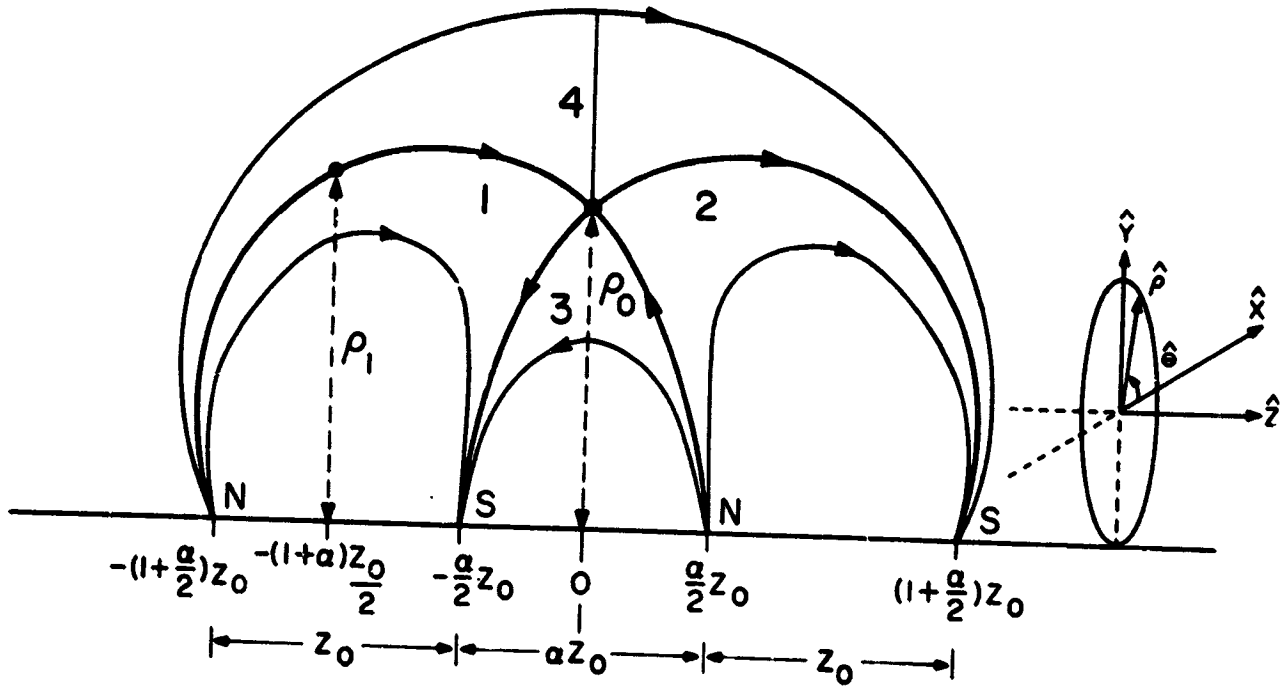


Figure 2. Four point sources of magnetic flux are assumed to be colinearly aligned with fields in the order NSNS. The overall topology is schematically shown with the separatrix as the darker lines. The separatrix defines flux cells which are labelled 1 through 4. The configuration is cylindrically symmetric about the z axis with the separator (line of neutral points) located along z = 0. The distance ρ_1 from $\rho = 0$ to the separatrix in cell 1 is used to compute the flux in cell 1. The spots N, S, N, S are labelled 1, 2, 3, 4 from left to right.

The results of straight-forward calculations are shown in Figure 3 where the relative heights of the separator (ρ_0) and the separatrix (ρ_1) and the relative fluxes in cells 1 and 3 are plotted as functions of α. Given observations of the geometric factors α and Z_0 and of the sunspot flux $\phi_0 = \int B_{\parallel} dA$, where $\int dA$ is the sunspot area, we are able to compute the partition function $F \equiv \phi_3 = \phi_4 = \phi_0 - \phi_1 = \phi_0 - \phi_2$ for the relatively complex geometry of Figure 2. It remains a challenge to compute fluxes and partition functions for a more general configuration such as that in Figure 1.

3. DISCUSSION AND CONCLUSIONS

We now attempt to show how a knowledge of the cellular fluxes in the potential limit may be used to estimate the flare energy stored in the presence of conducting plasma. If this is successful, then flare energies can be determined from a knowledge of the photospheric field component $B_{\parallel}$ alone.

Storage of flux $\Delta\phi$ and free magnetic energy ΔU_m to drive the flare process can come about through (i) a relative displacement of the spots (temporal change in α and/or Z_0), (ii) a change in spot strength (temporal change in ϕ_0), or (iii) spot group rotations (not analyzed here). In the absence of a conducting plasma, changes (i) and (ii) simply change the flux partition function F according to Figure 3b. However, in the presence of plasma, as we shall presently see, these changes in F cannot take place pari pasu but are delayed. Flux $\Delta\phi$ and free energy ΔU_m are therefore stored, and we can use Figure 3b as a basis for determining this flux and energy storage. Baum et al. (1978a) showed that the energy and flux change during a large flare could correspond to only a few percent of the total energy and flux of the field threading a bipolar pair.

In Case (i) examination of Figure 3b shows that a few percent change in ϕ_1 or ϕ_3 requires only a small change in α (a small displacement). For Case (ii) with constant α, ϕ_1 and ϕ_3 are directly proportional to the change in spot flux ϕ_0 (e.g. a 1% change in ϕ_0 produces a 1% change in ϕ_1 and ϕ_3). The small magnitude of the changes and the spatial resolution required make observations difficult.

On the basis of laboratory studies (Bratenahl and Baum, 1976), we have acquired some knowledge of plasma effects in reconnecting systems. The presence of plasma controls the rate of flux transfer into the daughter cell. Thus with plasma present, Faraday's law, Ohm's law, and the relation $\Delta\phi = LI$ give

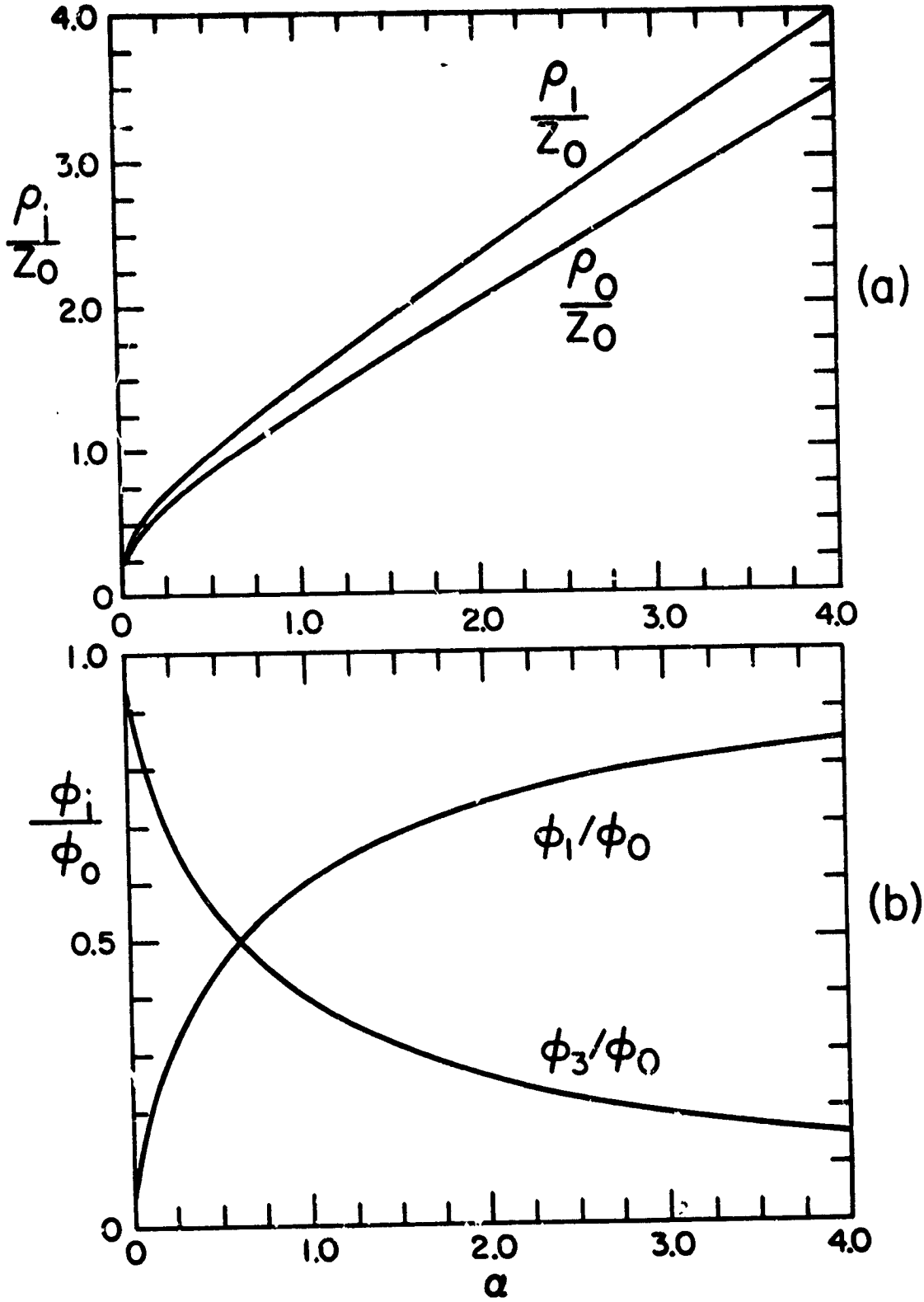


Figure 3. (a) Normalized separatrix heights calculated from the point charge model are plotted as a function of α. α is shown in Figure 2. The lower curve shows the separator as a function of α whereas the upper curve shows the height of the separatrix midway between spots 1 and 2. (b) The normalized calculated cellular fluxes are plotted as a function of α. ϕ_0 is the flux in the half space ($y \geq 0$) from a single spot. ϕ_1 is the flux in cell 1 and ϕ_3 is the flux in cell 3.

$$-\frac{d\phi_3}{dt} = IR = \frac{R}{L}\Delta\phi = V_S \tag{1}$$

Here I is the current induced along the separator, R is the effective resistance, L is the inductance associated with the current path, and $\Delta\phi$ is the stored flux associated with I (not to be confused with the background flux of the active region). On the other hand, with no plasma present the rate of change of flux in cell 3 is dF/dt which is greater than $d\phi_3/dt$ by the amount $d(\Delta\phi)/dt$:

$$\frac{dF}{dt} - \frac{d\phi_3}{dt} = \frac{d(\Delta\phi)}{dt} \tag{2}$$

Hence, the equation for flux storage $\Delta\phi$ is

$$\frac{dF}{dt} = \frac{d(\Delta\phi)}{dt} + \frac{R}{L}\Delta\phi \tag{3}$$

in which

$$\frac{dF}{dt} = \frac{\partial F}{\partial \phi_0}\frac{\partial \phi_0}{\partial t} + \frac{\partial F}{\partial a}\frac{\partial a}{\partial t} + \frac{\partial F}{\partial Z_0}\frac{\partial Z_0}{\partial t} \tag{4}$$

acts an an EMF. From (3) the flux storage for times $t > t_1$ is (assuming $d^2F/dt^2 = 0$)

$$\Delta\phi = \frac{L}{R}\frac{dF}{dt}\left\{1 - \left[1 - \frac{R}{L}\frac{\Delta\phi(t_1)}{\frac{dF}{dt}}\right]\exp\left[-\frac{R}{L}(t-t_1)\right]\right\} \tag{5}$$

and the stored energy is

$$\Delta U_m = \frac{(\Delta\phi)^2}{2L} \tag{6}$$

The details of the time release of this energy are presented elsewhere (Baum et al., 1978b) along with estimates of L and R. We note that (5) and (6) have the asymptotic flux and energy storage limits

$$\Delta\phi_\infty = \frac{L}{R}\left(\frac{dF}{dt}\right) \tag{7}$$

and

$$\Delta U_{m_\infty} = \frac{L}{2R^2}\left(\frac{dF}{dt}\right)^2 \tag{8}$$

From (1) and (7) we note that since $V_S = \frac{R}{L}\Delta\phi$, $V_S \to \frac{-dF}{dt}$ as $t \to \infty$. This is true for all times t in the no-plasma case. Hence, the plasma reconnection rate asymptotically approaches the vacuum rate.

In (8) we note that the energy storage capacity is proportional to the square of both the no-plasma reconnection rate dF/dt and the plasma conductivity. This latter condition makes it evident that the integrity of the energy storage reservoir depends on the stability of the conduction mechanism. Finally, we note that the foregoing no-plasma cellular flux analysis provides the formalism by which we can compute the energy available for flares in terms of changes in the parameters describing spot strengths and their geometrical arrangement.

In the potential limit by Faraday's law the vacuum separator voltage may be calculated from the time rate of change of flux. This was used as a driving EMF in calculating the separator electric field of reconnecting plasma systems and in computing the stored flare energy. The results will be of use in testing modified numerical solar magnetic flux codes capable of computing cellular flux ϕ_i and their changes $\Delta\phi_i$. Further, the results will be valuable in comparing with further laboratory data on bipolar pairs.

It is now important to look for correlations between cellular flux changes and flares rather than observing the less sensitive active region flux changes. To this end it would be appropriate (1) to develop a general computer code to compute such fluxes and (2) to assure that the solar probe will provide magnetogram measurements of $B_{\parallel}$ (and $B_{\perp}$ if possible) to supply accurate data for the new analysis. The applied analysis provides a new tool for use in attacking the flare problem as called for in the "Colgate report" (Colgate, 1978).

Finally it should be noted that while the present example assumes that the solar field emanates from four sunspots, that is not a requirement of this model. So long as the spatial distribution of $B_{\parallel}$ is known at several times, the method can be generalized to arbitrary field distributions.

ACKNOWLEDGMENTS

We appreciate helpful discussions with Earle Mayfield. This paper represents one phase of research carried out under National Science Foundation grant #ATM76-23463 for which we gratefully acknowledge support. We thank G. Kamin for valuable assistance.

REFERENCES

Baum, P. J., Bratenahl, A., Crockett, G., and Kamin, G., 1978a, UCR/IGPP-78/2. Submitted to Solar Physics.

Baum, P. J., Bratenahl, A., and Kamin, G., 1978b, Astrophys. J., scheduled Nov. 15.

Baum, P. J. Kamin, G., and Bratenahl, A., 1976, Trans. Am. Geophys. Union 57, 998

Bratenahl, A., and Baum, P. J., 1976, Geophys. J. Roy. Astron. Soc. 46, 259.

Colgate, S. A. (chairman, Study Committee), 1978, Space Plasma Physics: The Study of Solar-System Plasmas, Vol. 1. Reports of the Study Committee and Advocacy Panels, U. S. Nat. Academy of Sciences, Washington, D. C.

Evans, J. W., 1959, Astron. J. 64, 330.

Gopasyuk, S. I., Ogir, M. B., Severny, A. B., Shaposhnikova, E. F., 1963, Izv. Krym. Ast. Obs. 29, 15.

Howard, R., and Severny, A., 1963, Astrophys. J. 137, 1242.

Rust, D. M., 1976, Phil. Trans. Roy. Soc. Lon. A281, 427.

Schmidt, H. U., 1964, NASA SP-50, p. 107. U. S. Govt. printing office.

Svestka, Z., 1976, Solar Flares, D. Reidel, Dordrecht-Holland, p. 300.

Sweet, P. A., 1958. Nuovo Cimento Supplement 8, Series 10, p. 188.

Wu, S. T., and Smith, J. B., 1977, Private communication to P. J. Baum, KPNO magnetogram at 1431 UT.

Zvereva, A. M., and Severny, A. B., 1970, Izv. Krym. Ast. Obs. 41-42, 97.

SECTION IV.

SOLAR ENERGETIC PARTICLES

STUDIES OF ACCELERATION PROCESSES IN THE CORONA USING ION MEASUREMENTS ON THE SOLAR PROBE MISSION

George Gloeckler

Department of Physics and Astronomy
University of Maryland, College Park, MD 20742

1. Introduction

The energy spectra and composition of particles escaping from the Sun provide not only essential information on mechanisms responsible for their acceleration, but may also be used to characterize the regions where they are accelerated and confined and through which they propagate. The suprathermal energy range, which extends from solar wind energies (~1 KeV) to about 1 MeV/nucleon, is of special interest to studies of non-thermal acceleration processes because a large fraction of particles is likely to be accelerated into this energy range. One finds evidence for this in observations carried out near earth.

No direct measurements exist of suprathermal ions escaping from the Sun. Even if such ions are able to reach the orbit of earth, interplanetary processes will have modified their composition and energy spectra to such an extent as to erase most of the information they carried about processes on the Sun. The Solar Probe Mission thus offers a truly unique opportunity to observe these ions and to study their characteristics. Some of the basic questions we need to examine are: Is there an extended suprathermal tail? What are its characteristics? More specifically, how much and what material is accelerated to these suprathermal energies? How is it accelerated and where? Is this acceleration occurring all the time or is it transient? How much energy is transferred by these ions and where is this energy dumped?

2. Suprathermal Ion Measurements near 1 AU

Before turning to a more detailed discussion of scientific objectives for ion studies we will review briefly our knowledge of particles in the suprathermal energy range based on observations made near earth. During times of

no solar activity and when the local interplanetary medium is sufficiently unperturbed one observes the energy spectra shown in Fig. 1 (Mason et al, 1977, 1978). Below 1 MeV/nucleon the differential energy spectra of protons, He, C and O are rising. The relative abundances of these species are basically similar to the solar system composition. Notice that the observations stop at a few hundred KeV/nucleon. Most of the suprathermal region is, in fact, at present unexplored even at 1 AU. The origin of these particles is unknown, although it is believed that they may be accelerated in the interplanetary medium.

Following a solar flare, sub-MeV particles accelerated at the Sun begin arriving at earth, and reach their intensity about a day after the onset of the flare. In Fig. 2 are shown the energy spectra of such sub-MeV ions averaged over a one day period around maximum intensity following a 3 N, W64 flare on June 8, 1974 (Sciambi et al, 1977). In this case the spectra of He and CNO particles are flat at low energies. This flattening may be related to the acceleration process at the Sun. The energy spectra observed following smaller flares are generally steeper.

The relative abundances are observed to be variable. For example, values of the ratio of He to CNO at 40 KeV/nucleon vary from one event to the next and may be as low as 2 or as high as 20 (Sciambi et al, 1977). Such variability may again reflect conditions at the acceleration site. The relative composition also varies on a shorter time scale during a given flare particle event (Armstrong, et al, 1976) as is illustrated in Fig. 3, taken from Scholer et al, 1978. The top panel shows the intensity of oxygen following the June 8 solar flare. The middle and lower panels show the relative abundancies of Fe/O, C/O and He/H as a function of time at 0.6-1.0 MeV/nucleon

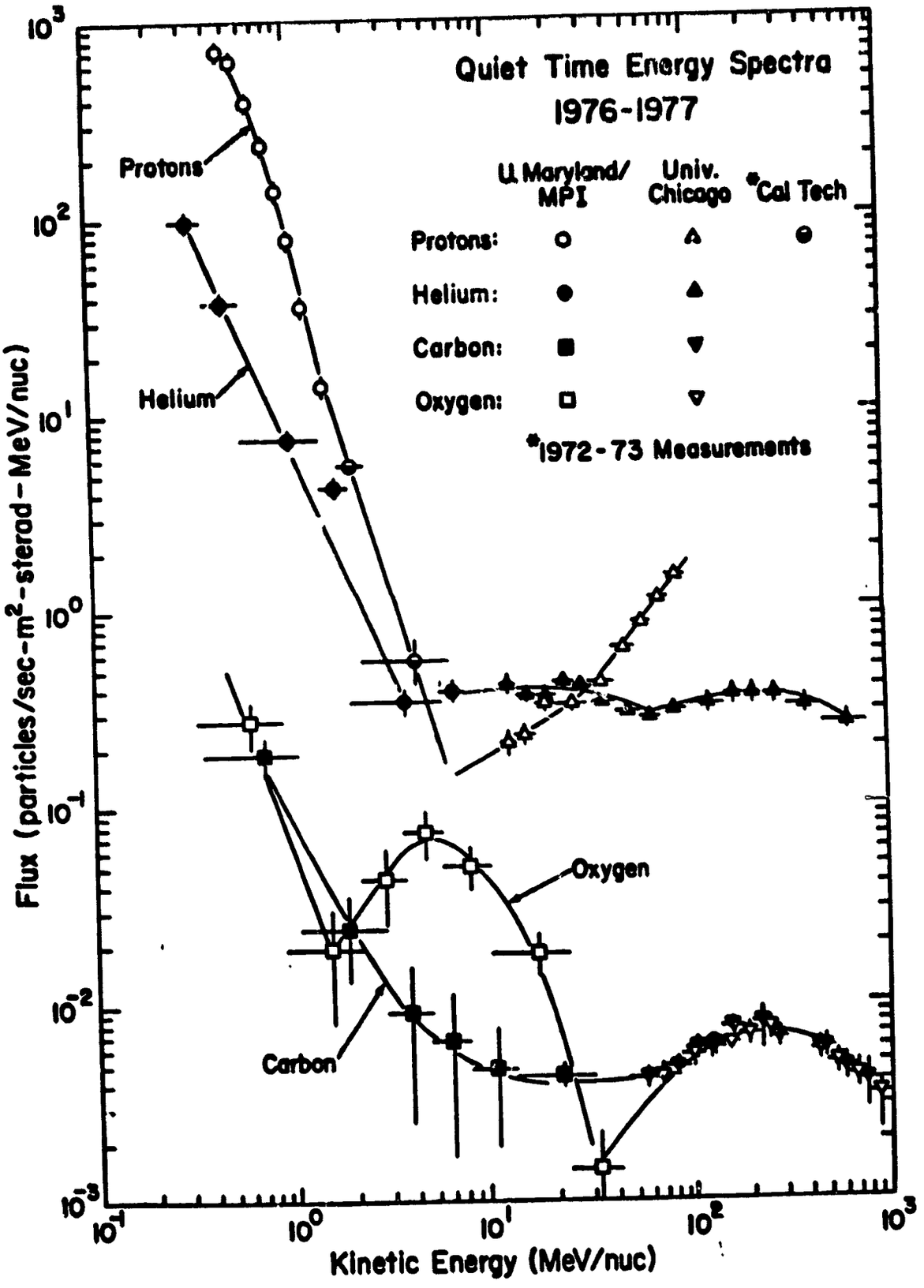


Figure 1. Differential energy spectra of protons, helium, carbon, and oxygen during "quiet times" near 1 AU. In the suprathermal energy range, the spectra increase with decreasing energy.

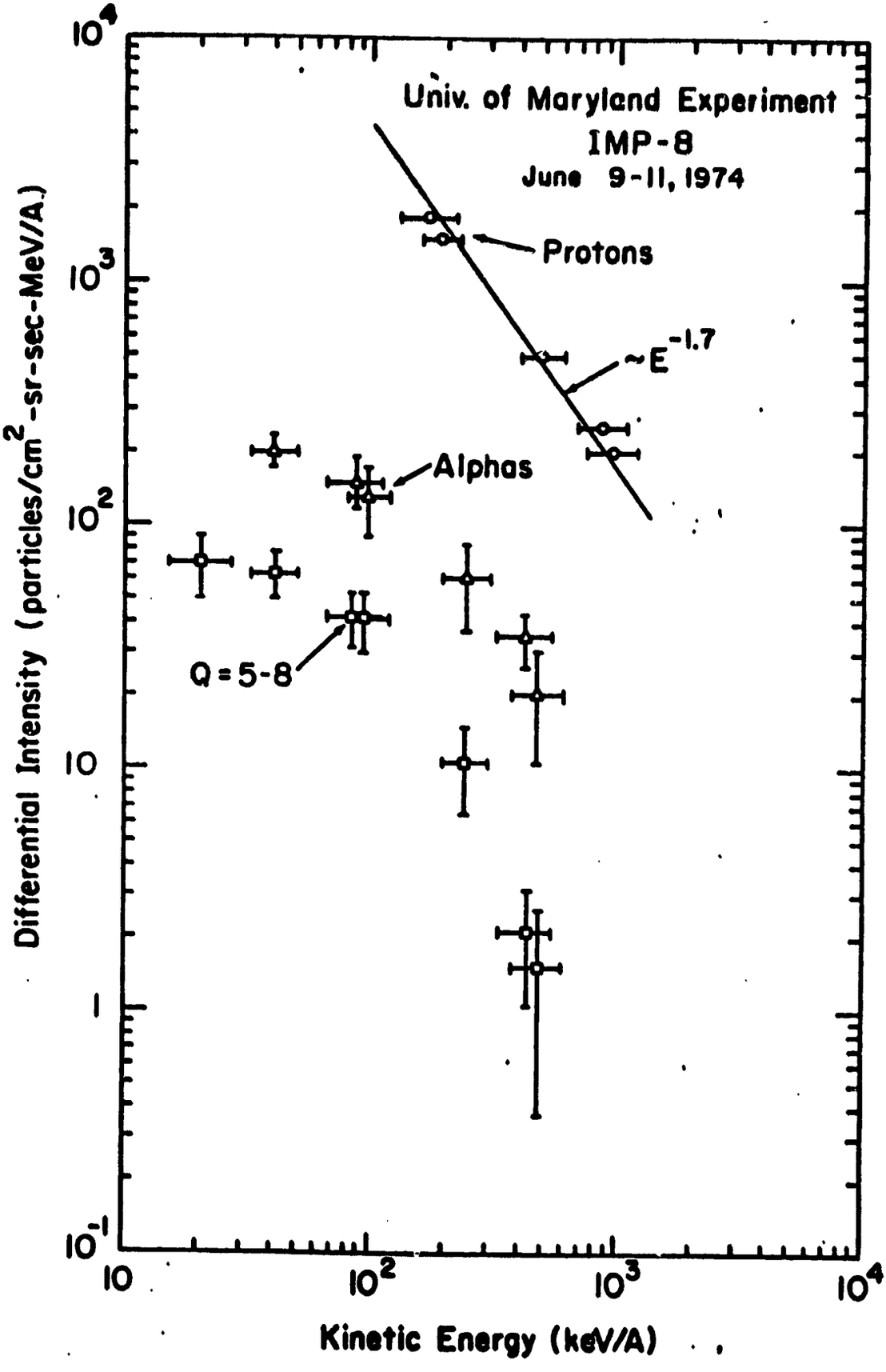


Figure 2. Proton, helium, and CNO spectra for the June 8, 1974, solar particle event. Both the helium and medium group spectra show a flattening at lower energies.

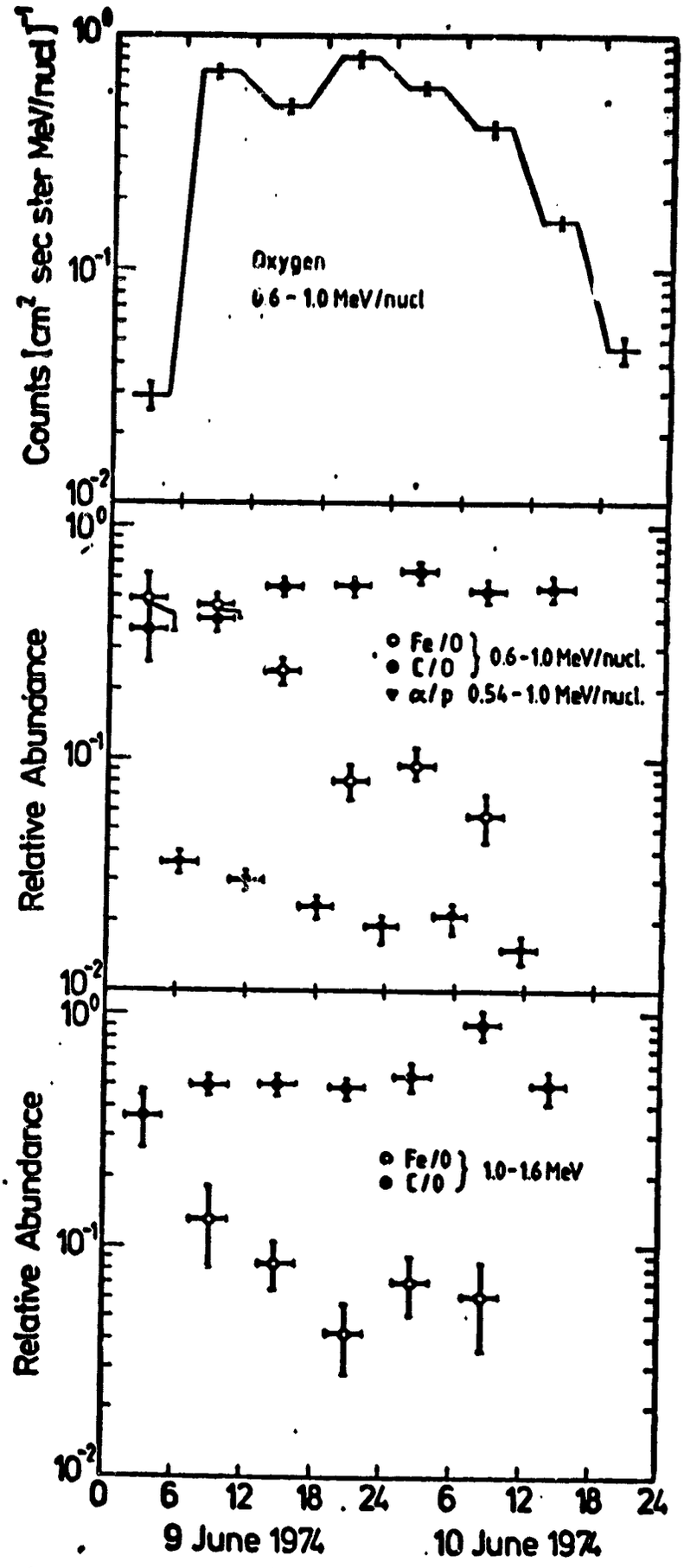


Figure 3. Temporal variations in (a) the intensity of 0.6 - 1.0 MeV/nucleon oxygen, (b) the Fe/O, C/O, and α/p ratios at 0.6 - 1.0 MeV/nucleon, and (c) the Fe/O and C/O ratios at 1.0 - 1.6 MeV/nucleon during the June 8, 1974, solar particle event.

and 1.0-1.6 MeV/nucleon respectively. Notice the large changes, particularly in the Fe/O ratio, during this event. It is not clear whether these variations reflect changing conditions in the acceleration regions or are due to propagational effects.

In a sizable fraction of solar flare particle events the composition is quite unusual. ^{3}He is observed to be enormously enriched, often exceeding the abundance of ^{4}He and approaching that of protons (see for example Gloeckler, 1975 and references therein). An example of this is shown in Fig. 4 (Hovestadt, private communication) where for a small solar flare event on May 8, 1974 are plotted the spectra of H, ^{3}He and ^{4}He. Notice the large abundance of ^{3}He compared to ^{4}He and protons. Almost without exception such ^{3}He rich flares also have enriched abundances of heavier elements (Hovestadt et al, 1975), in particular of iron. This is illustrated in Fig. 5 where the iron spectrum is seen to extend down to ~10 KeV/nucleon and its abundance is comparable to that of oxygen (Gloeckler et al, 1976). In such "iron-rich" events one generally finds low C/O ratios. To explain the ^{3}He and iron anomalies in solar flare particle events one can suppose either strong compositional inhomogeneities in the corona, or look for preferential heating and/or acceleration of certain ions. In fact, Fisk (1977) has proposed a plasma mechanism which is able to heat ^{3}He and some ions of Fe and preferentially inject these ions into the flare acceleration process.

In contrast to the enormous variability one finds in the isotopic and chemical compositions of solar flare particles, Sciambi et al (1977) find that the observed charge states of sub-MeV solar carbon and oxygen show a consistent absence of any variability with either energy, time or from one event to the next. The mean charge states observed for C and O are 5.7 and 6.2 respectively, and are quite similar to those for the solar wind.

C-3

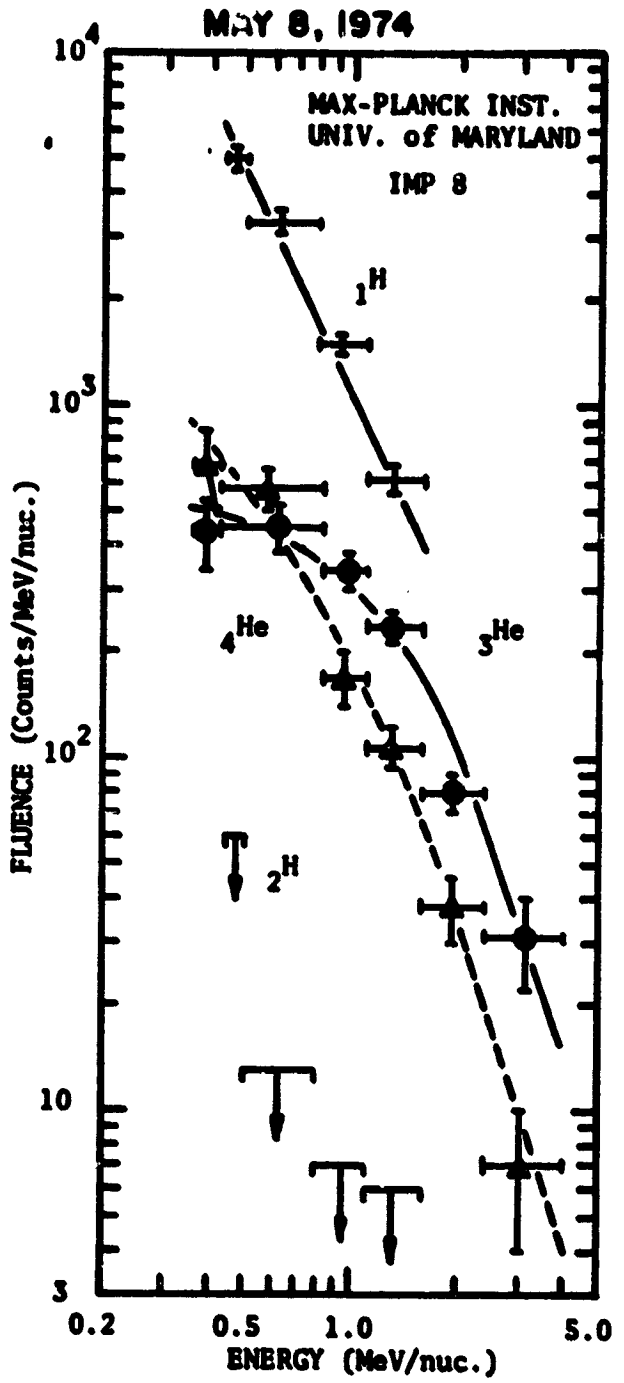


Fig. 4 Differential energy spectra of protons, ^{3}He and ^{4}He (dashed curve) for the May 8, 1974 ^{3}He-rich event. Comparing the ^{3}He and ^{4}He spectra one observes that the ^{3}He/^{4}He ratio decreases with decreasing energy.

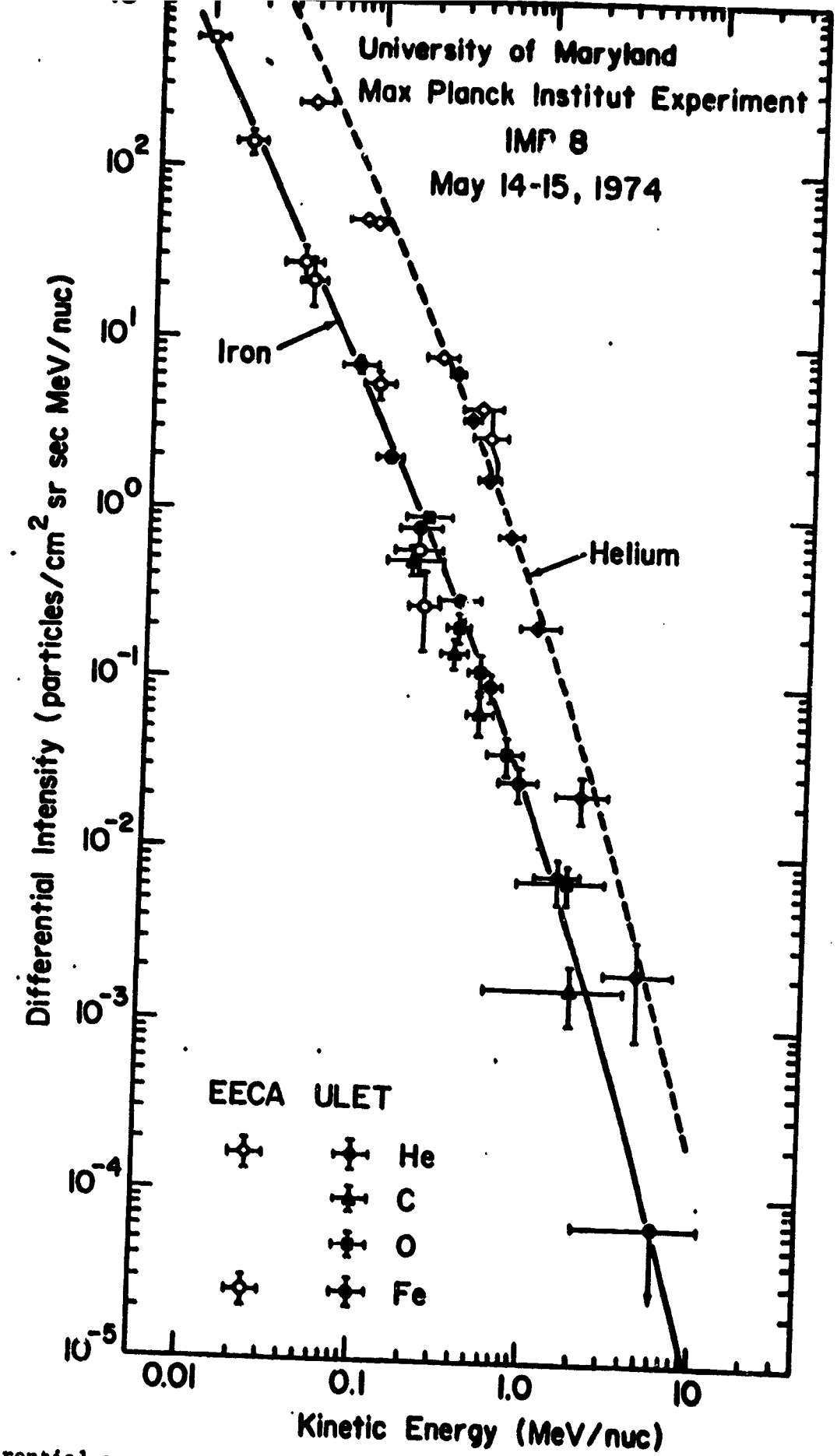


Figure 5. Differential energy spectra for He, C, O, and Fe for the May 14, 1974, iron-rich solar particle event. The abundance of Fe is about equal to that of oxygen. The C/O ratio is considerably below the 0.5 value of most events.

Evidence for interplanetary acceleration of ions to energies up to a few MeV/nucleon is provided in the observations of energetic storm particle events (ESP) and in co-rotating particle streams. ESP events are associated with interplanetary shock waves which often accompany solar flares. Particles in co-rotating streams are generally believed to be accelerated in co-rotating stream-stream interaction regions which are a common feature of the interplanetary medium beyond several AU's. Analysis of the characteristic energy spectra and composition of particles in these events as well as of the systematic changes observed in these parameters should lead to the identification and better understanding of the acceleration mechanisms involved.

3. Studies of Suprathermal Ions on the Solar Probe Spacecraft

Returning now to studies of ions on the Solar Probe Mission we can list two broad problem areas: The first is concerned with the acceleration of particles at the Sun and requires observations of ions below $\sim 10\ R_\odot$. The second set of problems is directed at examining physical processes in the inner heliosphere, primarily within a heliocentric radius of ~ 0.5 AU. In Table 1 are listed some of these questions and the types of measurements which are required for their resolution.

It is important to establish the location of acceleration regions and to characterize the conditions, such as temperature, density, plasma wave fields, etc. in these regions. For example, are suprathermal ions accelerated in regions where the solar wind is believed to originate? A detailed comparison of the composition and charge states of the solar wind with simultaneous and similar measurements of suprathermal particles is required as well as observations of how these distributions change with energy. Probing the nature of acceleration processes again requires a measurement of the energy

TABLE 1. ION STUDIES ON THE SOLAR PROBE MISSION

I. Acceleration Of Ions At The Sun ($\lesssim 10\ R_\odot$)

SPECIFIC PROBLEM	APPROACH
LOCATIONS OF AND CONDITIONS IN THE ACCELERATION REGIONS	DETAILED COMPARISON OF COMPOSITION AND CHARGE STATE DISTRIBUTIONS OF THE SOLAR WIND AND SUPRATHERMAL IONS
CORONAL HOLES OPEN FIELD LINE REGIONS TEMPERATURE WAVE FIELDS	MEASUREMENTS OF CHARGE STATES OF SUPRATHERMAL IONS AS A FUNCTION OF ENERGY
NATURE OF ACCELERATION PROCESSES	MEASUREMENTS OF ENERGY SPECTRA AND COMPOSITION OF SUPRATHERMAL IONS
NON-ADIABATIC HEATING ALFVEN WAVE ACCELERATION FRICTIONAL EFFECTS	
COMPOSITIONAL VARIABILITY IN THE SOLAR CORONA	MEASUREMENTS OF THE COMPOSITION AND CHARGE STATE DISTRIBUTIONS OF MAJOR ELEMENTS AS A FUNCTION OF ENERGY
SCALE SIZE OF INHOMOGENEITIES DENSITY AND TEMPERATURE VARIATIONS	
COMPOSITIONAL ANOMALIES (^{3}He, Fe)	

II. Processes In The Inner Heliosphere ($\lesssim 1$ AU)

STATISTICAL/SHOCK ACCELERATION IN LOW BETA PLASMA REGIONS	DETAILED COMPARISON OF THE ENERGY SPECTRA, INTENSITIES AND COMPOSITION OF SUPRATHERMAL IONS, SOLAR WIND, AND ENERGETIC PARTICLES
ADIABATIC ENERGY LOSSES	MEASUREMENTS OF ANISOTROPIES
INTERPLANETARY PROPOGATION AND RADIAL GRADIENTS OF SOLAR FLARE IONS CO-ROTATING STREAM PARTICLES	CORRELATIONS WITH THE INTERPLANETARY MAGNETIC FIELD COMPARISONS WITH 1 AU MEASUREMENTS OF SUPRATHERMAL AND ENERGETIC IONS

spectra of a number of ion species. Our observations near earth indicate considerable variability in the composition of flare particles. Are there similar and perhaps related compositional variations in the suprathermal ion population near the Sun? How do these in turn reflect compositional inhomogeities in the s ar corona?

The second category includes studies of statistical and shock acceleration processes in low beta plasma regions. (Beta is expected to decrease from $\sim$1 at 1 AU to $\lesssim 10^{-3}$ in the lower corona.) Interplanetary effects on suprathermal ions are particularly important and can be studied by comparing measurements made on the Solar Probe spacecraft with similar observations at $\sim$1 AU. We can learn much about the propagation of particles in the inner heliosphere by examining radial gradients, anisotropies and adiabatic energy losses of a number of representative ion species as a function of energy.

4. Instrumentation for Suprathermal Ion Measurements on the Solar Probe Spacecraft

The instrumentation for carrying out the required suprathermal ion measurements must be sufficiently flexible to operate under all conceivable conditions likely to be encountered in an exploratory mission. In Table 2 are listed some of the necessary capabilities of an ion composition experiment on the Solar Probe Mission. The energy range should extend from that of the solar wind to about 1 MeV/charge. It is essential to have a large intensity dynamic range. Energy spectra of various ions should be measured with a resolution of less than 10% to resolve spectral features. The charge and mass resolution should be sufficient to identify charge states of major elements up to Fe and at least the He isotopes.

The instrument should have a high degree of immunity to background and have provisions for some degree of on-board data compression and storage to reduce telemetry requirements. Limits on weight and power will be additional constraints.

TABLE 2.

REQUIRED ION MEASUREMENTS ON THE SOLAR PROBE MISSION

ENERGY RANGE OF IONS	$\sim$ 3 - 1000 KE V/Q
INTENSITY DYNAMIC RANGE	$\sim 10^{10}$

RESOLUTION

-ENERGY:	$\Delta E/E \lesssim 0.1$
-TIME:	$\lesssim$ 1 SEC TO $\sim$ 1 MIN
-CHARGE STATES:	MUST RESOLVE CHARGE STATES OF MAJOR ELEMENTS
-MASS:	MUST RESOLVE ELEMENTS UP TO FE, AND ISOTOPES ^{3}HE AND ^{4}HE

OTHER REQUIREMENTS

- -BACKGROUND IMMUNITY TO SIG AND ENERGETIC PARTICLES
- -ON-BOARD DATA REDUCTION TO REDUCE TELEMENTRY REQUIREMENTS TO 50-100 BPS
- -MUST OPERATE AT RELATIVELY HIGH TEMPERATURES (45-50° C)

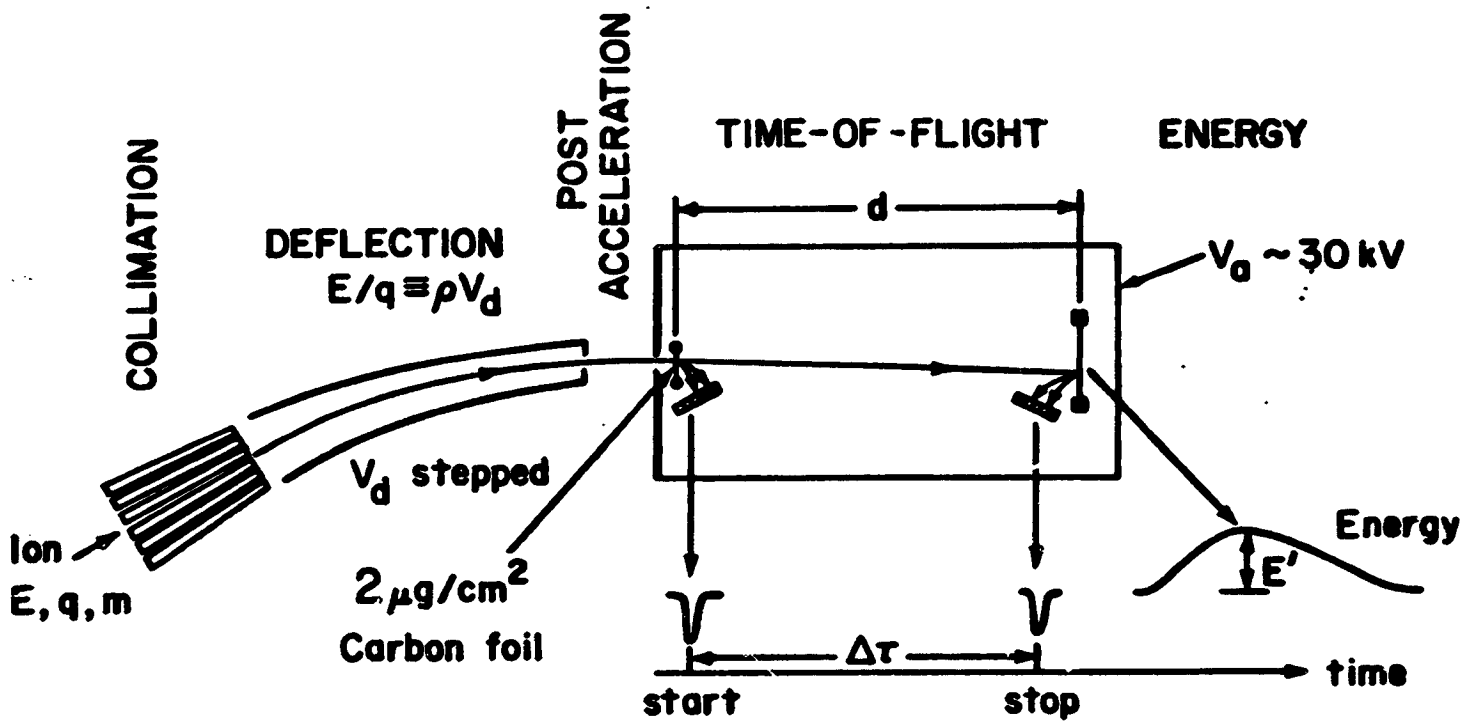
PROBLEM AREAS

- -SUNWARD LOOK DIRECTION
- -ANISOTROPY MEASUREMENTS ON A NON-SPINNING SPACECRAFT
- -THERMAL DESIGN

The major problem areas which are apparent at this time are how to achieve the required sunward look direction, and how to perform within the weight constraints anisotropy measurements on a non-spinning spacecraft. The thermal design of all instruments on this spacecraft is likely to be a major concern.

An instrument which could meet the requirements listed here is represented schematically in Fig. 6 (see Gloeckler (1977) for a more detailed description). It consists of an electrostatic deflection system, modest post acceleration and a time of flight vs. energy system. Briefly, an ion characterized by energy E, charge q and mass m enters the curved deflection system driven by a stepped voltage V_d. Only ions which have the appropriate E/q ratio are able to exit the deflection system at any given time. After post-acceleration the ions enter the TOF system consisting of a start and stop detector separated by ~10 cm. The ion signals its passage through the start detector by emitting secondary electrons from surfaces of a thin foil which are then detected by a microchannel plate assembly. Similarly, the stop signal is provided by secondary electrons emitted from the surface of a solid state detector which is also used to measure the energy of the ion. As is indicated at the bottom of the figure, it is possible to determine E, q, and m of the incoming ion by combining the information provided by simultaneous measurements of the time-of-flight and energy with the knowledge of the deflection and post-acceleration voltages.

A possible configuration of the instrument is shown in Fig. 7. In this configuration two deflection systems are used to obtain the required energy range and to provide simultaneous measurements at two energies at any given time. (This system could also be used to measure the solar wind composition.) The compact TOF system may be shielded to provide high degrees of immunity to background due to large fluxes of energetic particles.



$$m/q = \alpha(\Delta\tau)^2 \qquad q = \beta E'/(V_a + \rho V_d) \qquad E = q\rho V_d$$

$$E/q = \rho V_d \qquad \alpha = 2(V_a + \rho V_d)/d^2 \qquad 1/\beta = \text{nuclear defect}$$

Figure 6. Schematic representation of the measurement technique for suprathermal ions. Post-acceleration is required at the low end of the suprathermal energy range.

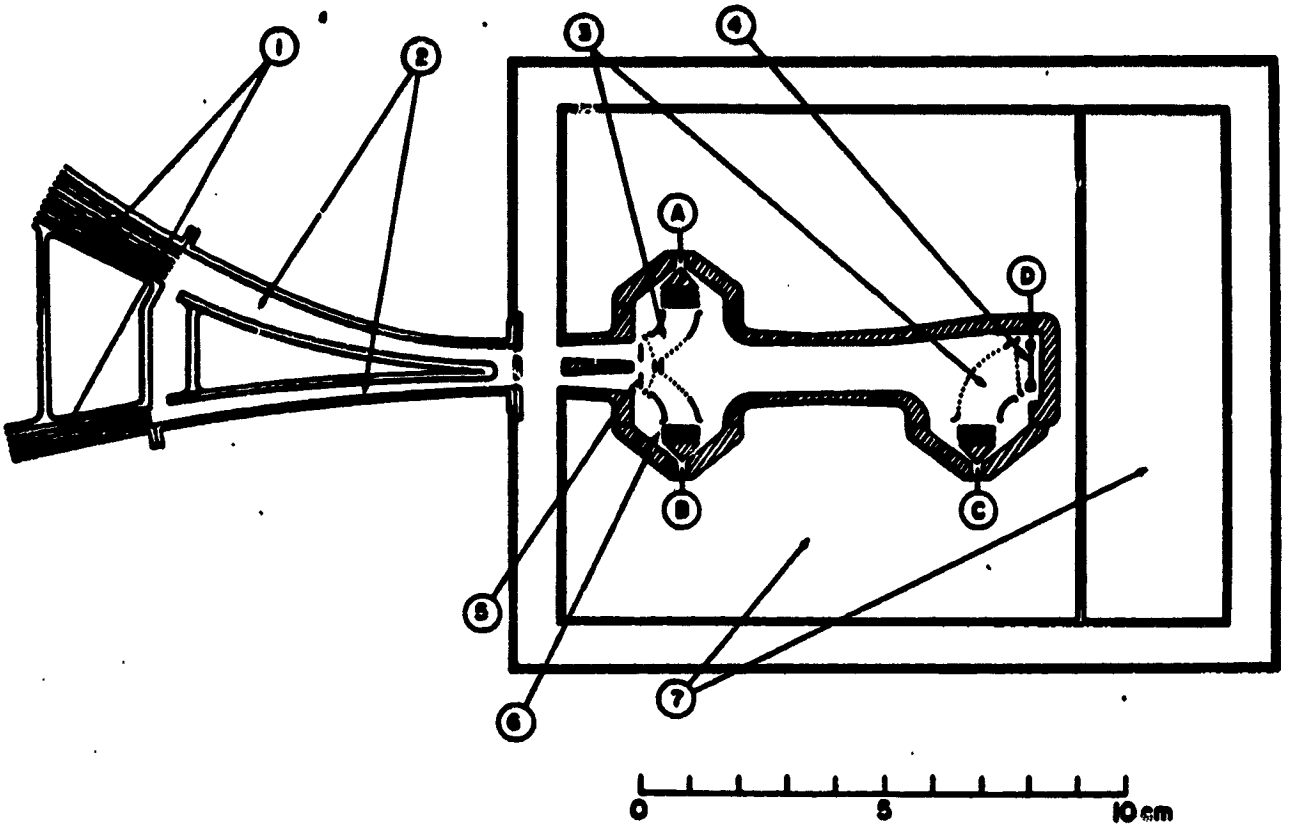


Figure 7. Simplified cross-section of a suprathermal ion detector which includes post-acceleration and two electrostatic deflection analyzers: (1) multi-slit focusing collimator; (2) electrostatic deflection analyzer; (3) acceleration and deflection; (4) rectangular solid state detector; (5) carbon foils; (6) chevron microchannel plate assemblies; and (7) housings for fast electronics.

A possible solution for achieving both a sunward view directions and two dimensional anisotropy measurements is illustrated in Fig. 8 which shows a schematic representation of the Solar Probe Spacecraft and the Ion Composition Experiment. The entrance aperture of the instrument looks into a larger deflection system consisting of an electrically grounded outer grid, a segmented inner deflection plate and a grounded support. Segments of the inner deflector plates are turned on electrically, thus selecting the look direction. With all segments turned on one would look into the sunward direction, with all segments off the instrument would view the anti-solar direction. On a spacecraft with a spinning section this scheme would permit 3 dimensional anisotropy measurements. Without a spinning section, some 3 dimensional anisotropy measurements could still be made by placing a cluster of simple ion telescopes with carefully chosen view directions on the side opposite to the main instrument.

5. Acknowledgments

Discussions with L. A. Fisk and W. C. Feldman are gratefully acknowledged. This work was supported in part by NASA under grants NGR-21-002-224 and NGR-21-002-316.

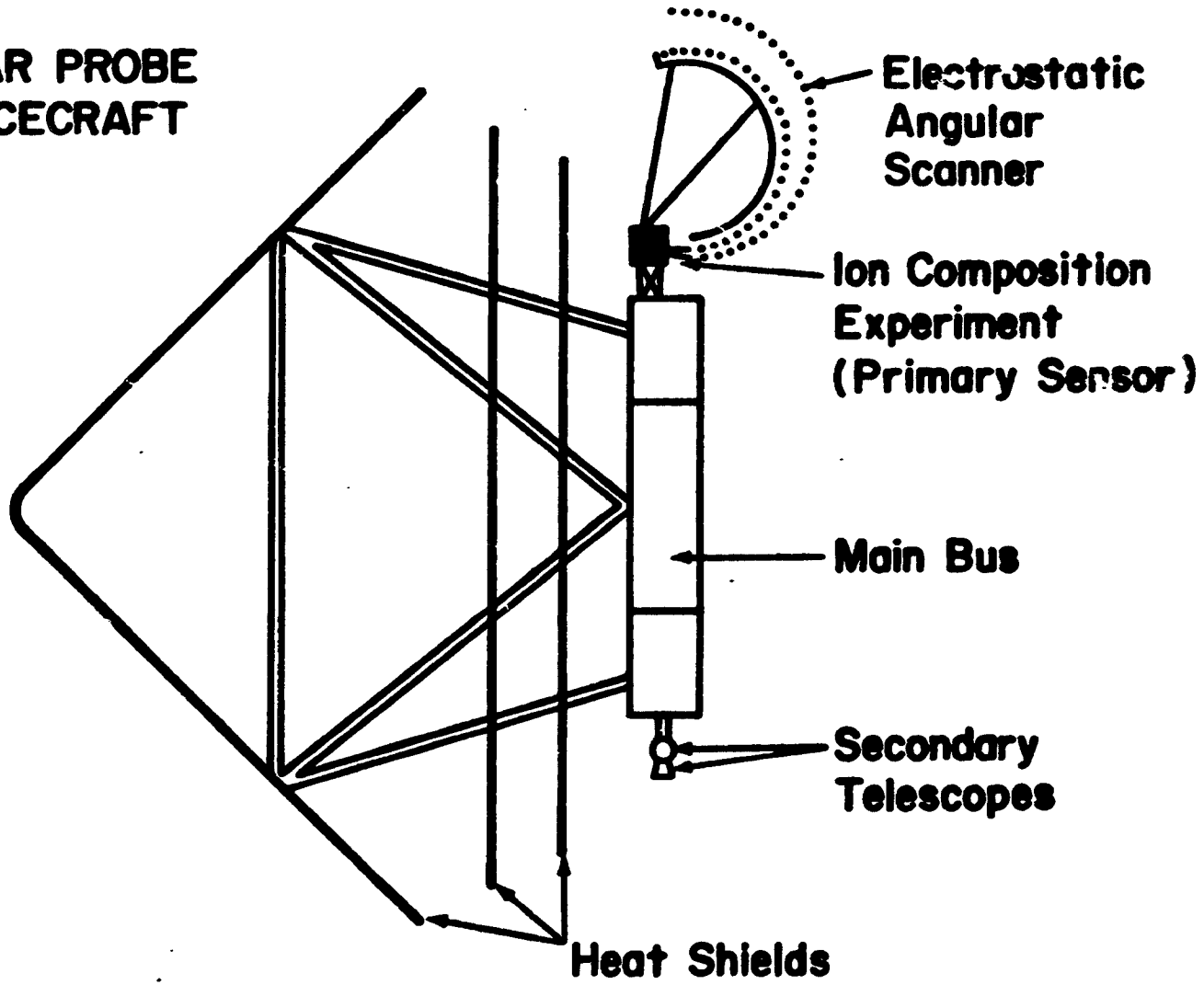


Figure 8. Possible location of the ion composition experiment on the Solar Probe spacecraft. The electrostatic angular scanner (see text) may be used to provide the necessary sunward view and 3-dimensional anisotropy measurements on a spinning spacecraft. On a non-spinning spacecraft, secondary telescopes with carefully chosen view direction would provide some 3-dimensional anisotropy measurements.

REFERENCES

Armstrong, T.P., S.M. Krimigis, D. Hovestadt, B. Klecker, and G. Gloeckler, Observation of temporal and special variations in the Fe/O charge composition of the the solar particle event of 4 July, 1974, Solar Physics 49, 395 (1976).

Fisk, L.A., ^{3}He-rich flares: a possible explanation University of New Hampshire preprint, December 1977.

Gloeckler, G., A versatile detector system to measure the charge states, man compositions and energy spectra of interplanetary and magnetospheric ions, University of Maryland Technical Report TR 77-043, (1977).

Gloeckler, G., R.K. Sciambi, C.Y. Fan, and D. Hovestadt, A direct measurement of the charge states of energetic iron emitted by the Sun, Ap J Letters 209, L. 93 (1976).

Gloeckler, G., Low energy particle composition, 14th International Cosmic Ray Conference, Munich, Germany, August 15-29, 1975, Conference Papers 11, 3784 (1975).

Hovestadt, D., B. Klecker, O. Vollmen, G. Gloeckler, and C.Y. Fan, Heavy particle emmision of unusual composition from the Sun, 14th International Cosmic Ray Conference, Munich, Germany, August 15-29, 1975, Conference Papers 5, 1613 (1975).

Mason, G.M., G. Gloeckler, C.Y. Fan and D. Hovestadt, Energy spectra of quiet-time interplanetary protons, helium and heavier nuclei below 10 MeV/nucleon near solar minimum, Trans. Am. 6 ev plays Union, 58, No. 12, 1204 (1977); to be published (1978).

Scholer, M., D. Hovestadt, B. Klecker, and G. Gloeckler, and C.Y. Fan, Temporal development of the energetic particle composition during solar flares, J. Geophys Res. to be published (1978).

Sciambi, R.K., G. Gloeckler, C. Y. Fan, and D. Hovestadt, Direct measurement of the ionization states of energetic carbon and oxygen emmitted by the Sun, Ap. J. 214, 316 (1977).

Electron Studies of Acceleration Processes in the Corona

N78-32982

R.P. Lin, Space Sciences Laboratory,
University of California, Berkeley, CA 94709

I. Introduction

The Solar Probe Mission can obtain unique and crucially important measurements of electron acceleration, storage, and propagation processes in the corona. In addition, electron measurements on Solar Probe can provide a novel way to probe the magnetic field structure of the corona below the spacecraft. We describe the various energetic electron phenomena which will be sampled by the Solar Probe and then suggest some new techniques to probe coronal structures.

II. Coronal Electron Phenomena

Quiet time electrons

Figure 1 shows the spectrum of electrons at solar quiet time observed in the near-earth interplanetary medium (Lin et al., 1972). Solar wind electrons make up the lowest energy, $\lesssim$1 keV portion of the spectrum. Electrons from a few hundred keV to $\sim$10 MeV apparently are of Jovian origin. Above ~10 MeV is the cosmic ray electron population. However, the origin of a few keV to a few hundred keV electrons is unknown. It seems likely these are solar, perhaps originating in some quasi-continuous acceleration process near the sun. Electron measurements on Solar Probe should help to discover their origin.

Storms of type III bursts (Fainberg and Stone, 1970) are a phenomenon which appear to be unique to the hectometric wavelength $\sim$0.3 to 3 MH_2 radio range (Figure 2). These wavelengths correspond to coronal distances of a few to a few tens of solar radii. In these storms both fast drifting

type II-like bursts and underlying continuum are observed. The bursts occur on the average about once every 10 sec at a few MHz; tens to hundreds

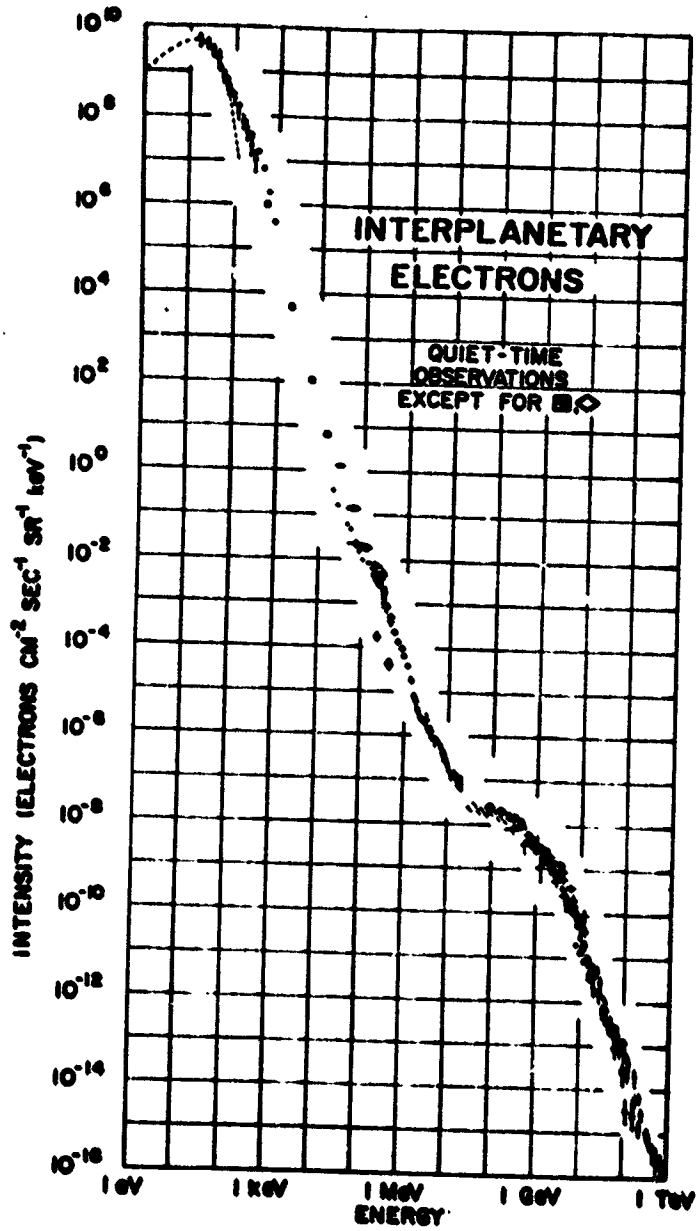


Figure 1. The quiet-time interplanetary electron spectrum from solar wind to galactic cosmic ray energies (Lin et al., 1972). The dashed line indicates a 1.5 X 10^5 K Maxwellian. Below ~100 keV the electrons are believed to be solar in origin.

of thousands occur in a typical storm. The storms appear to be associated with active regions and type I noise storms at higher frequencies, and they are a common phenomenon at times of enhanced solar activity. Under current plans, the Solar Probe will pass right through the type III storm heights at times near solar maximum, thus permitting in situ observations of the type III storm electron acceleration process.

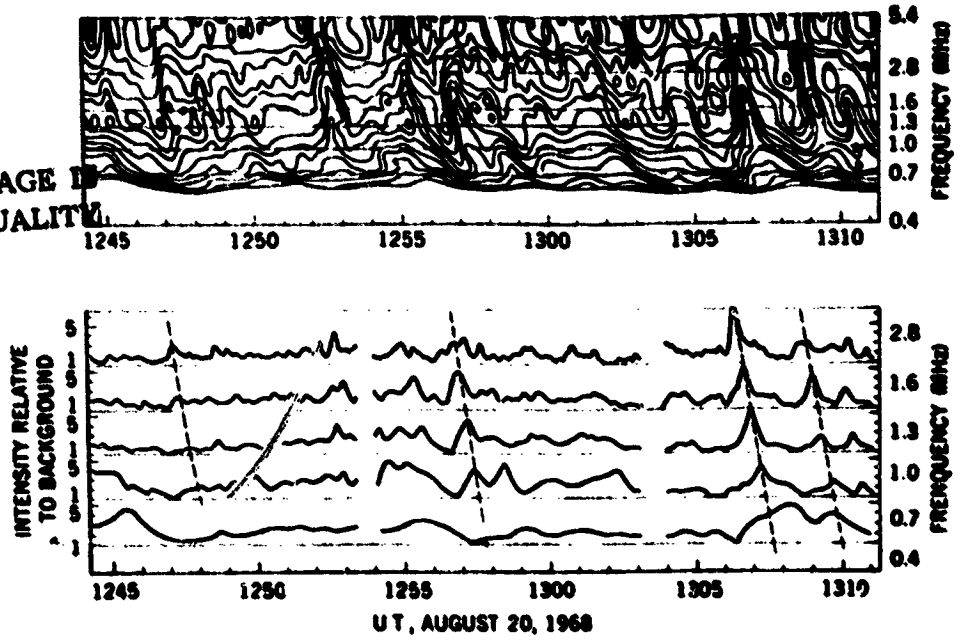


Figure 2. A 30-min segment of data of a solar storm of type III bursts near CMP. The drift in frequency, limited burst bandwidth, and large numbers of small bursts can be seen (Fainberg and Stone, 1970).

Impulsive electron acceleration; flares and type III bursts

Electron acceleration to $\sim 10\text{-}10^2$ keV commonly occurs during the flash phase of even small flares and subflares. Type III bursts indicate that impulsive electron acceleration occurs perhaps at a lower intensity level, even more frequently than do flares (although some optical activity is usually observed to accompany these bursts). The accelerated electrons

which escape are observed in the interplanetary medium (Figure 3), but usually not before substantial interplanetary scattering and diffusion of the particles has "erased" most information on the temporal and pitch angle structure of the acceleration and/or escape processes. Even for "scatter-free" events, just the velocity dispersion at these energies is sufficient to erase the available temporal information. From analysis of the diffusion characteristics of these electron events, it appears that most of the interplanetary scattering and diffusion occur well outside $\sim 10\ R_\odot$. Thus, the Solar Probe mission provides a unique opportunity to observe the temporal and angular structure of the acceleration free from the "curtain" of interplanetary propagation. Hard x-ray and type III burst

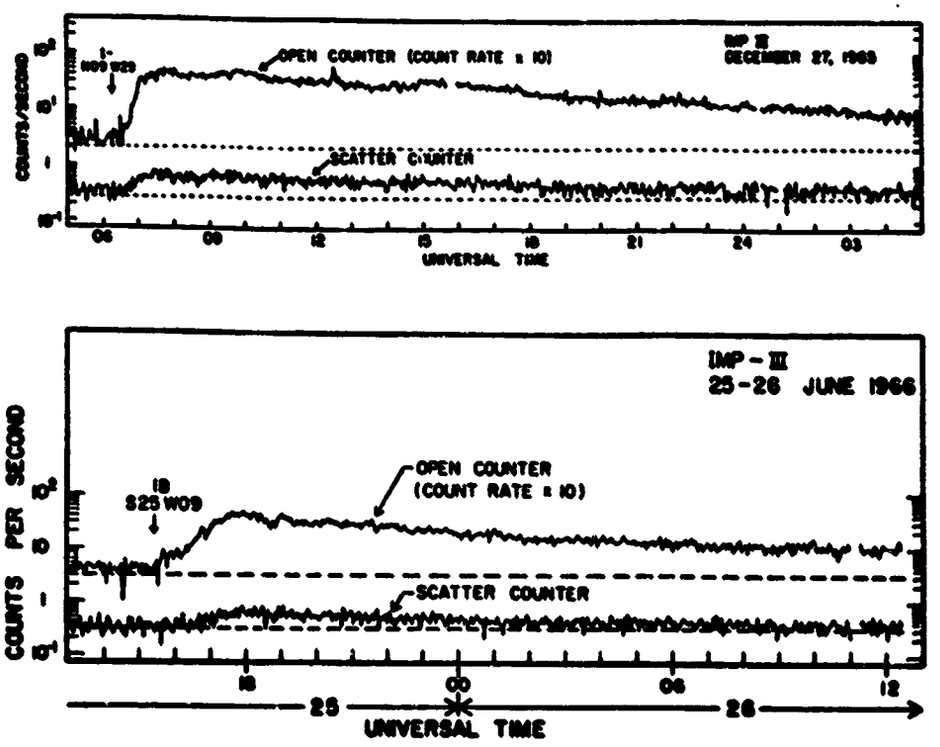


Figure 3. Two electron events observed at >45 keV energies (from Lin, 1974).

observations (see Figure 4) indicate that the acceleration time scales are typically seconds. The Solar Probe at distances less than $\sim$ ten solar radii may see the accelerated and escaping electrons unaffected by scattering. As described in section II, such observations can provide information on the location, density, magnetic field strength of the acceleration region, and detailed information on the temporal and spectral behavior of the acceleration process.

Particle storage in the upper corona

Observations of the bursts (Figure 5) at frequencies extending down to $\lesssim$1 MHz indicate that loop structures are present out to $\gtrsim 35\ R_\odot$. These loops are likely to be transient phenomena, but based on the Skylab coronagraph results, loop transients of this type should occur very frequently (a few per day) near solar maximum. Loop structures extending into the upper corona where ambient densities are $\lesssim 10^7 cm^{-3}$ provide excellent long term storage regions for energetic electrons and ions. It is possible that the bulk of the $10-10^2$ keV electrons accelerated in the impulsive phase of the flare escape to these regions. This would be consistent with the thin-target interpretation of the hard x-ray emission (Datlow and Lin, 1973). Also, there is substantial evidence from solar particle observations that storage and/or continual acceleration often occurs near the sun. Because the density is low, energetic electrons in these regions will be effectively invisible from the earth. The passage of Solar Probe through these regions near solar maximum should have an excellent chance of intersecting these storage regions.

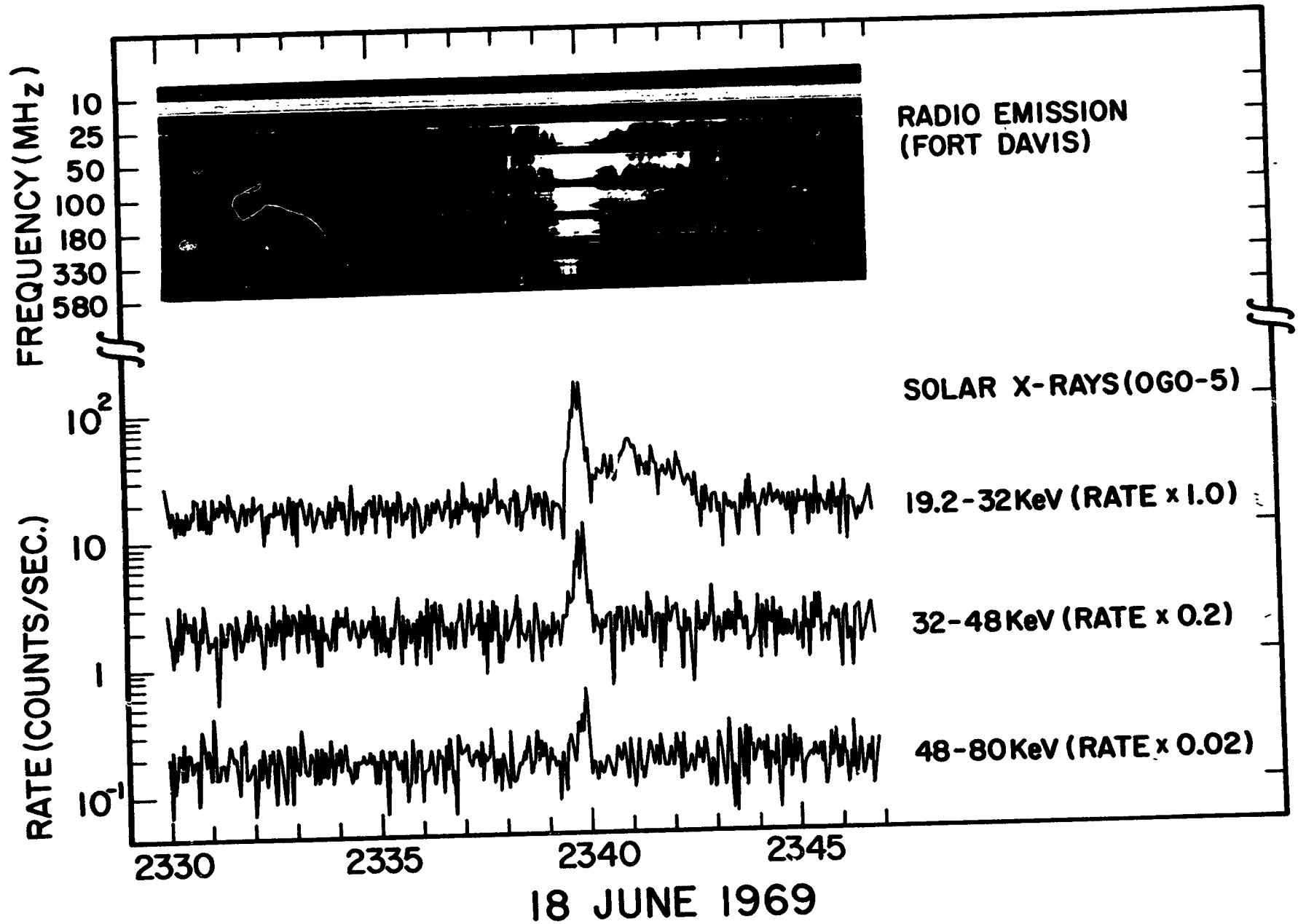


Figure 4. Simultaneous hard X-ray and Type III radio bursts (Kane, 1974). Note the correspondence of the X-ray peaks to

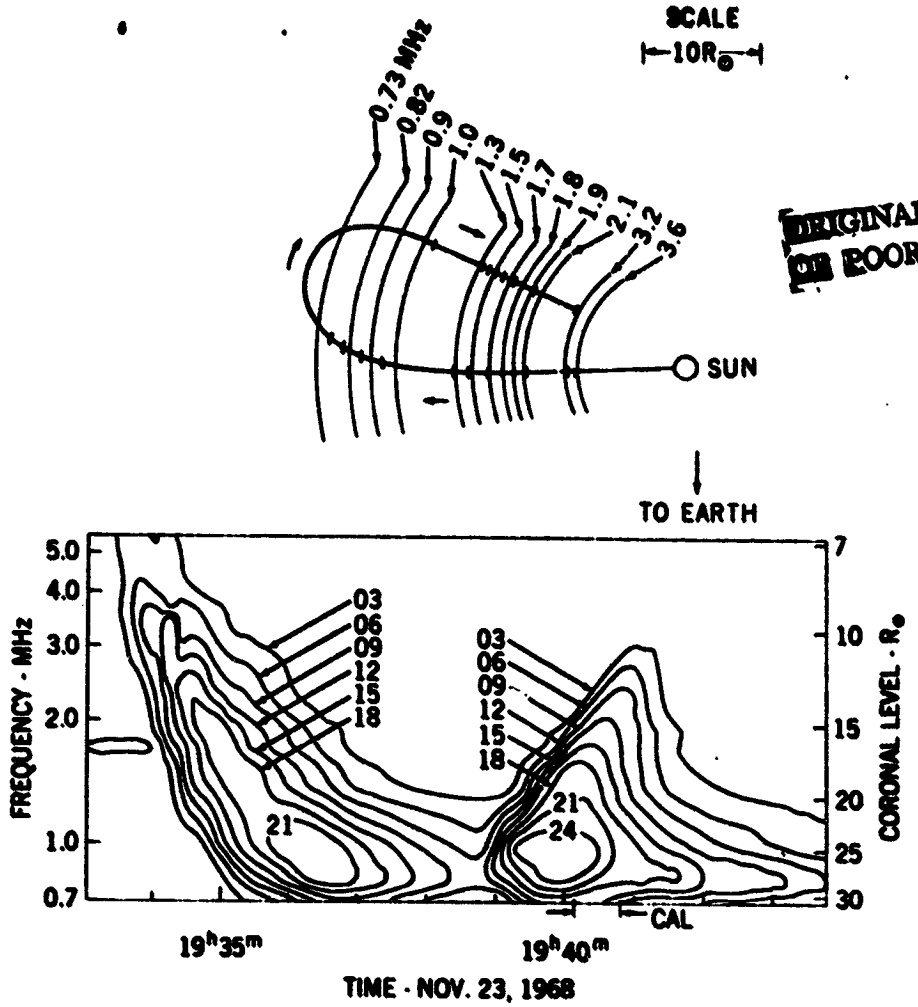


Figure 5. (a) Dynamic spectrum of a hectometer U burst. Numbers on the figure indicate intensity in decibels relative to cosmic noise background. Note the frequency limit of the return branch. (b) The magnetic bottle inferred from the U burst. The model is based on the RAE-1 density scale and an assumed exciter speed of 0.3 c. Bars on the figure show the positions determined from the dynamic spectrum relative to the start of the burst (Stone and Fainberg, 1971).

Shocks and second stage acceleration

Current evidence suggests that energetic ions and relativistic electrons are accelerated by shock waves in a second acceleration stage following the impulsive phase of large flares (see Lin, 1974 for review). The observations of type II bursts and associated type III bursts and "herring bone" structure are strong evidence for electron acceleration by shocks in the corona. Type II bursts have been observed to ≤ 1 MHz (Figure 6) and the shock waves are often observed near 1 A.U. several

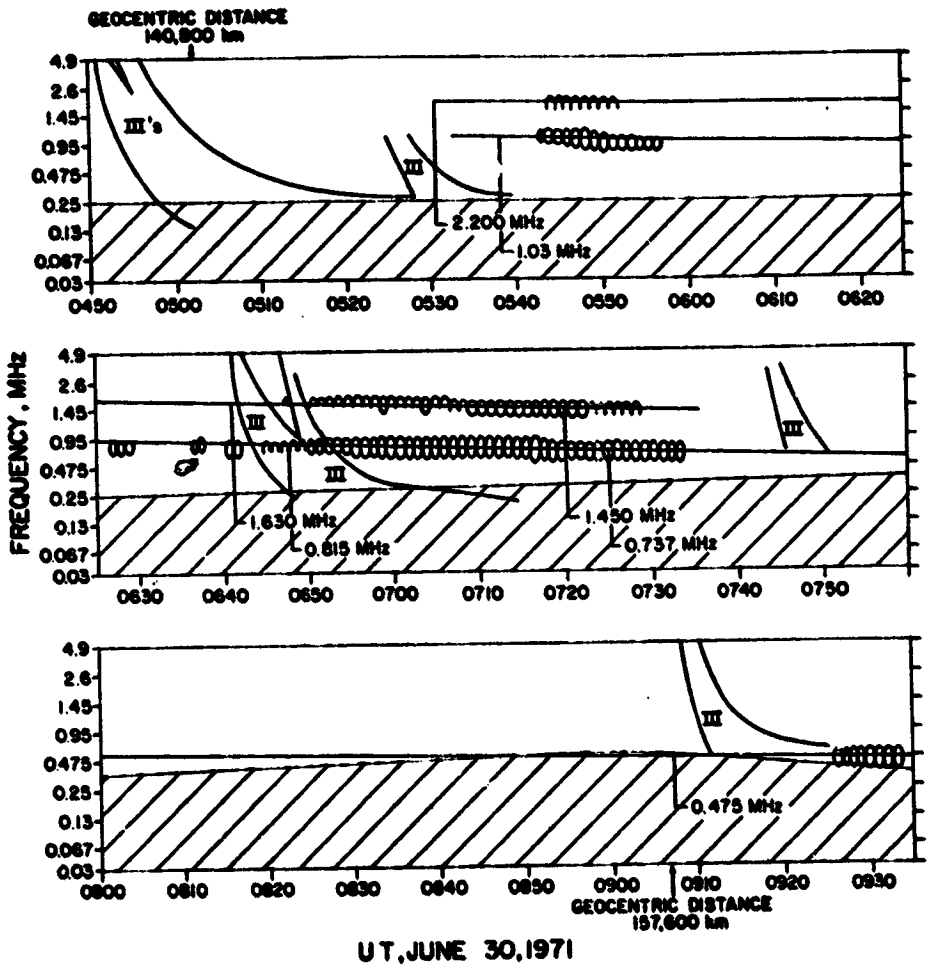


Figure 6. The first type II radio burst to be observed at frequencies below the ionospheric cutoff occurred on June 30, 1971; a tracing made from the intensity contour plot of the IMP-6 radio data is shown here (Malitson et al., 1973). The shaded portion of the record at the lower frequencies represents magnetospheric noise.

days later. Ions are often observed to be accelerated by these interplanetary shocks to ∿1-10 MeV energies. Although these shock type II radio burst events are relatively rare even near solar maximum (∿few/month) the direct in situ observations of such an event near the sun would go far toward identifying the specific processes by which particles are accelerated to very high energies by the sun. It would be useful to have some shock observations even at distances of tens of solar radii. Clearly the chances of observing a shock event would be increased by a short orbital period (≲ 1 year) for the solar probe.

III. Electron Measurements on Solar Probe

As the Solar Probe mission approaches the sun the integrated effects of pitch angle scattering and diffusion in the interplanetary medium eventually become negligible compared to the effect of the magnetic field divergence. Then the motion of charged particles can be described as essentially adiabatic. Conservation of the first adiabatic invariant then implies that the pitch angle, α, of a charged particle is related to the magnetic field at every point in its trajectory by the equation

$$\frac{\sin^2 \alpha}{B} = \text{constant}$$

Figure 7 is a schematic illustration of what would be observed by the Solar Probe if it was in this adiabatic region when an impulsive electron acceleration event such as a flare or type III burst occurs. Some of the downward moving particles would reach a dense region of the solar atmosphere where they would produce hard x-rays. These would be observed by x-ray detectors at the Solar Probe with the time delay indicated in

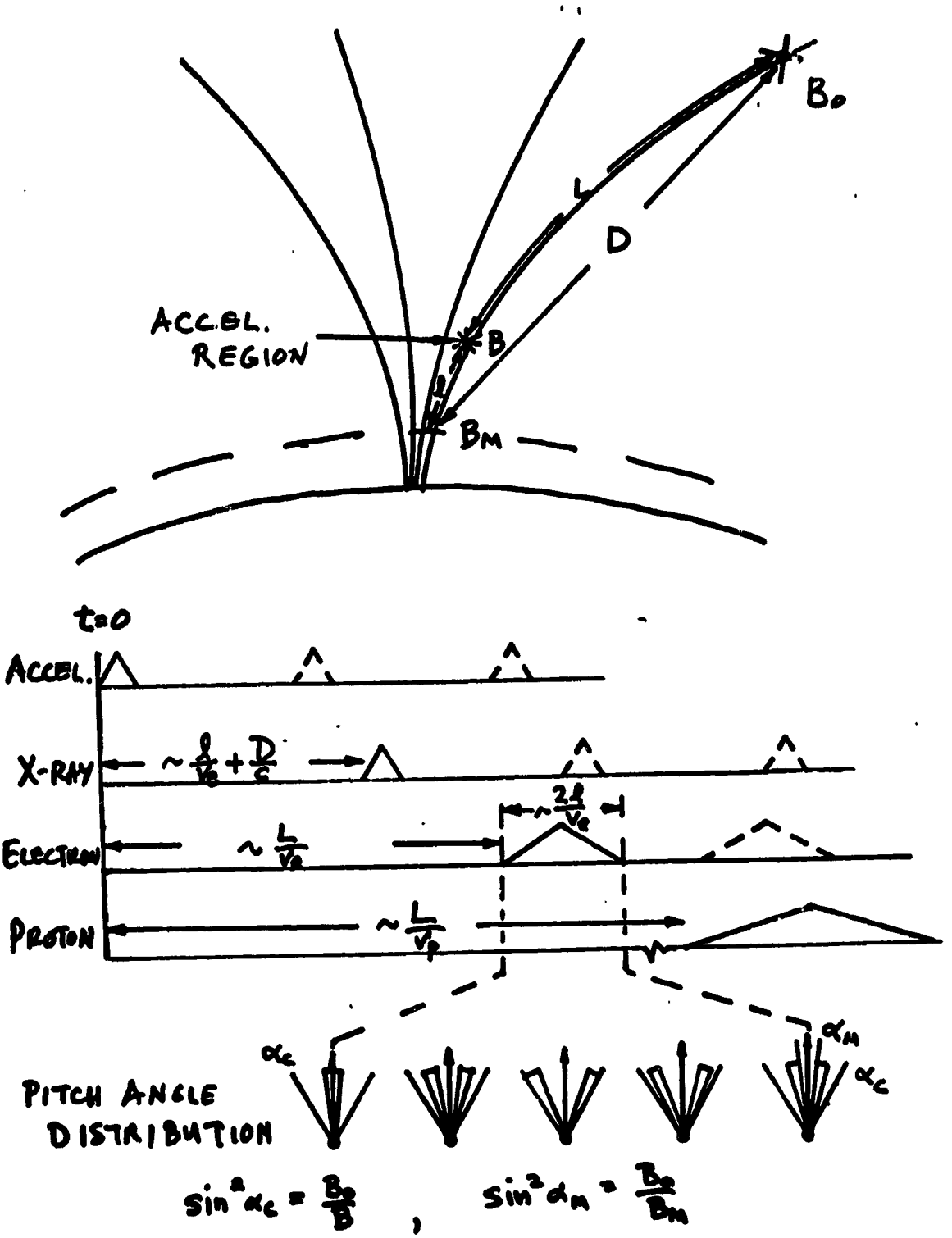


Figure 7. Schematic illustration of an impulsive acceleration event (above) and observations by Solar Probe (below).

Figure 7. A rudimentary hard x-ray imaging detector could locate this region on the solar disk. A short time later the upward moving electrons would be observed going by the satellite if the propagation conditions are favorable. The first arriving particles are those with the smallest pitch angles. Eventually, as illustrated, the particles originally accelerated with 90° pitch angles arrive with the maximum pitch angle of $\alpha_c = \text{arc sin } (\sqrt{B_o/B})$. The particles which originally were headed toward the surface and were mirrored by the stronger fields near the surface are observed. However, these have a minimum pitch angle of $\alpha_e = \text{arc sin } (\sqrt{B_o/B_e})$ due to those downward moving particles with small pitch angles which were able to reach low enough altitudes to be lost by Coulomb collisions (α_e is commonly known as the loss cone angle in magnetospheric studies). Thus, the magnetic field strength in both the acceleration and loss regions can be obtained. The field lines can be traced by comparing the hard x-ray source position to the spacecraft position. Comparison of the timing of the x-ray and electron pulses gives the distances along the field line, ℓ and L. By observing different energy particles the magnetic field strength at different stopping altitudes will be sampled. Also the presence of parallel electric fields can be detected from comparing the pitch angle distribution for different energy particles. Finally, the question of whether low energy protons are also simultaneously accelerated in these impulsive events is likely to be resolved since close to the sun ($\lesssim 10\ R_o$) the effects of propagation - velocity, dispersion, scattering and energy changes - should be insignificant. Figure 8 indicates the energy loss range and travel times of electrons and ions of $\sim$1 to 10^3 keV energy. Also indicated are the approximate

column densities traversed during their travel down to various heights in the corona. Based on Figure 8 it appears that field line lengths can be measured by electrons to $\sim 0.1\ R_\odot$, and the solar atmosphere and its magnetic field can be probed from below the transition region to $>1\ R_\odot$ height in the corona by using different energy particles.

It is also possible to probe the corona in the absence of impulsive events. The pitch angle distribution for quiet time particles should also show a loss cone related as above to the field strength at the particle stopping height. These determinations of the coronal field strength can be used to map the coronal magnetic field and determine whether the field at the spacecraft originates in the weak or strong solar surface field regions.

Based on the frequencey of electron acceleration in flares and type III bursts near solar maximum we expect that tens of events will be observed by Solar Probe in a single passage inside $10\ R_\odot$. Furthermore, the particle fluxes from an impulsive event should be intense enough for these measurements since they scale as R^{-3}. However, as indicated by Figures 7 and 8, the time scales are short so that data storage and/or high telemetry bit rate will be required.

Acknowledgements. I wish to thank S.R. Kane for his assistance. This research was supported in part by NASA grant NGL 05-003-017.

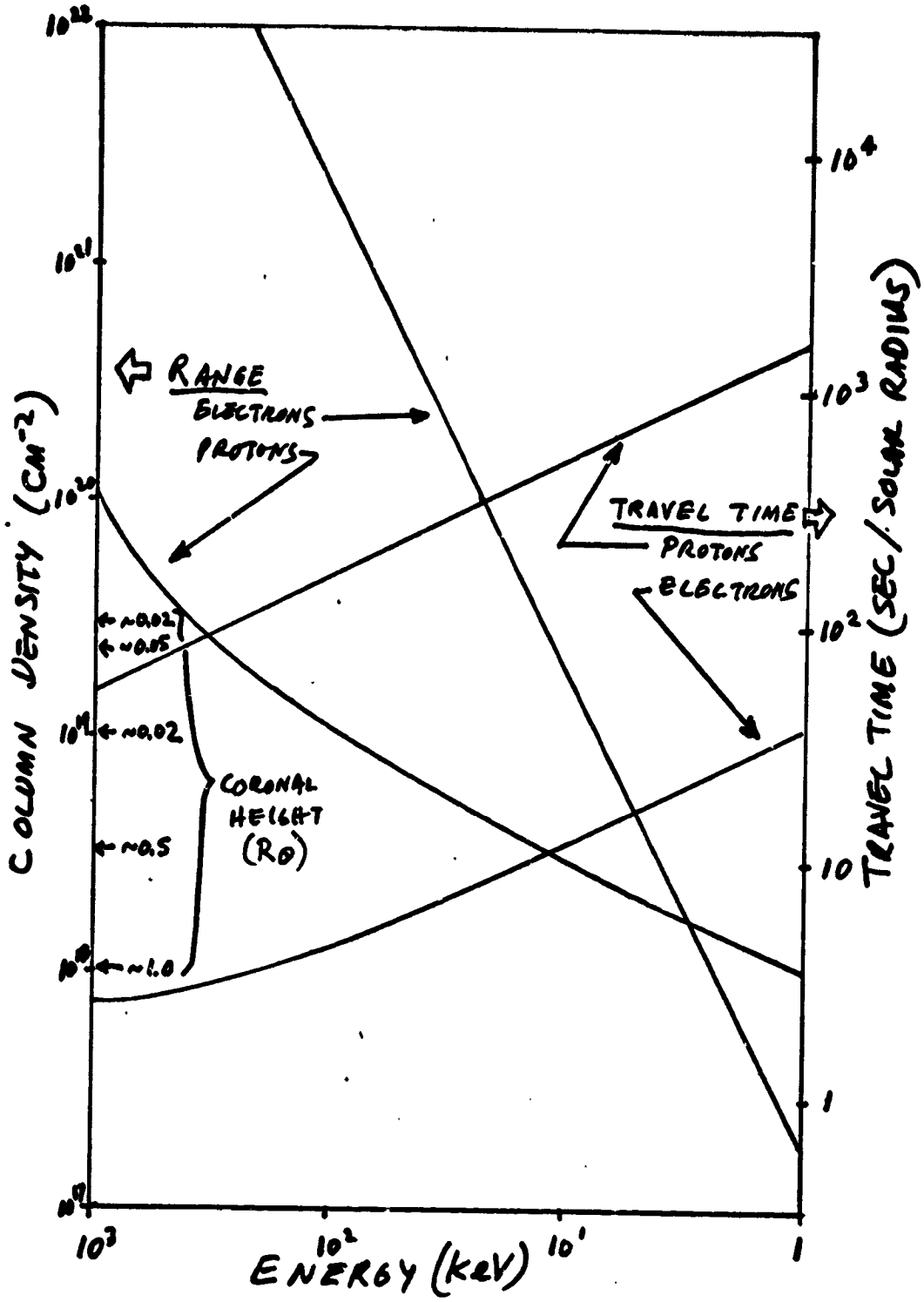


Figure 8. Plotted here are the particle range in column density (cm^{-2}) and the time to travel 1 $R_\odot$ for electrons and protons.

References

Datlowe, D. and Lin, R.P. 1973, Solar Phys. 32, 459.

Fainberg, J. and Stone R.G. 1970, Solar Phys. 15, 222.

Kane S.R. 1974, in I.A.U. Symposium 57, Coronal Disturbances (ed.) Gordon Newkirk, Jr., D. Reidel, 105.

Lin, R.P. 1974, Space Science Rev. 15, 189.

Lin, R.P., Anderson, K.A., and Cline, T.L. 1972, Phys. Rev. Letters 29, 1035.

Malitson, H.H., Fainberg, J., and Stone, R.G. 1973, Astrophys. Letters 14, 111.

Stone, R.G. and Fainberg, J. 1971, Solar Phys. 20, 106.

D19

N78-32983

PROBING THE SOLAR CORONA WITH ENERGETIC PARTICLES

M. A. I. Van Hollebeke*
Laboratory for High Energy Astrophysics
NASA/Goddard Space Flight Center
Greenbelt, Maryland 20771

Abstract

Fundamental questions about solar flare phenomena can be answered by the study of energetic particles in space. In this paper, representative sets of recent experimental results and theoretical studies are very briefly summarized to show the current knowledge of energetic particle phenomena based mainly on observations near 1 AU. It is shown that information on the properties of the source region and acceleration processes deduced from observations of particles with energy > 1 MeV is essentially limited by a lack of knowledge of the coronal storage and particle escape into the interplanetary medium. Measurements close to the sun of quantities such as particle spectra, charge and isotopic composition, and the flow of particles as a function of time, together with measurements of the plasma, magnetic field, X-rays, and γ-rays will not only reveal the properties of the source region and acceleration mechanisms but will also provide a better knowledge of the physics of the corona. In addition, these measurements may reveal the existence of large scale inhomogeneities in the solar corona.

The general requirements for making these measurements on the Solar Probe are listed. An energetic particle detector system which has the capability of resolving the charge and isotopic composition up to 50 MeV/nucleon may successfully measure the above quantities provided that it can perform under both very low and average intensity levels. The time of encounter with the Sun with respect to the solar cycle as well as the total time spent in the very inner solar system are important factors in the probability of observing a solar-flare-associated event. Alternative measurements in the event of non-flare conditions are also listed.

*Also University of Maryland, College Park, Maryland 20742

I. Introduction

Observations of energetic particle populations provide information on the nature of the source region as well as on the acceleration, injection, and transport processes. For solar cosmic rays, the charge composition and, to a lesser extent, the energy spectra measured at 1 AU can provide information on both the properties of the source region and the nature of the acceleration process. Isotopic composition measured at 1 AU can also give information on the dynamics of the source region. However, the quality of the information on those processes deduced from observations at 1 AU is strongly dependent on the distortion of the observed quantities by the injection process and storage. Little is known about the nature of the injection process and possible coronal storage. Studies of the time histories of flare associated events observed at 1 AU have revealed the existence of a "fast propagation region" extending up to 60° from the flare site and a coronal transport process that is not strongly rigidity dependent.

Models have been developed to account for those observations. However, the exact nature of the injection process, the physical mechanism of coronal propagation, as well as possible solar coronal storage are strongly dependent on the detailed knowledge of the properties and structure of the coronal magnetic field. These quantities can only be determined by measurements very close to the Sun.

II. Observations in the Interplanetary Medium; Results and Limitations

Below ∿20 MeV/nucleon, a great variety of particle events have been observed in the interplanetary medium. They tentatively can be grouped in five categories: (a) flare associated particles; (b) corotating particles associated with the interaction regions between low-speed and high-speed solar wind streams; (c) planetary particles accelerated in the magnetospheres of planets such as Earth or Jupiter; (d) energetic stor_ particles seen in conjunction with very strong interplanetary shock waves; (e) a more or less continuous solar component which may come from either the superposition of many small events or the frequent injection from several active regions.

Above ∿20 MeV/nucleon, the observed particle population is mainly accelerated in solar flares.

As one approaches the sun, it is expected that interplanetary and planetary particle events will diminish in intensity, as observed by Helios 1 and 2 for

corotating events (Van Hollebeke et al., 1978), while the flare associated particle events will be the dominant component together with the low energy continuous solar component which would be expected to become well resolved.

In this section, the information provided by studies of flare associated particle events observed in the interplanetary medium are briefly reviewed.

Figure 1 shows several particle intensity profiles detected at 1 AU during solar cycle 20. These time histories are typical of the observations of different particle species in interplanetary space. They represent the combined effects of source properties, acceleration, coronal transport, release into the interplanetary medium, and interplanetary propagation.

A major step in understanding the characteristics of the propagation has been made by statistical analyses of parameters associated with the particle intensity profiles, e.g.: the particle anisotropy as a function of time; the event visibility as a function of the position of the flare relative to the observer; and the heliolongitude distribution of the onset time and rise time to maximum intensity. In addition, better knowledge of both coronal and interplanetary transport has been obtained by studying the variation of the spectrum as a function of the flare position relative to the observer or by fitting the observed time intensity profile with a density profile calculated from various models of interplanetary propagation and coronal transport.

Figure 2 shows one possible scenario for particle propagation which is based on the above studies and has received the widest acceptance. After being accelerated in the solar flare (black box), particles are either released within about 1 hour after the flare event in a so-called "fast propagation region" or propagate in the corona where they may lose energy before escaping into the interplanetary medium. In this scenario, the "fast propagation region" may extend up to 40-60° from the flare site and has nearly uniform access to interplanetary field lines.

The propagation in interplanetary space is weakly dependent on the particle rigidity (Reinhard and Wibberenz, 1974; Ma Sung et al., 1975) and can be well described by focussed diffusion (Earl, 1976). The propagation in the corona is nearly independent of energy and rigidity (e.g., Ma Sung, 1977) and the particle escape rate into the interplanetary medium has a weak dependence on velocity (Van Hollebeke, Ma Sung and McDonald, 1975).

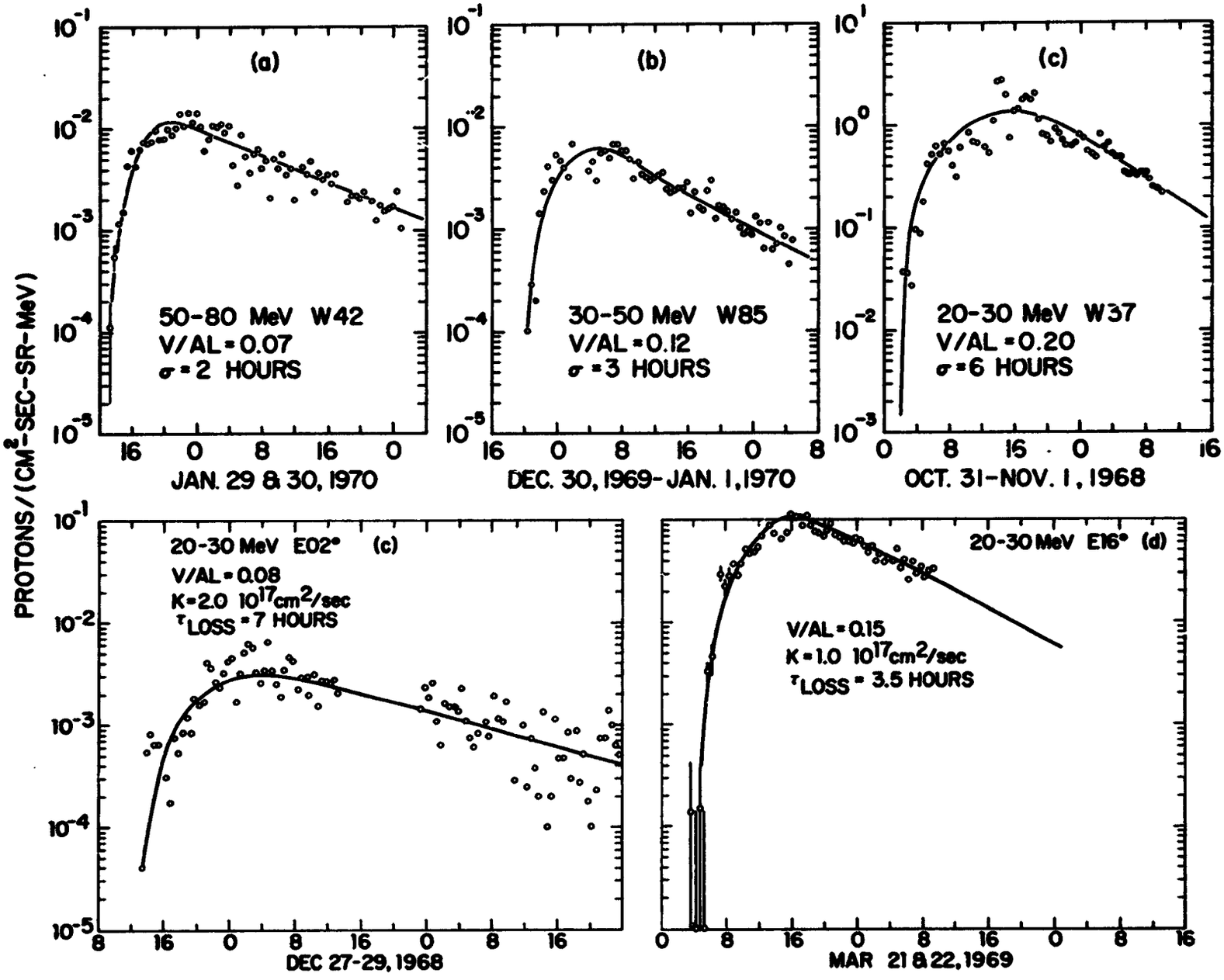


Figure 1. A few examples of proton time histories detected at 1 AU by the GSFC cosmic ray telescope on IMP 4/5. The solid curves are the best fit of the data using a model which includes focused diffusion in the interplanetary medium, indicated by the parameter V/AL, a gaussian injection profile described by σ and a coronal propagation and escape time described respectively by K and τ_{LOSS}. (Ma Sung, 1977.)

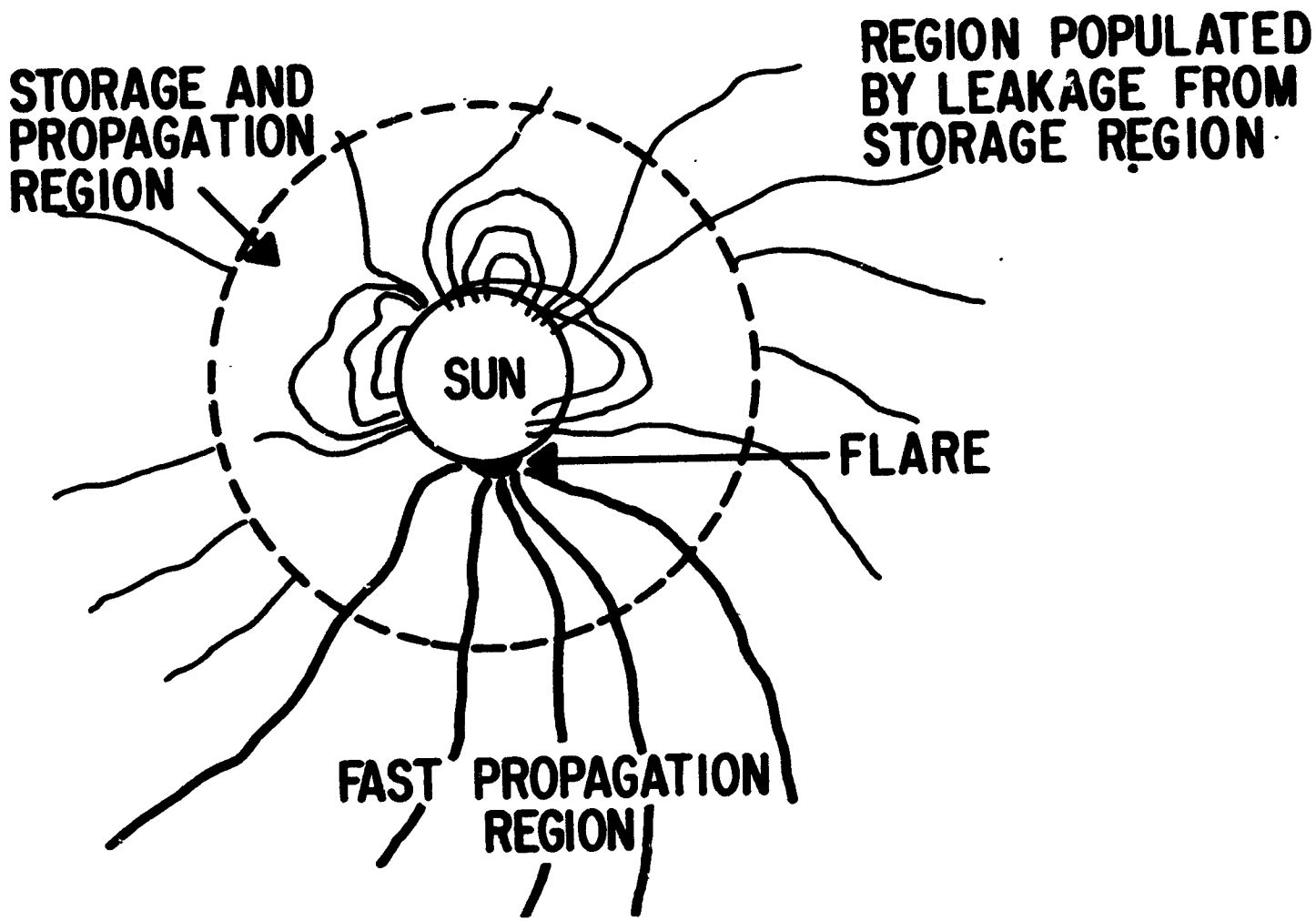


Figure 2. One possible scenario for particle propagation based on observations at 1 AU. After being accelerated in the solar flare (black box), particles are either released in the fast propagation region or propagate in the corona where they may lose their energy before escaping into the interplanetary medium.

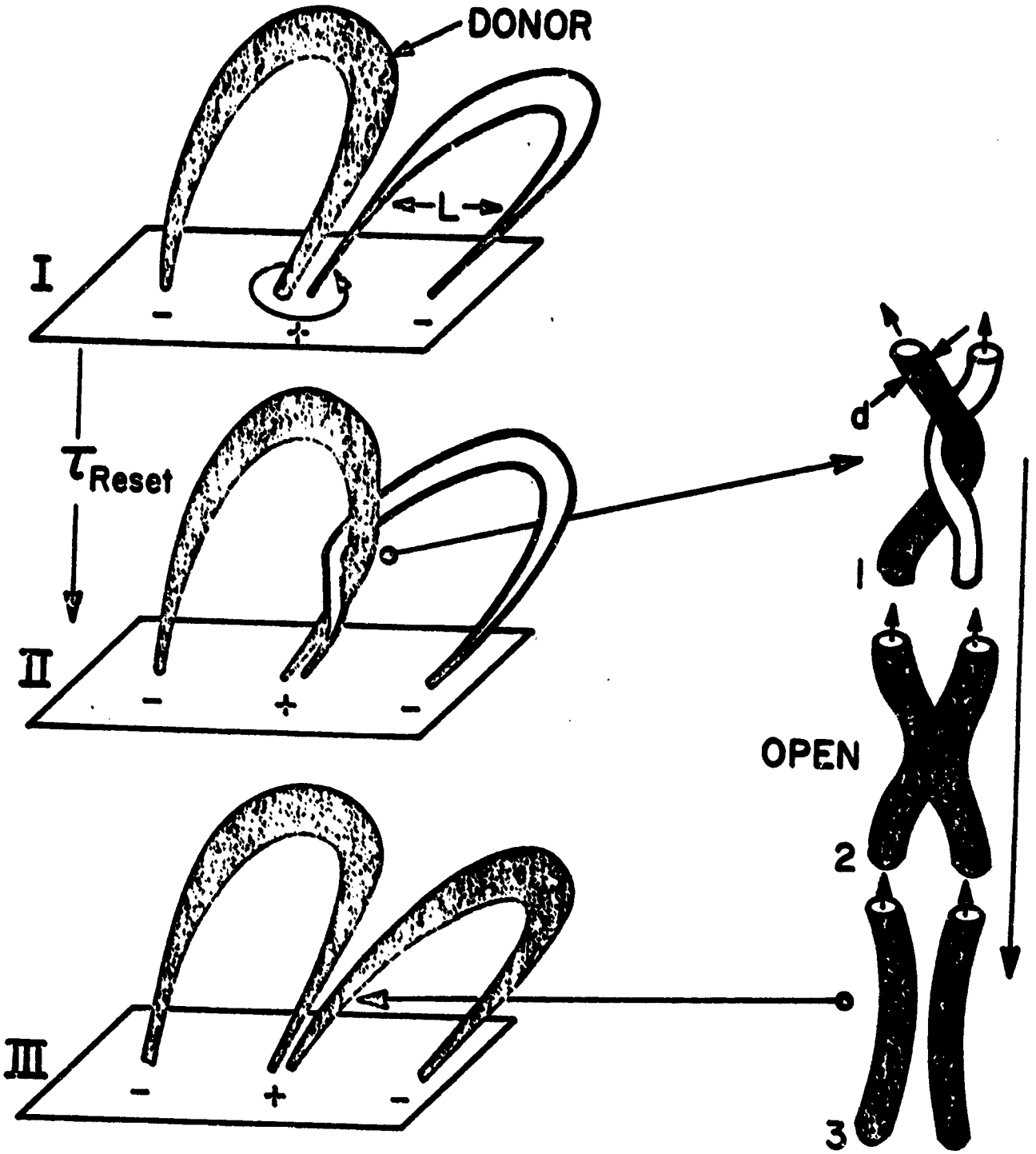
This sketchy picture, which summarizes the most recent observations of particle propagation, has given important clues to understanding coronal propagation, particle release, acceleration and source properties. However, as we shall see below, there are a number of limitations which are expected to be resolved with observations very close to the Sun.

A. Coronal Propagation

The requirement of a lack of rigidity dependence of the particle transport places a serious constraint on any propagation model, since it argues strongly against all the proposed diffusion-type models. Recently, Newkirk and Wentzel (1978) have suggested a "bird-cage" model which reproduces the observed rate of azimuthal propagation in a rigidity-independent manner. In this model, the particles are confined to extensive magnetic flux loops for a certain time and are then discontinuously transferred to adjacent flux tubes by field-line reconnection produced by rearrangement of the field in the super-granulation network flow (Figure 3). For this model to work, it is essential that the time scale of reorganization of the field be greater than the time of merging between two flux tubes which, in turn, is greater than the time for the slowest particle to traverse a flux tube. While this model is attractive, the nature of the reconnection process and escape remains unclear. Other details, such as energy loss, still need to be worked out. Magnetic loops are known to extend up to 4-5 $R_\odot$. It is expected that a cosmic ray detector on a spacecraft such as the Solar Probe could sample both the stored and escaping particles.

B. Particle Escape

The release of particles in the "fast propagation region" must take place in a limited time interval. As shown by the mass ejections associated with flare particle events (Kahler et al., 1978), the release seems to be catastrophic in nature. Schatten and Mullan (1977) have proposed that coronal magnetic bottles, produced by flares, serve as temporary traps for solar cosmic rays (Figure 4). The expension of these bottles allows fast azimuthal propagation over a region up to 40-60° from the flare site independent of the particle energy. As the result of the Rayleigh-Taylor instability, which occurs when ρ_1, the density inside the bottle becomes smaller than ρ_2, the density of the gas being pushed ahead of the bottle, cosmic rays originally trapped inside the bottle are released by field line interchange. Schatten and Mullan suggested that the energy losses due to the bottle expansion can be offset by cosmic ray acceleration by second-order



$$\tau_{OPEN} = \frac{d}{fV_A}$$

Figure 3. A schematic diagram of the twisting and reconnection of a pair of flux tubes which allows particle transport in the corona as proposed by the "bird cage" model of Newkirk and Wentzel (1978).

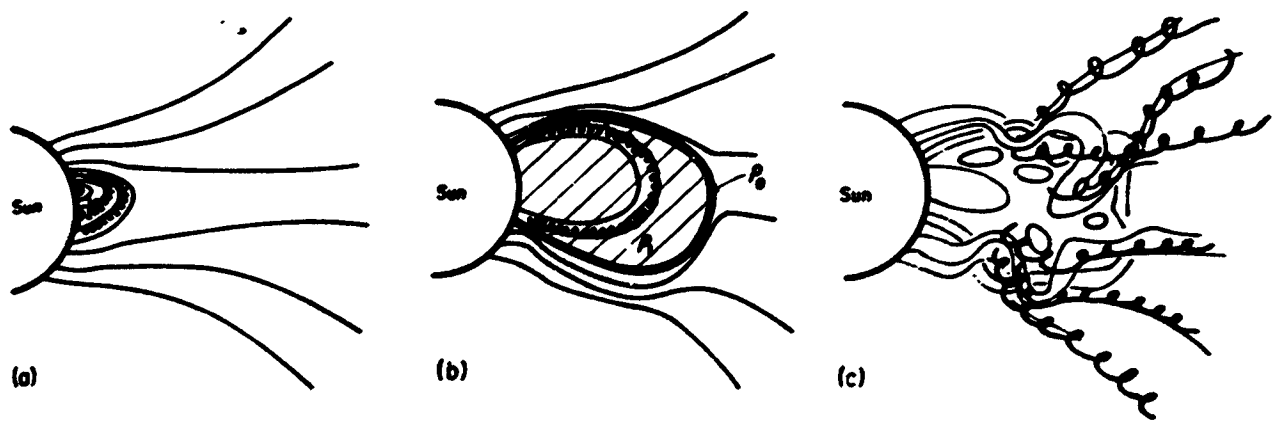


Figure 4. A schematic diagram of Schatten and Mullan's (1977) model for the fast escape of energetic particles from coronal magnetic battles.

Fermi processes. In this case, acceleration may last until either the shocks disappear or trapping of the particles is no longer effective.

From such a model, it is obvious that complex processes, which still remain to be studied, may occur during particle release. These processes may significantly distort the properties of the original particle population accelerated in the flare region. Observations close to the sun of the structure and evolution of the coronal magnetic field and of energetic particles and plasma are expected to resolve the problem of particle escape.

C. Acceleration and Properties of the Source Region

Information on the energy spectra of protons and electrons and their ratios can give important clues about the acceleration mechanism. Particle spectra are also an essential complement to the information provided by X-rays and gamma-rays. It has been shown (Lin, 1974; Van Hollebeke et al., 1975) that interplanetary propagation effects are minimized if the energy spectrum is measured in interplanetary space at the time of maximum particle intensity. Furthermore, it has also been shown that the particle distribution measured in the "fast propagation region" where coronal transport can be neglected should closely reflect the particle distribution in the flare site provided the release process does not significantly distort the spectrum. The energy spectra of protons, electrons, and a few other particle species have been determined under these conditions. From $\sim$10 MeV to 80 MeV, the proton spectrum is well described by a power law in kinetic energy with the spectral index γ (defined as $dJ/dE \propto E^{-\gamma}$) ranging between 2 and 3.2 (Van Hollebeke et al., 1975); however, below 10 MeV, the observed spectrum is generally flatter (Figure 5). This indicates either a flattening intrinsic to the acceleration mechanism or different propagation conditions at low and high energies. Together with X-ray and γ-ray measurements, direct measurements of the energy spectra of the different particle species close to the Sun are necessary to understand the acceleration mechanism.

In addition to the energy spectra, charge composition also provides data on the acceleration processes. Furthermore, charge composition reflects properties of the source region which are difficult to measure by other methods. It has recently become obvious that in many instances the energetic particle composition can depart drastically from the expected composition of the solar atmosphere and can show large variability from event to event (Figure 6). These enhancements

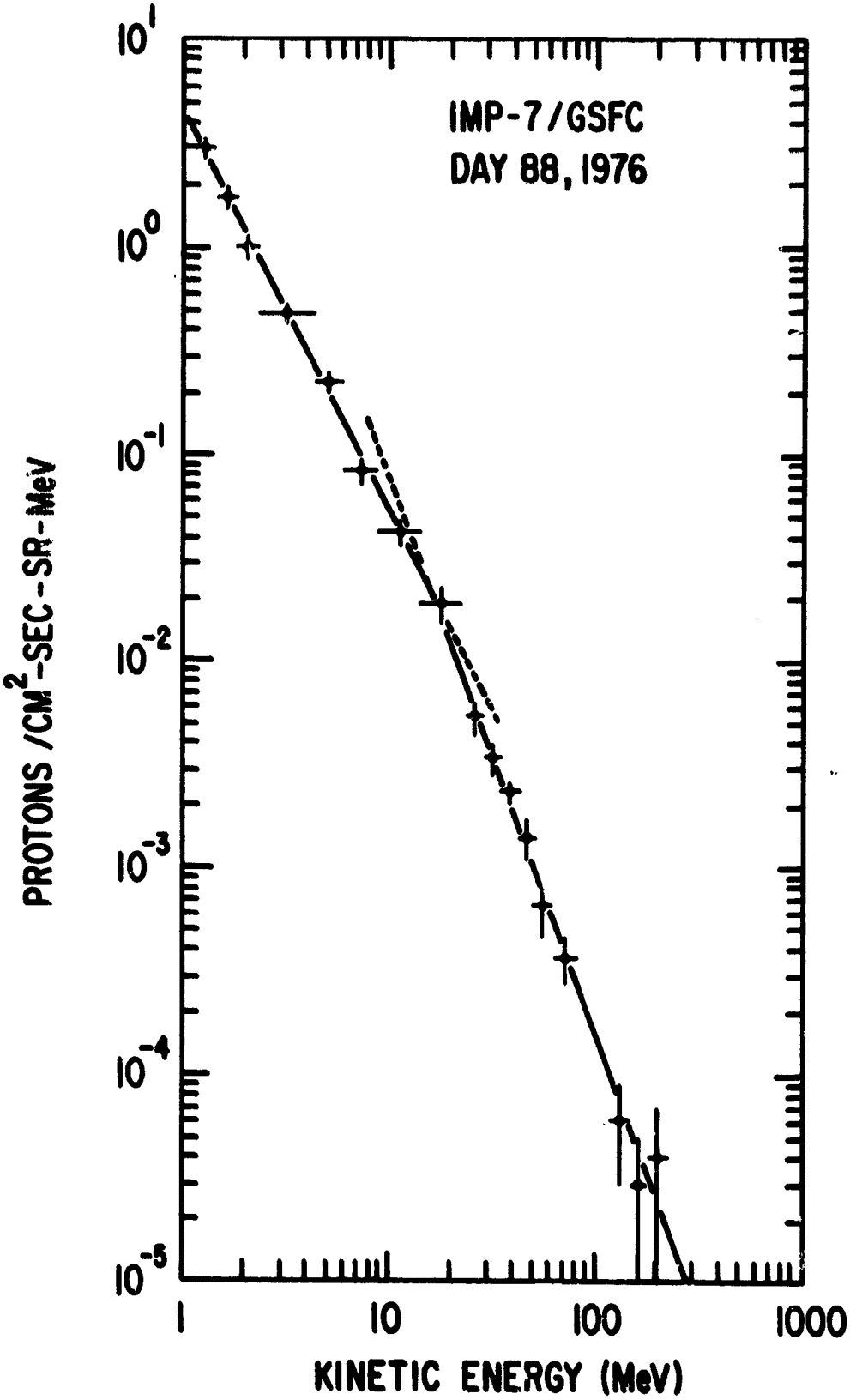


Figure 5. A characteristic proton energy spectrum obtained at the time of maximum particle intensity.

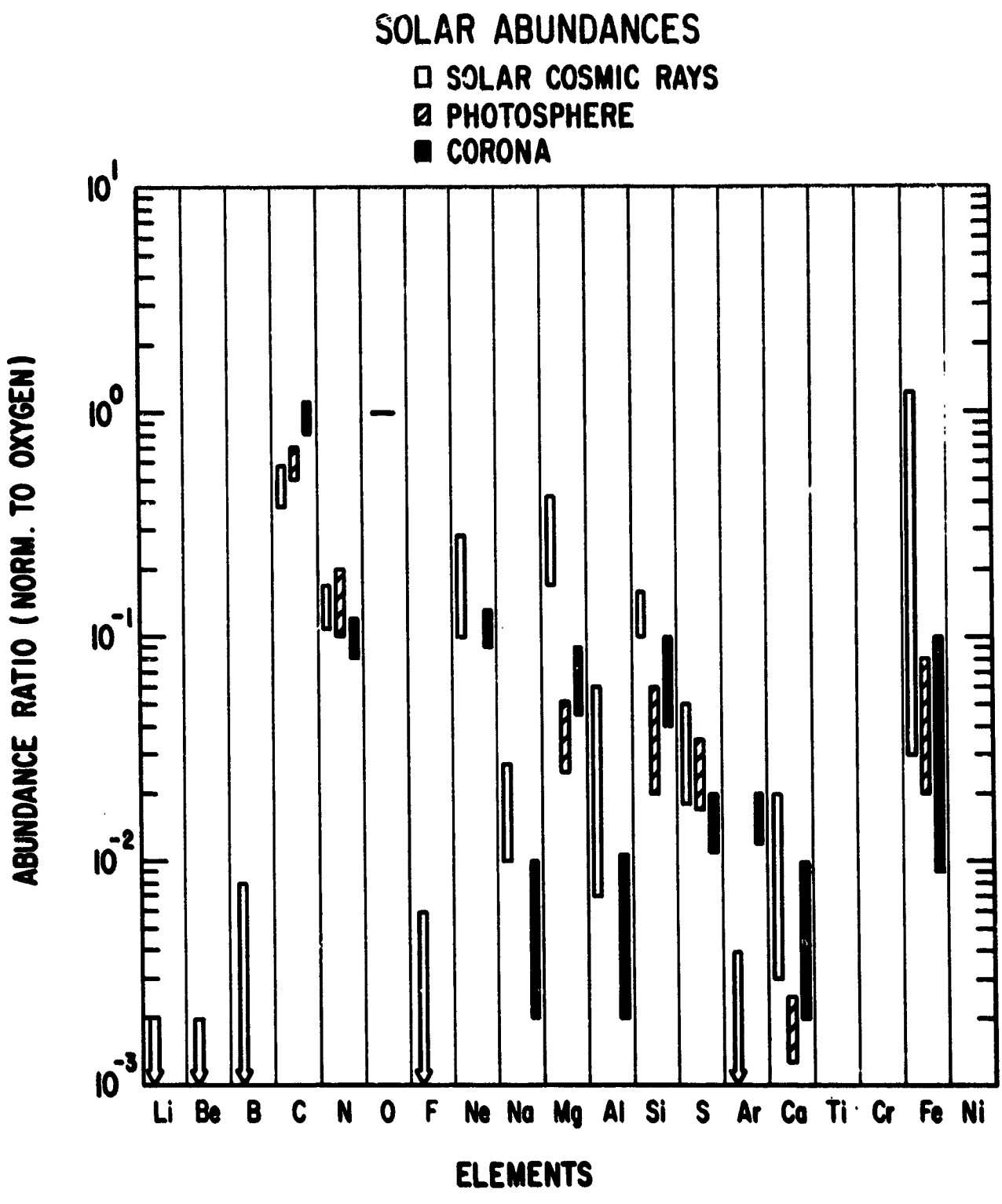


Figure 6. A comparison of solar cosmic ray, photospheric and coronal abundances from McDonald et al. (1974), updated with more recent observations of the solar cosmic ray composition.

are not limited to low energies but have been observed at energies up to 200 MeV/nucleon, in the case of Fe/O during the September 24, 1977 event (Dietrich and Simpson, 1978). It is not clear how these differences arise. They can be due either to inhomogeneities in the source region, acceleration processes, or spallation in the acceleration region. In addition, low energy particles may be further accelerated in the interplanetary medium.

Besides providing definite information on solar abundance, measurements of these quantities close to the Sun, can be compared to X-ray and γ-ray measurements to determine the timing of the high energy particle acceleration, the energy content in protons and nuclei, and the type of emission or nuclear reactions taking place at the surface of the sun.

III. General Requirements

The main quantities to be measured to provide information on the properties of the source region, acceleration, particle escape and coronal propagation are the energy spectra and charge and isotopic abundances of energetic particles as functions of time and direction. Since timing is an important factor in understanding acceleration processes, the cosmic ray detector should have sufficiently high sensitivity to be able to measure the onset time of any flare-associated particle event within a fraction of a minute, in order to compare with X-rays, γ-rays and radio events. This is particularly true for protons and electrons.

The probability of occurrence of a flare-associated particle event is strongly dependent on the solar cycle. Figure 7 shows the number of >20 MeV proton events with peak intensity >0.5 protons/cm^2-sec-ster detected per year at 1 AU as a function of solar activity as indicated by sunspot number for solar cycle 20. At the lower energies and lower intensity levels, it is often difficult to associate a particle event observed at 1 AU with a solar flare. Closer to the sun, because the decay time is shorter and the particle event is better defined in time, the number of events > 10 MeV with peak intensity > 10^{-4} particles/cm^2-sec-sr-MeV is expected to be a factor of 3 to 5 greater than that in Figure 7. From the point of view of event probability, it is therefore desirable that the Solar Probe reach perihelion during solar maximum.

However, even during solar minimum, a number of active regions exist which produce subflares at a rate of 1-2 per day. These subflares may contribute to the continuous low energy solar particle component observed in the interplanetary medium and their effects may be resolved at a closer distance to the sun. Also, even though corotating particle streams are believed to result from acceleration in the interplanetary medium, the origin of the accelerated ions is not known.

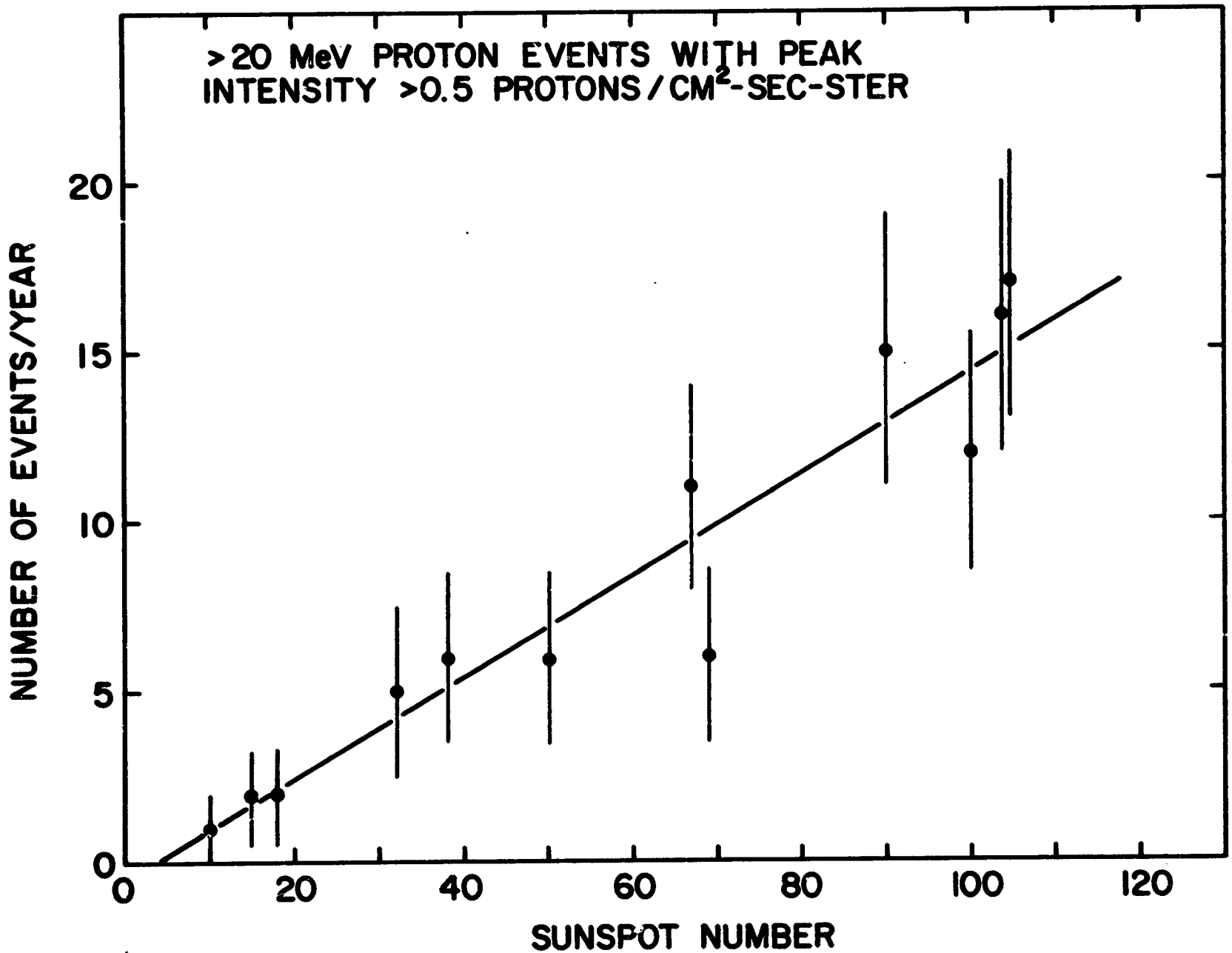


Figure 7. Number of > 20 MeV proton events with peak intensity > 0.5 protons/cm^2-sec-ster detected per year at 1 AU as a function of sunspot number for solar cycle 20.

These events, which are mainly observed during solar minimum, are associated with high-speed solar wind streams and may originate either from the tail of the supertherma1 distribution of the plasma or from very low energy particles. In any event, the origin of these particles can be found close to the sun.

In order to know the intrinsic characteristics of solar particle events, it is important either to sample a significant number of events (this requires measurements in the very inner solar system for a long time period) or else to make measurements simultaneously close to the sun and in the interplanetary medium (for example 1 AU).

References

Dietrich, W. F. and J. A. Simpson, 1978 to be published.

Earl, J. A., Astrophys. J. 205, 900, 1976.

Kahler, S. W., E. Hildner and M. A. I. Van Hollebeke, Solar Physics 1978 (in press).

Lin, R. P., Space Science Rev., 16, 189, 1974.

Ma-Sung, L. S., 1977, Ph.D. Thesis, University of Maryland, NASA/GSFC X-660-77-113.

Ma-Sung, L. S., M. A. I. Van Hollebeke and F. B. McDonald, Proc. 14th Int. Cosmic Ray Conf., Munich, 5, 1767, 1975.

McDonald, F. B., C. E. Fichtel and L. A. Fisk, 1974, High Energy Particles and Quanta in Astrophysics, ed. F. B. McDonald and C. E. Fichtel, p. 212.

Newkirk, G. Jr. and D. G. Wentzel, J. Geophys. Res., 83, 2009, 1978.

Reinhard, R. and G. Wibberenz, Solar Phys. 36, 473, 1974.

Schatten, K. H. and D. J. Mullan, J. Geophys. Res. 82, 5609, 1977.

Van Hollebeke, M.A.I., L. Ma-Sung and F. B. McDonald, Solar Physics, 41, 189, 1975.

Van Hollebeke, M.A.I., F. B. McDonald, T. von Rosenvinge and J. H. Trainor, NASA/GSFC TM 78040, J. Geophys. Res. 1978 (in press).

D20

N78-32984

CORONAL PROPAGATION AND STORAGE AT ENERGIES ~ 1 MeV/NUCLEON

E. C. Roelof
S. M. Krimigis
R. E. Gold
Applied Physics Laboratory
The Johns Hopkins University
Laurel, Maryland 20810

ABSTRACT

Representative energetic particle events relating to quasi-stationary particle populations, solar flare events, and solar composition are discussed in the context of similar observations likely to be made during Solar Probe encounter. Particular emphasis is given to inter-relationships with coronal magnetic structure and plasma processes. General instrumental requirements are listed, and it is concluded that energetic particle detectors meeting these requirements will successfully complement other essential experiments on the Solar Probe.

1. INTRODUCTION

We present selected examples relevant to the Solar Probe mission from studies by our group at JHU/APL and colleagues at collaborating institutions of energetic particles and coronal structure during the decline and minimum of Solar Cycle 20 (1973 - 1976). The Jupiter-assisted Solar Probe trajectory will yield encounter in 1989 or 1990, almost the same phase of Solar Cycle 22 as that in which we are right now for Solar Cycle 20. We are currently finding occassional events in our 1977 - 1978 IMP 7, 8 and Voyager 1, 2 data that are similar to those from 1973 - 1976 which we are going to discuss. We therefore regard these examples as being indicative of the types of energetic particle phenomena which may be measured by the Solar Probe inside 1 AU, and we shall attempt to extrapolate the characteristics we have deduced to the perihelion environment.

2. QUASI-STATIONARY PARTICLE FLUXES AND CORONAL STRUCTURE

The unprecedented radial and angular velocities which will be achieved by Solar Probe require us to transform our earth-based perception of the relationship between spatial structure and temporal evolution of solar particle fluxes. With the spacecraft spending less than 10 days inside the orbit of Mercury and transiting an entire hemisphere of the sun in less than 14 hours, quasi-stationary particle distributions will produce rapidly varying fluxes as measured on the spacecraft.

The first example, taken from a study by *Levine and Roelof (1977)*, is a broad recurrent minimum in 1.9 - 4.5 MeV protons centered over a growing coronal hole (CH2) in August, 1973. Figure 1 shows the hole at the center of the disk, surrounded by low coronal magnetic neutral lines inferred from absorbtion features in Hα filtergrams *(McIntosh, 1975)* over which lie loop-like soft x-ray emission features evident in the American Science and Engineering, Inc. photographs from their spectrographic telescope on Skylab. Such a ring-like loop arcade suggests open field lines above the holes, and this is confirmed by the potential-field calculations of coronal field lines by *Levine et al.*,

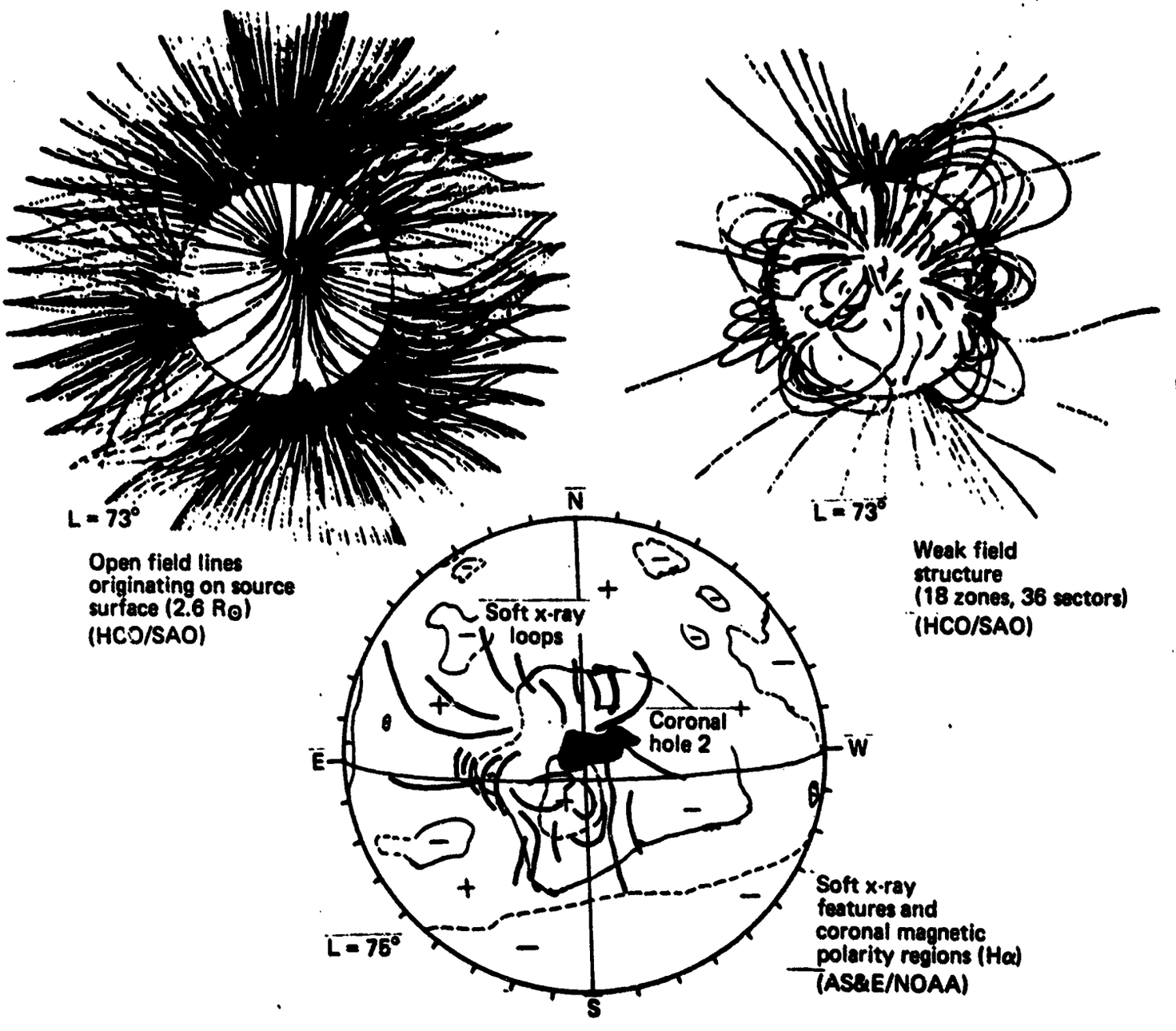


Figure 1 Coronal structure in the vicinity of a small equatorial coronal hole *(Levine and Roelof, 1977)*. Center: Hα inferred neutral lines and soft x-ray loops. Upper left: open field lines originating on a 2.6 $R_\odot$ source surface calculated from a potential field. Upper right: closed field lines from the same calculation.

(1977). Open field lines originating on the "source surface" at 2.6 $R_\odot$ are shown on the left, and most of those in the visible hemisphere lead down into CH2. On the right, the calculated closed field lines show general agreement with the x-ray loop ring arcade.

Hourly-averaged ~ 3 MeV proton fluxes from the JHU/APL detector on IMP 7 are shown in Figure 2. They have been "mapped back" to their estimated high coronal emission longitudes assuming the large-scale interplanetary magnetic field was "frozen-into" a radially propagating solar wind. Hourly averages of solar wind speed were taken from the MIT plasma observations on the same spacecraft. For reference, the same $H\alpha$ inferred neutral line structure and x-ray coronal hole boundary shown in Figure 1 are plotted again on the same longitude scale below the particle fluxes in heliographic coordinates. The proton flux minimum is some 2 orders of magnitude below the adjacent flux distributions and spans more than 100° in heliolongitude. We know this is a quasi-stationary particle population because similar profiles were observed over these longitudes on two previous rotations (1601 and 1603).

If a similar spatial structure happens to lie under the Solar Probe perihelion, one will measure a hundred-fold flux decrease in only 4 hours, followed by an equally rapid recovery. These time rates of flux variation are comparable to the rise times of some flare-associated ~ 3 MeV events observed near 1 AU. The desirability should be apparent for monitoring particle populations from 1 AU during Solar Probe encounter in order to identify such quasi-stationary distributions (and thus reduce the possibility for confusion with temporal evolution).

Figure 3 reveals an interesting upper bound on the altitude at which these ~ 3 MeV protons are distributed in the corona. Mapped back solar wind velocities are shown above a montage of AS&E x-ray images for Carrington Rotations 1603 and 1604 *(Nolte et al., 1976)*. The ~ 3 MeV particle minimum was observed on both these rotations from 30° to 120° heliolongitude. The mapped solar wind shows a small 400 - 500 km s^{-1} stream above the growing coronal hole at 75°. This stream sits exactly in the center of the particle minimum. If we accept the contemporary field models for coronal holes with solar wind streams above them, the coronal magnetic field lines should be predominantly open, at least near

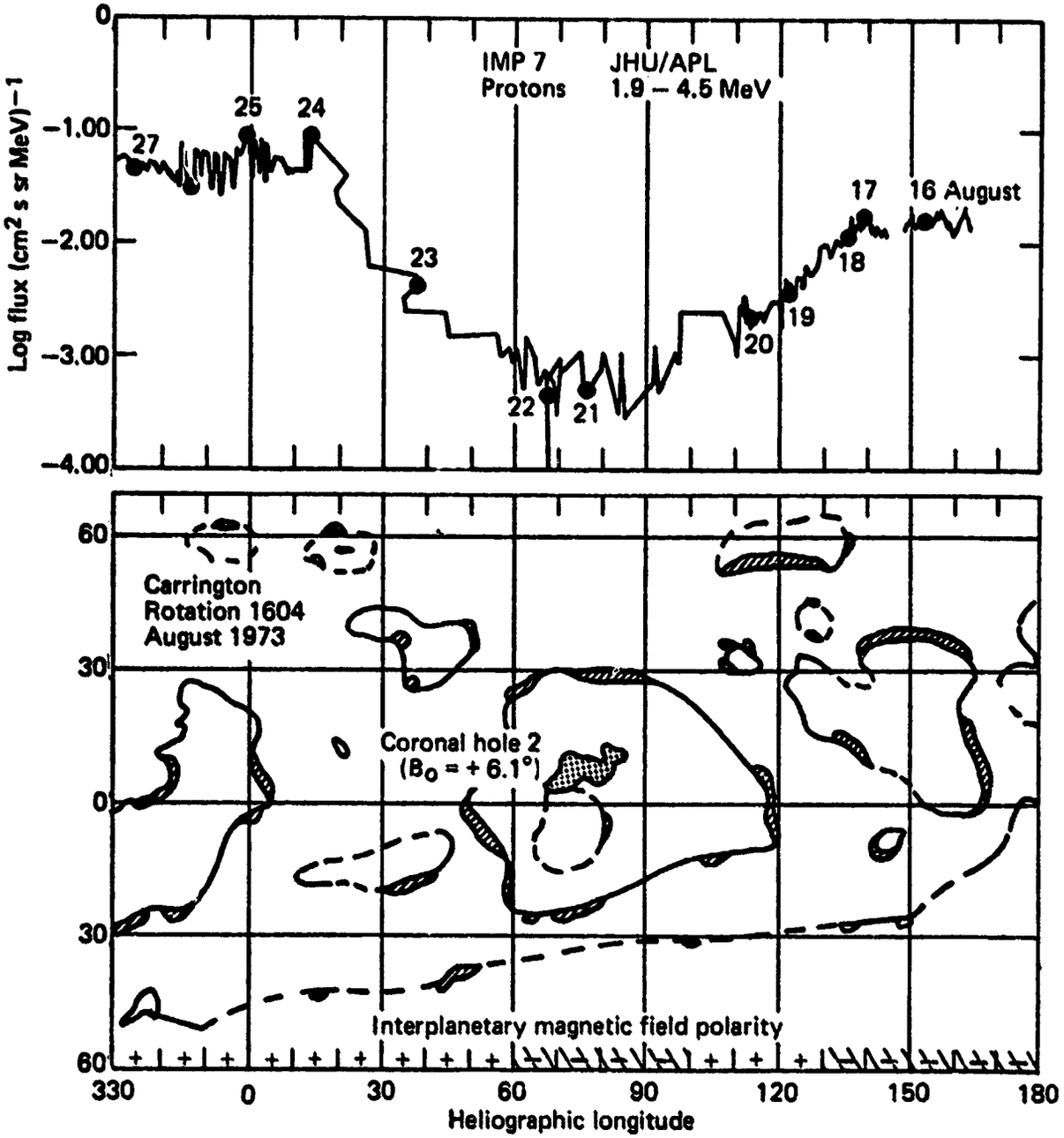


Figure 2 Quasi-stationary energetic particle population above a coronal hole in August, 1973. Lower: Hα absorbtion feature and coronal hole (from Figure 1) in heliographic coordinates. Upper: 1.9 - 4.5 MeV hourly-averaged proton fluxes "mapped back" to their estimated high coronal emission longitudes using simultaneous measurements of solar wind velocity on the IMP 7 spacecraft *(Levine and Roelof, 1977)*.

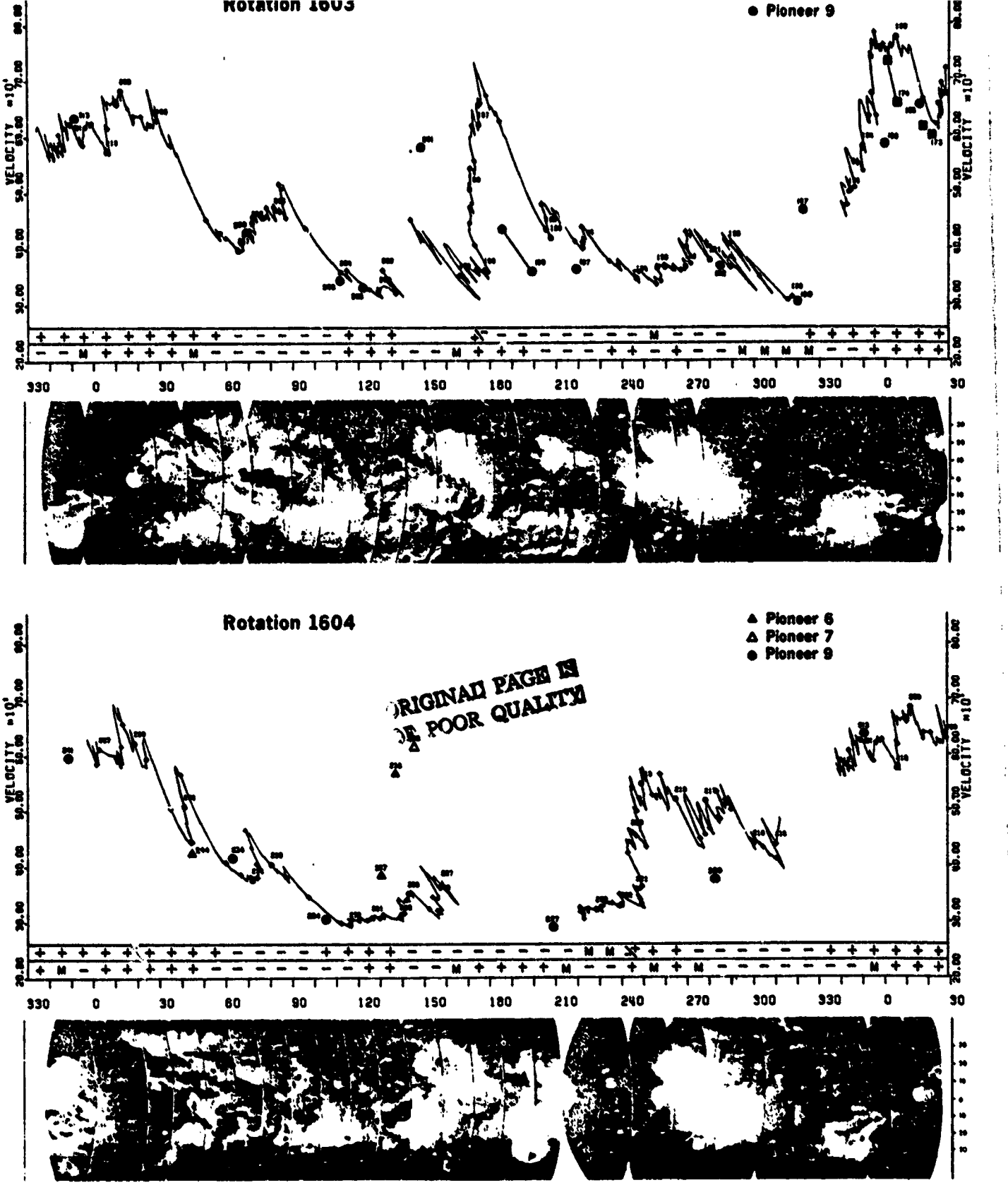


Figure 3 Mapped-back solar wind velocities (MIT) and montage of soft x-ray photographs (AS&E) from Skylab, July-August 1974 *(Nolte et al., 1976)*. Note the small (400 - 500 km s^{-1}) solar wind stream emerging from CH2 (75° heliolongitude) centered within the minimum of the particle distribution shown in Figure 2.

and above the altitude of the MHD critical surfaces. The particles, however, must be distributed on *closed* field lines, and therefore must be transported and temporarily stored well below the altitude of the MHD critical points.

However, the altitude of the critical surface is not a very restrictive upper bound, as may be seen in Figure 4, taken from *Nolte et al., (1977)*. Configuration A was deduced from a stream above a meridional coronal hole, while configuration B resulted from the edge of a stream from the equatorward edge of a polar coronal hole. In their deduction of the three dimensional configuration of the Alfvén surface (i.e., the locus of points where the solar wind bulk speed exceeds the local Alfvén velocity), they followed the range of critical altitude (15 - 50 $R_\odot$) estimated by *Weber and Davis (1967)*. Thus it appears possible that the Solar Probe could possibly cut through coronal storage region of $\leq$ 1 MeV particles if perihelion is < 10 $R_\odot$ (i.e., well below the altitudes of the critical surface).

3. SOLAR FLARE INJECTION AND SOLAR COMPOSITION

The essential point concerning energetic particles from flares is that the injection usually has a finite duration, not only for large events lasting many days, but also for small events lasting $\sim$ 1 day. Figure 5, from a recent review by *Roelof and Krimigis (1977)*, shows the > 0.22 MeV electrons from the large September 10, 1974 flare event measured by the JHU/APL electrons on IMP 8. Similar large, long-lived events are currently being observed, with series in September and November 1977, and April - May 1978, so we could expect them again 11 years from now near Solar Probe encounter. The 1 -11 Å x-ray fluxes plotted in Figure 5d reveal the large flares on September 10 and 13, while the persistent electron anisotropy is displayed in Figure 5b. Anisotropy measurements were interrupted after $\sim$ 50 hours because IMP 8 entered the magnetosheath, but we can infer that electron injection lasted at least this long because the fluxes were continually rising all the while. Shown in Figure 5a for illustrative purpose only are coronal regions of Type IV radio emission (measured in a limb event) which accompany such large electron events.

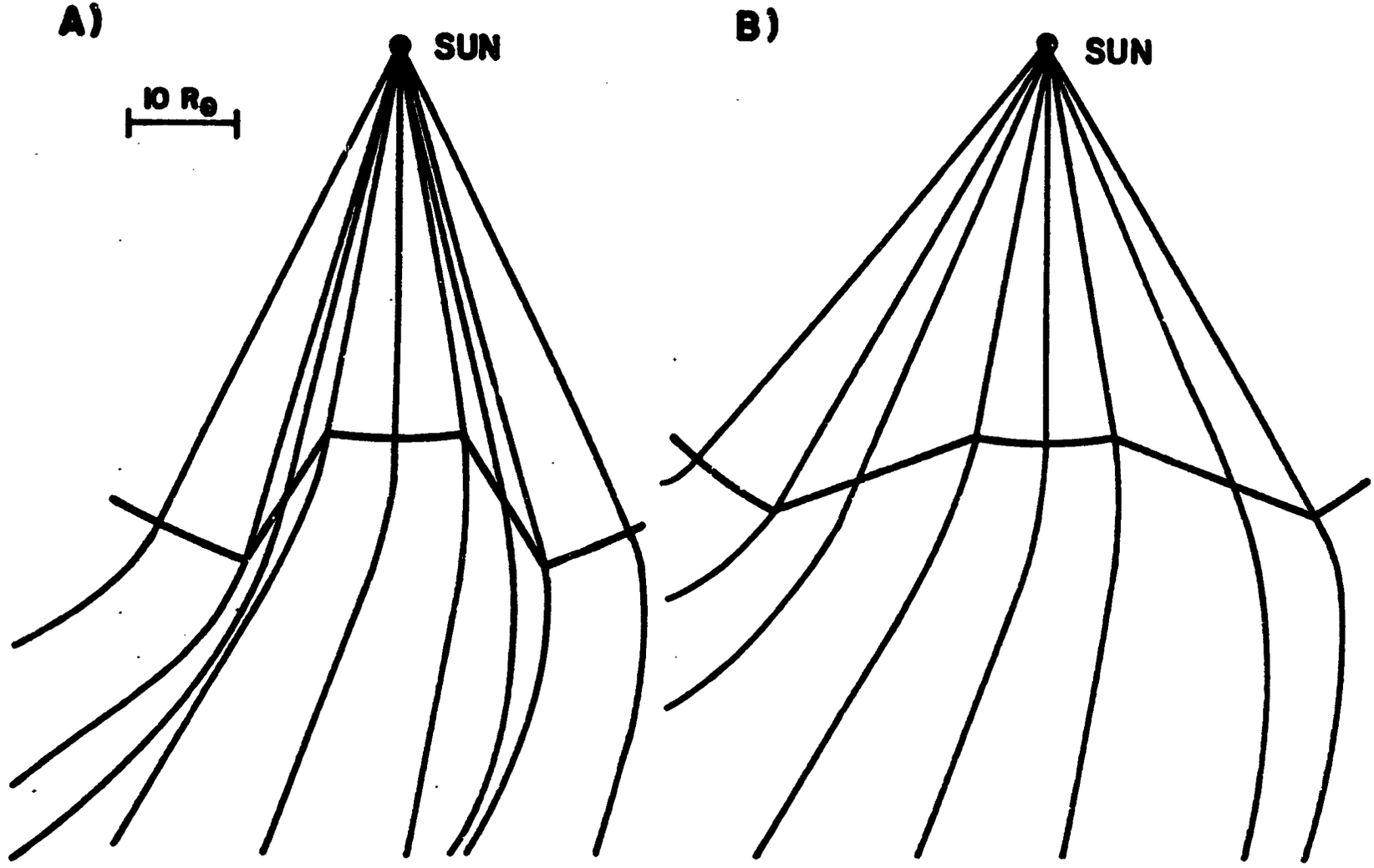


Figure 4 Altitude dependence of the Alfvén critical point estimated from interplanetary data *(Wolte et al., 1977)* for solar wind streams (A) at the edge of the extension of a polar coronal hole, and (B) over an equator-crossing coronal hole. Altitudes are within the range (15 - 50 $R_\odot$) estimated by *Webber and Davis (1967)* and are well above Solar Probe perihelion.

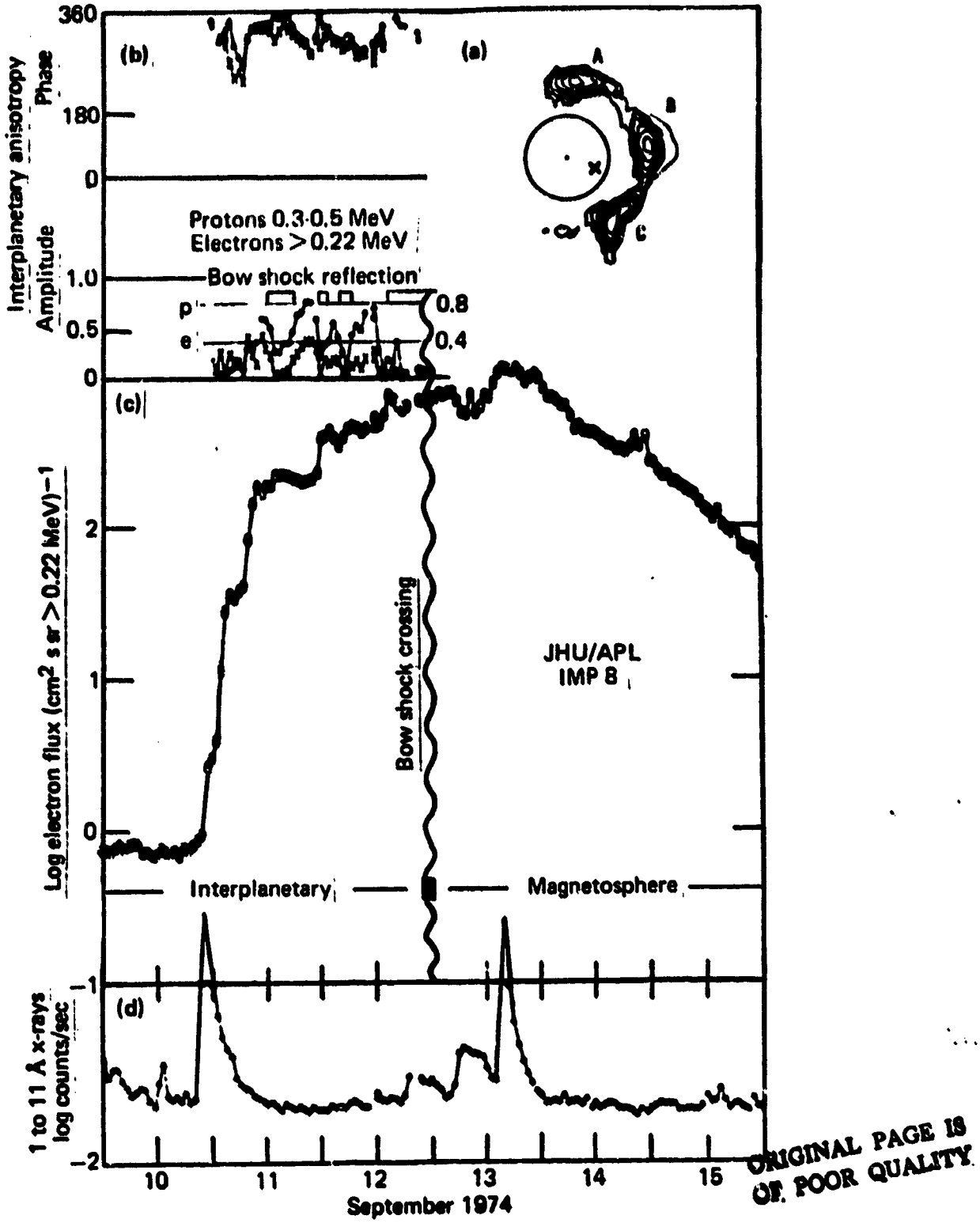


Figure 5 Long-lived injection of >0.22 MeV flare electrons. a) Type IV radio emission regions in the corona typical of the early phases of this class of event; b) electron and low energy proton anisotropy phase and amplitude (distorted by occasional reflection from the bow shock); c) electron intensity, which continues to rise during the persistent anisotropic streaming (anisotropy measurements are interrupted by entrance into the magnetosheath); d) 1 - 11 Å x-ray flux indicating large flares on September 10 and 13. From Roelof and Krimigis (1977).

Extended injection also occurs in smaller events such as the ~ 1 day proton event observed March 28, 1976 on Helios 2 at 0.5 AU. Figure 6, taken from an analysis of interplanetary propagation by *Green and Roelof (1978)* using the "scatter-free" approximation, shows the anisotropy and intensity histories for 4 - 12 MeV and 27 - 37 MeV protons. The theory allows a direct deduction of the coronal injection history, and it can be seen from Figure 6 that the e-folding times are ~ 6 hours for the lower energy channel and ~ 3 hours for the higher, whereas the interplanetary intensities decay with characteristic times only a few times that of the injection. Since there must be strong negative gradients associated with the strongly anisotropic phase of flare events, Solar Probe may sample many more short-duration "micro-events" than are observed at 1 AU. The fact that, during the extended injection phase of such events, the particle flux will be highly collimated along the field lines, means that some provision should be made for measuring field-aligned fluxes, even near perihelion.

As a final example, we point out that rather crucial low energy composition measurements may have to be made in small but highly anisotropic flare-associated events. The fascinating class of ^{3}He-rich and Z-rich events observed during 1972 - 1976 has been associated with small-flare activity from western hemisphere active regions by *Zwickl et al., (1978)*. Figure 7 shows the one-hour average sectored rates from the JHU/APL IMP 7 and 8 experiments for species $Z \geq 3$ at ~ 1 MeV/nuc during four events. The anisotropies are strong and persist up to event maximum, implying continual coronal injection. Very enhanced abundances of elements up to and including iron are observed in these events. After correlating interplanetary and solar data with the Z-rich events, *Zwickl et al.*, concluded that the circumstances conducive to producing these events are those summarized in Figure 8.

The format is an r-φ plot of distance from the sun versus heliolongitude. An Alfvén critical surface similar to those shown in Figure 4 is sketched in near 0.2 AU to indicate the source of a solar wind stream. The Z-rich events (which are usually also ^{3}He-rich) tend to occur near the leading edge of a stream when there is

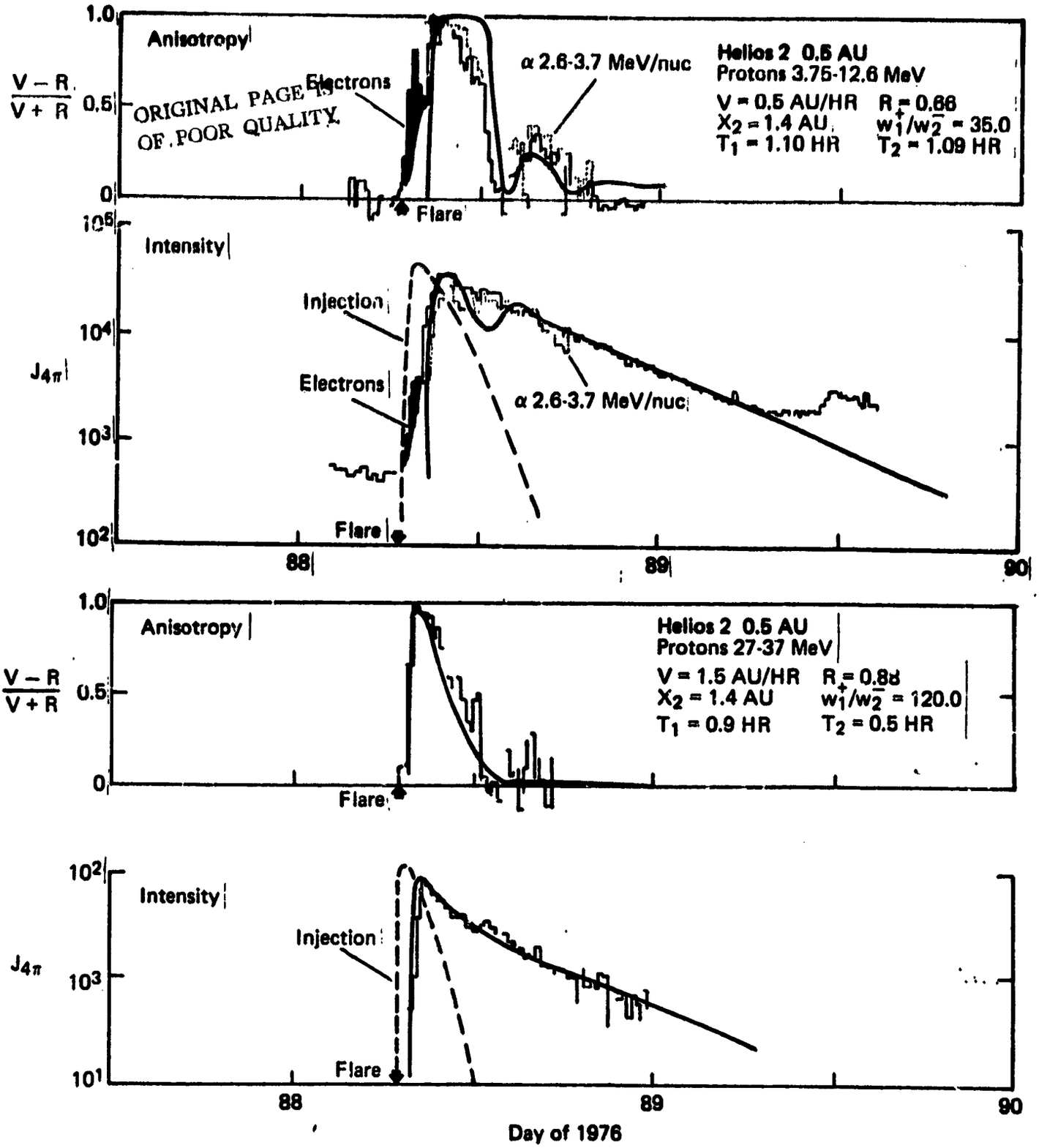


Figure 6 Solar flare particle event observed by the Kiel experiment on Helios 2, March 28 (day 88), 1976. Curves show theoretical fits to "scatter-free" propagation for anisotropy parameter (V-R/V+R) and omnidirectional flux ($J_{4\pi}$). Dashed intensity curves are deduced coronal injection time dependence. From *Green and Roelof (1978).*

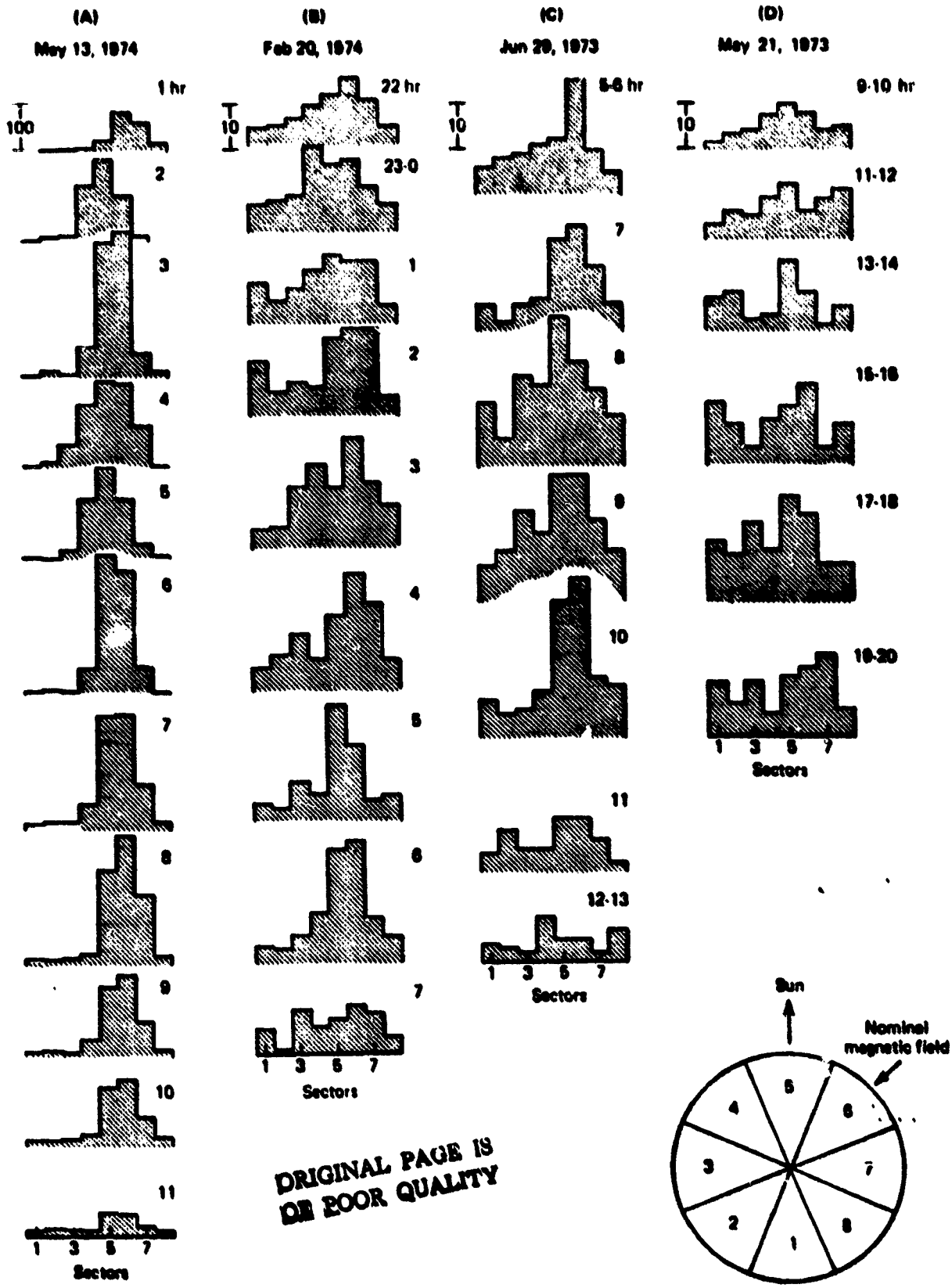


Figure 7 Sector rates (in the ecliptic plane) for four "Z-rich" solar particle events measured on IMP 7 and 8 (*Zwickl et al., 1978*). This class of brief ($\lesssim$ 1 day) event displays strong anisotropies as well as compositional enhancements of heavy ions ($\sim$ 1 MeV/nuc) up through Fe, and of ^{3}He.

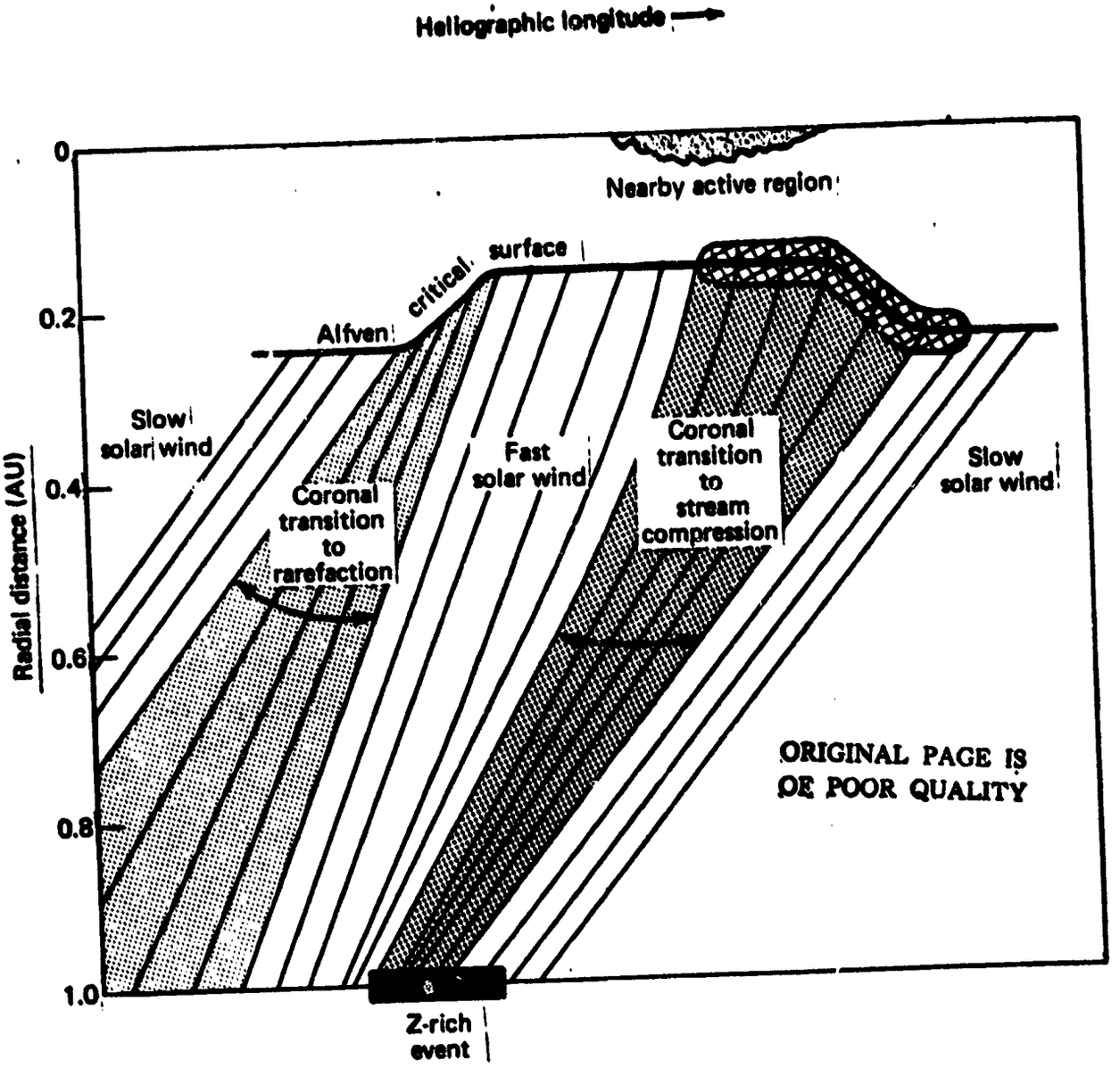


Figure 8 Coronal configuration deduced from correlation of solar and interplanetary observations during Z-rich events which tend to occur near the rise of solar wind streams originating near small active regions with subflare activity *(Zwickl et al., 1978)*. Particle abundances may be enriched prior to flare acceleration in nearly active region, perhaps due to plasma dynamics of western boundary of high speed solar wind source region. Transit of such a structure by the Solar Probe could reveal inter-relationships of plasma heating and energetic ion acceleration.

an active region near the western edge of the solar wind source boundary. *Zwickl et al.*, suggest that the dynamical plasma interactions which occur as fast plasma interacts with the slow plasma to the west of the high speed stream may be instrumental in producing the high-Z enhancements in the substrate plasma prior to flare acceleration. Thus one of the most interesting coronal configurations (from the point of view of solar composition and dynamics), which the Solar Probe could transit may well be the western edge of a solar wind source region.

4. CONCLUSIONS

It seems to us that these examples demonstrate that valuable new information on basic solar processes will be derived from measurements of energetic particles ~ 1 MeV regardless of the state of the corona and solar activity during Solar Probe encounter.

(i) If activity is low, the distribution of quasi-stationary particle populations will be directly related to coronal magnetic structure.

(ii) If there is even moderate flare activity, there is a reasonable probability of measuring the injection function during the extended phase of strong anisotropy, thus deriving information on the acceleration process unobtainable near 1 AU.

(iii) The inter-relationships observed among energetic particle composition, plasma dynamics and coronal structure will stimulate more comprehensive and consistent models of plasma heating and particle acceleration mechanisms.

Finally, the ideal low-energy solar particle instrumentation must have high sensitivity, a very large dynamic range, and the ability to measure field-aligned flux anisotropies. All of these capabilities should hold for electrons as well as ions from hydrogen up to iron. Then the low-energy particle detectors can fulfill their potential in complementing the other essential experiments on the Solar Probe.

ACKNOWLEDGEMENTS

The research summarized in this paper was supported in part at JHU/APL by NASA and the Air Force Geophysics Laboratory via Contract N00024-78-C-5384 between the Johns Hopkins University and the Department of the Navy, and by the National Science Foundation under grants ATM-76-23816 and ATM-76-81099.

D21

N78-32985

LOW-ENERGY PROTONS: GRADIENTS IN INTERPLANETARY SPACE AND DISTRIBUTION IN THE SOLAR CORONA

G.A. Stevens, J.J. van Rooijen, Z. Svestka and C. de Jager
Space Research Laboratory, Utrecht, The Netherlands

ABSTRACT

First and second order anisotropy measurements are proposed as a tool for studying the coronal source function and interplanetary propagation of low energy protons. Optimum orbit and attitude requirements are suggested for a three telescope system. Some limitations with regard to the lower energy limit for a feasible set-up are discussed.

INTRODUCTION

Proton fluxes in the energy range above 10 MeV are usually interpreted in terms of a sudden release from the Sun followed by a diffusion-like propagation in interplanetary space. This picture, however, does not fit the observations at sub-MeV energies for which one observes long lasting, directed streams, corotating with the Sun without apparent evolution other than fluctuations. From this, one infers that hardly any diffusion takes place perpendicular to the magnetic field lines and that the source function is extended in space and time. Recently appreciable effort has been spent in mapping the fluxes back to the corona and relating the structure thus obtained to what is known about the coronal magnetic field (see the contribution to this workshop by E.C. Roelof, S.M. Krimigis and R.E. Gold).

SOLAR PROBE ORBIT

The Solar Probe Mission offers a unique opportunity to separate the source contributions from the influence of the interplanetary medium. Referring to Figure 1 where the distance between the spacecraft and the Sun is plotted against time, one observes that the entire region within 190 $R_\odot$ (0.9 AU) is traversed in 50 days. In this period, the Sun, with the interplanetary magnetic field structure attached to it, will rotate twice around its axis. So, for an orbit in the ecliptic, the satellite will see either one or three solar rotations depending on whether the orbital inclination is 0 or 180 degrees. Obviously the latter is optimal for our purposes; it enables one to see the same field line three or four times at different distances from the Sun within 50 days. The numbered points on the curve in Figure 1 identify the crossings of selected field lines connected to the solar surface and stretched out in an Archimedian spiral. Points 2 and 3 are chosen such that the corresponding field lines will be crossed at a distance of 20 $R_\odot$, the estimated average location of the Alfvenic critical point. Below this point, the field lines may form closed loops, and at larger distances they must be stretched out into interplanetary space. All the nominal field lines encountered for the first

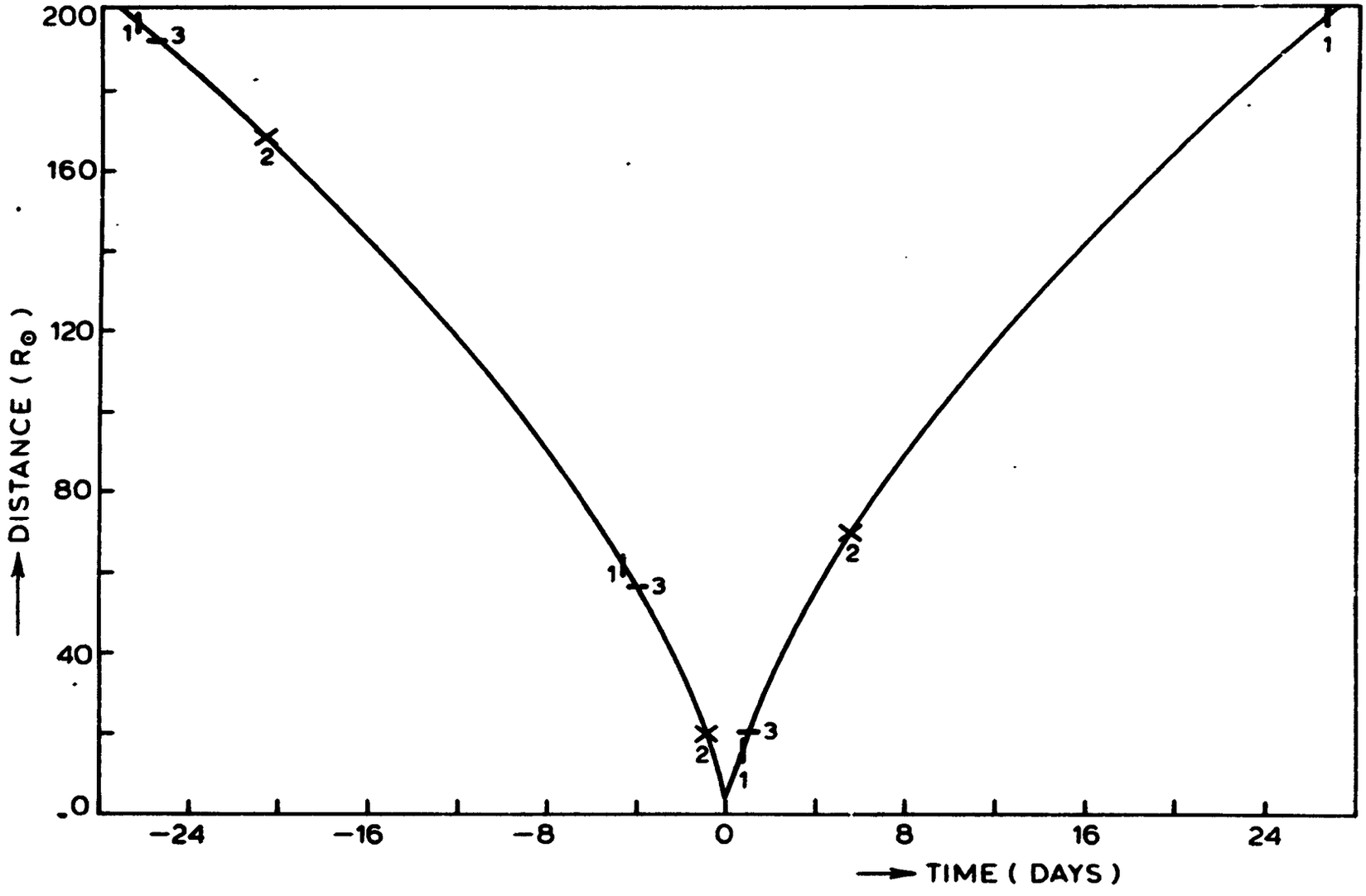


Figure 1. Distance in solar radii versus time for an elliptic orbit with perihelion of 4 $R_\odot$ and aphelion of 5.2 AU, within 28 days around perihelion. The numbers correspond to selected magnetic field lines of archimedean spiral form ($\omega_\odot$ = 14.18°/day, V = 400 km/s), for the case of anti-parallel orbital and solar angular momenta.

time between points 2 ($r = 168\ R_\odot$, $t = -20.8$ days) and 3 ($r = 56\ R_\odot$, $t = -4.1$ days) will be seen both outside and inside the Alfvenic critical point. Because the energetic particle fluxes are controlled by the magnetic field structure, a detailed study of the influence of the transition region between the Sun and interplanetary space will be valuable.

ANISOTROPY STUDIES

A valuable tool for studying the source function and interplanetary propagation is the accurate measurement of the particle flux anisotropies in at least the first and second orders. Mathematically, this is the determination of a second order expansion of the particle velocity distribution function in Legendre polynomials in μ, the cosine of the pitch angle. An in situ measurement of the magnetic field by itself is not sufficient to determine its global character. The characteristics of the particle distribution function will provide essential additional information. The presence of a first order term with the simultaneous absence of a second order term is a good indicator of an open structure, whereas the reverse situation would point to a closed structure.

In the interplanetary medium, the anisotropy is mainly the result of pitch angle scattering from the stochastic field fluctuations superimposed on the mean field as well as the adiabatic focussing towards this mean field due to a decrease of the field strength radially outward. Consider the following simplified equation for transport along the z-axis:

$$\frac{\partial f(z,\vec{v},t)}{\partial t} + v\mu \frac{\partial f}{\partial z} = D \frac{\partial}{\partial \mu} (1-\mu^2) \frac{\partial f}{\partial \mu} - \alpha(1-\mu^2) \frac{\partial f}{\partial \mu} \; ; \; z > 0 \qquad (1)$$

where $\vec{v}$ is the particle velocity, D is the scattering diffusion coefficient ($\frac{1}{2D}$ is the isotropisation time), and $\alpha = -\frac{v}{2} \frac{\partial \ln B(z)}{\partial z}$ (E.C. Roelof, 1969) governs the adiabatic focussing ($\alpha > 0$). If the distribution function is reasonably smooth in the parameter μ, it can be described by a truncated expansion of the form:

$$f(z,\vec{v},t) = f_o(z,v,t) + P_1(\mu)\, f_1(z,v,t) + P_2(\mu)\, f_2(z,v,t) \ . \qquad (2)$$

Substitution of (2) into (1) shows that third order terms are generated, the effect of which can be suppressed by evaluating (1) at the specific μ-values ($\mu_i = 0, \pm\sqrt{\frac{3}{5}}$) which cancel these $P_3(\mu)$ terms. For stationary conditions (time independent source at $z = 0$; $t >> \frac{z}{v}, \frac{1}{D}, \frac{1}{\alpha}$), one readily finds:

$$\left.\begin{aligned} f_1(z) &= c \exp(-2\alpha \tfrac{z}{v}) \\ f_2(z) &= \tfrac{\alpha}{3D} f_1(z) \end{aligned}\right\} \qquad (3)$$

Close to the source, the observations directly reflect the properties of the source ($D = \alpha = 0$).

EXPERIMENTAL ASPECTS

Some conditions for obtaining valuable information on the anisotropies must be mentioned. In order to cope with all possible directions of the magnetic field, at least three identical telescopes in a spinning spacecraft must be used with the rotation circle divided into 7 sectors. When the magnetic field is not directed along the spin-axis, in principle the third order can be included in the measurements. A spin-stabilized spacecraft would require at least 6 telescopes to obtain a reasonable portion of the required information.

As the entire orbit is in a region where the magnetic field is frozen into the plasma, the measured distribution function must be corrected for the spacecraft velocity relative to the plasma (Compton-Getting effect). Different telescope directions correspond to different energies in the plasma frame, and there is an interference between this artificial modulation and that present in the plasma frame. When the plasma flow velocity and the magnetic field are aligned, the result is an enhancement of the observed first order term and a suppression of the second order term. A true mapping of the actual distribution function is impossible because the energy and angular resolution have a

finite width. Computer simulations set up for a similar problem with an experiment aboard the ISEE-C (DFH; a cooperation between ESTEC Noordwijk, IC London, SRL Utrecht) reveal that for steepness parameters $\lesssim 5$ and plasma flow velocities $\lesssim 1000$ km/s, a conceptual experiment with 5 energy channels up to 3 MeV should not extend its lower limit much below 200 keV.

SUMMARY

To accomplish our objectives, the orbital and attitude requirements would be:

1. An ecliptic orbit with its angular momentum ánti-parallel to the solar spin.
2. A spinning spacecraft, with the spin-axis preferably directed towards the Sun, spinning at a rate of $\gtrsim 12$ rpm (to obtain a time resolution of 10 s).

The conceptual instrument package would then be characterized by:

1. Three identical telescopes placed at the zeroes of P_3 (cos θ), viz. $\theta = 39^\circ, 90^\circ, 141^\circ$ with respect to the spin-axis.
2. Energy range: (0.1) 0.2 MeV - 3 MeV split-up into (6) 5 channels. From the technical point of view, the lower limit in energy and the type of detectors to be used are dependent on the maximum temperature reached in the spacecraft.
3. Rotation circle divided into 7 sectors for all energy channels thus leading to a bit rate of about 100 b/s.

REFERENCES

E.C. Roelof, Lectures in High Energy Astrophysics, 1969 (Washington, D.C. NASA SP-199).

SECTION V.

SOLAR NEUTRONS

SOLAR NEUTRON SPECTROSCOPY NEAR THE SUN

J. A. Simpson
Enrico Fermi Institute and Department of Physics
University of Chicago

Introduction

No neutrons of solar origin have been detected unambiguously even though we know that the Sun must be a source of neutrons covering a wide energy range from both quasi-steady sources and discrete, violent events involving high energy nuclear interactions. Historically, the first suggestion that solar neutrons would be interesting was by Biermann et al. (1951). The fact that only upper limits on solar neutron fluxes have so far been reported (Lockwood, 1973; Lockwood et al., 1973; Kirsch, 1973; Balogh, 1976; and references therein) is understandable since all experimental attempts to identify fluxes of solar neutrons have been made in Earth satellites or balloons where two physical constraints make detection of solar neutrons an extremely difficult problem. First the approximately 12-minute half-life of the neutron reduces drastically the flux of low-energy neutrons which could survive to 1 AU. Second, local neutron background originating in the spacecraft or upper atmosphere--mainly from cosmic ray interactions--has placed severe limits on the sensitivity for detection at any neutron energy.

In Figures 1 and 2, we show the resultant dependence of neutron fluxes from both the radioactive decay of the neutron and the $1/R^2$ dependence as a function of radial distance R, normalized to unity at Earth and at the surface of the Sun, respectively. We have proposed, over the past approximately 15 years (Simpson, 1963, 1969, 1970; Anglin et al., 1973) that this unique radial dependence for solar neutrons could be utilized to identify neutrons of solar origin and separate solar neutrons from neutron background; it is this dependence which makes the Solar Probe Mission an attractive candidate for carrying out solar neutron experiments. Recently Balogh (1976) and Ahluwalia (1975) also have discussed this approach to the solar neutron detection problem. The purpose of this brief note is to summarize the importance of solar neutron observations for advancing our understanding of phenomena crucial for solar physics,

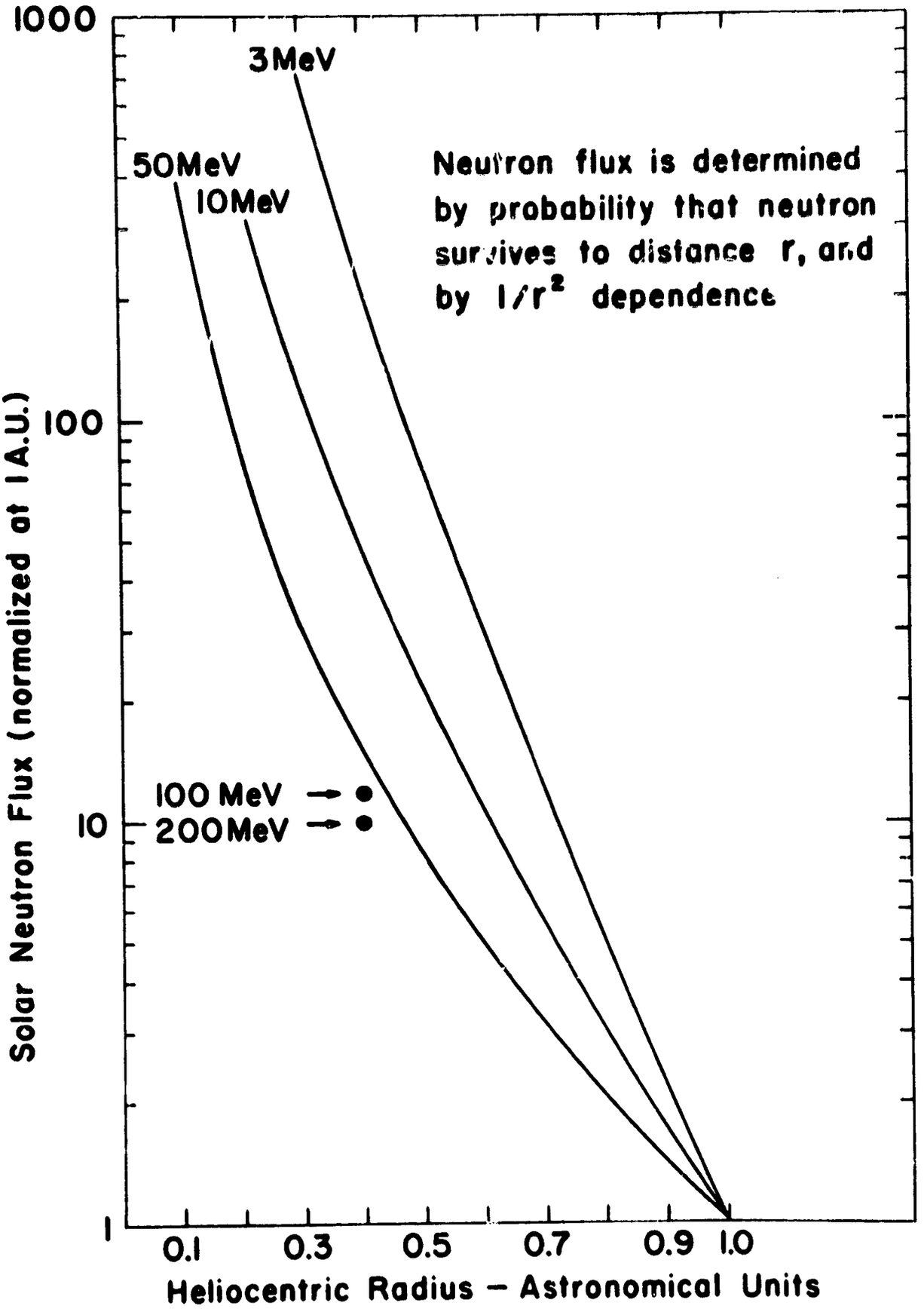


Figure 1 Variation of the solar neutron flux with distance from the Sun for neutron energies of 3, 10, and 50 MeV, normalized to 1.0 at the orbit of Earth.

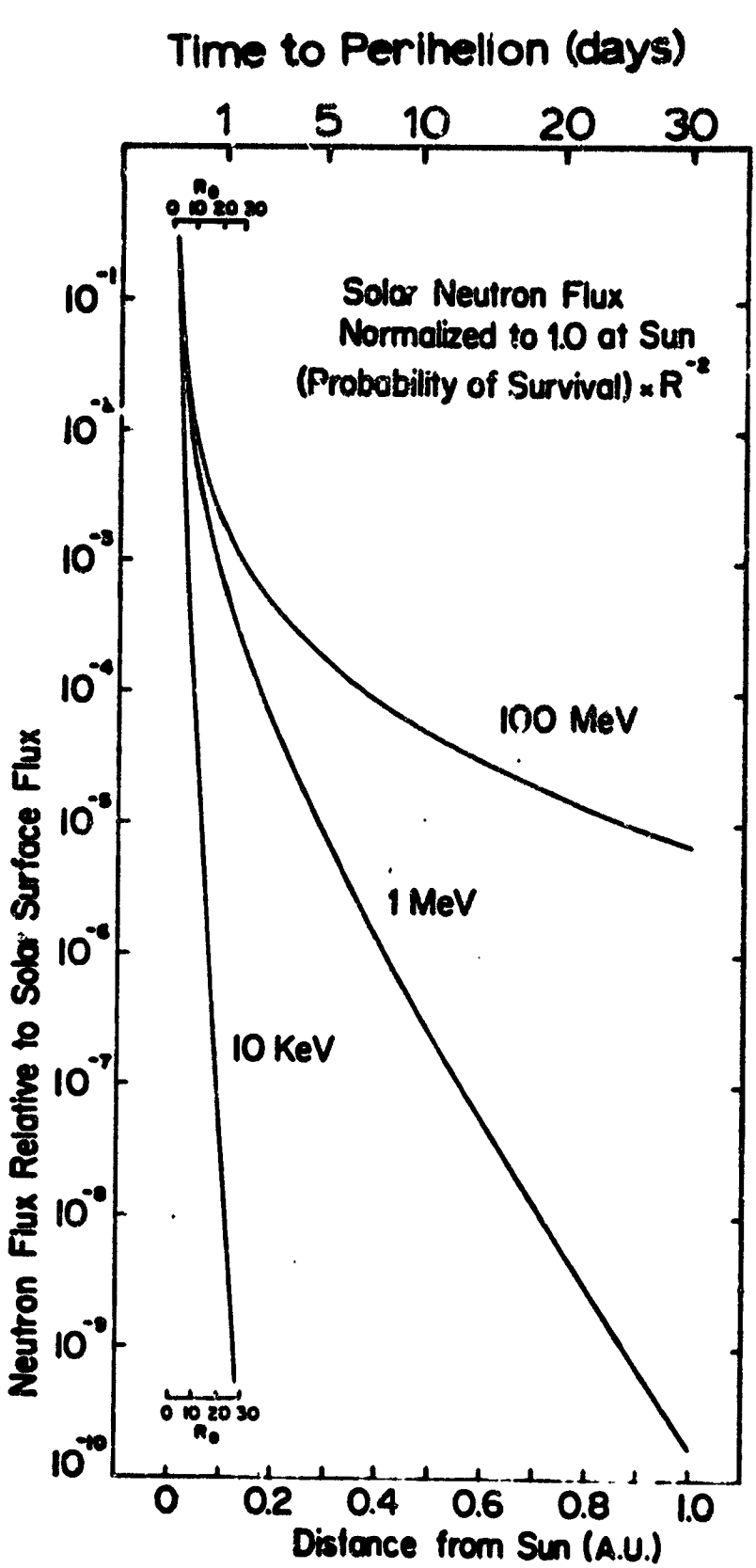


Figure 2 Variation of the solar neutron flux with distance from the Sun for neutron energies of 10 keV and 1 and 100 MeV, normalized to 1.0 at the surface of the Sun. The upper scale indicates the time to perihelion for a typical solar probe Jupiter Gravity Assist trajectory.

and to suggest the constraints which must be placed on the Solar Probe Mission for carrying out studies of this kind.

There are two essential assumptions to keep in mind during the discussion which follows in this note. First, in all discussions of the solar neutron experiment we tacitly assume that there will be simultaneous observations of gamma-rays - indeed, gamma-ray and neutron observations can be undertaken as a single experiment. Second, for an initial "discovery and explore" type of experiment, the emphasis should be upon obtaining energy spectra of the solar neutrons over as wide an energy range as possible rather than concentrating on high resolution "spectroscopy" at the expense of energy range and sensitivity.

The Probable Sources of Solar Neutrons

We may expect to find distinct populations of neutrons arriving from several processes:

1) The interaction of solar flare accelerated nuclei with photospheric, chromospheric and coronal material will yield a wide variety of nuclear interactions which produce, in addition to secondary charged nuclei, gamma-rays and neutrons. In the Appendix we estimate the neutron flux expected from the flare of August 4, 1972. This estimate is based upon a knowledge of the gamma-ray observations of Chupp et al. (1973) and the theoretical model of Ramaty et al. (1975). This leads to an integrated flux plotted as a function of radial distance as shown in Figure 3 (solar flare neutrons will be in the range 1-100 MeV).

2) Many local regions near the photosphere have been shown by x-ray and other observations to be hot spots, probably the result of bombardment by electrons and nucleons trapped in small-scale magnetic fields. Since such regions may be wide-spread on the Sun, their integral effect may constitute a quasi-steady source of neutrons as observed from a distant spacecraft. (These neutrons are likely to be in the energy range below 10 MeV).

3) Since the neutrons have a ~12-minute half-life, the neutrons produced deep in the Sun will have decayed before they could diffuse appreciable distances towards the photospheric surface. However, any neutron production near the photospheric boundary may result in the appearance of thermal and superthermal neutrons

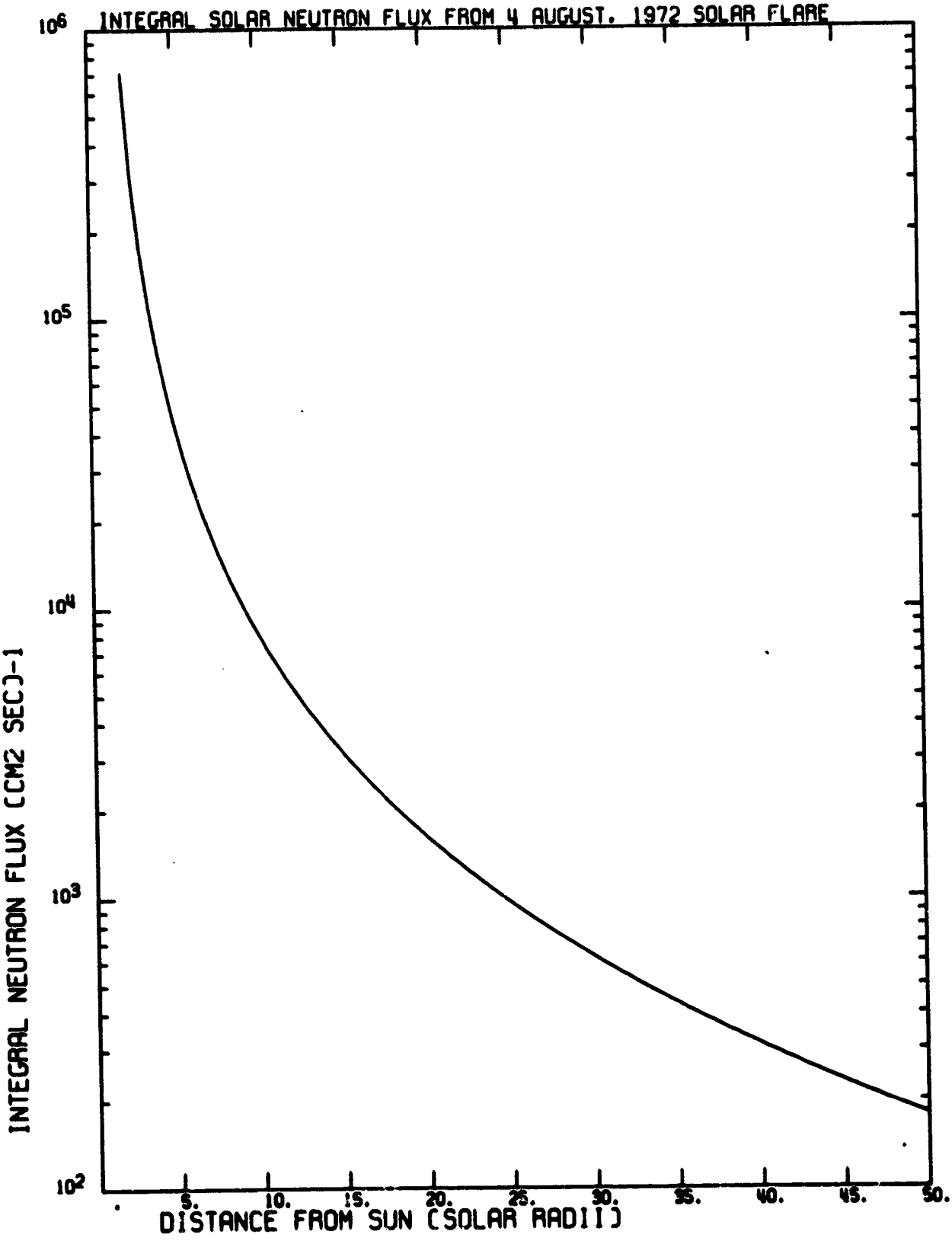


Figure 3 Calculated integral solar neutron flux expected from the August 4, 1972 solar flare from 2 to 50 solar radii.

diffusing outward beyond the photospheric boundary (at least to the "augmentation distance"). (These thermalized neutrons will be below 10^{-2} MeV).

4) Since there now exists considerable evidence for proton and helium nuclei being accelerated in association with either interplanetary shocks or magnetic field merging in magnetospheres, it is likely that similar mechanisms are operative over extensive regions of the chromosphere or corona, where accelerated helium nuclei could yield α-n reactions.

It is highly likely that other classes of neutron sources--probably more important than those discussed above--will be discovered. However, until experiments are carried out, the relative importance of the different sources of solar neutrons will largely be speculation.

The Significance of Solar Neutron Measurements

Solar neutrons provide a different kind of information on a) nuclear interactions, b) acceleration, and c) storage of high energy charged particles in the solar coronal magnetic fields. Solar neutrons are emitted promptly, are not deflected in magnetic fields, and have a distribution in the observer's frame of reference which is highly dependent on the kinematics and geometry of the nuclear interactions leading to their production. Thus, combined with gamma-ray observations, they are a new analytical tool for investigating a wide range of unsolved acceleration and storage problems, as well as the problems of He^3-rich flares and preferential enhancement of nuclear species.

Ramaty et al. (1975) in their studies of nuclear interaction products from charged particle acceleration at flare sites, have prepared tables of typical nuclear interactions which yield neutrons and gamma rays. These tables are reproduced as Table 1. In Table 2 we have summarized a few of the more prominant reactions and their products. To illustrate the importance of neutron measurements, let us consider some examples:

1) The following idealized case bears on the question of determining the abundance of He^3 at the flare (see also Imbragimov and Kocharov, 1975; Ramaty et al., 1975). Neutrons are captured at the Sun principally by either H^1 or He^3 in the reaction $n + H^1 \rightarrow H^2 + \gamma$,

Neutron producing reactions

Reaction	Threshold (MeV/nucleon)
1. $p + {}^{1}H \rightarrow n + p + \pi^{+}$	292.3
2. $p + {}^{4}He \rightarrow {}^{3}He + p + n + (\pi)$	25.7
$\rightarrow {}^{2}H + 2p + n + (\pi)$	32.6
$\rightarrow 3p + 2n + (\pi)$	35.4
3. $p + {}^{12}C \rightarrow n + \cdots$	19.6
$p + {}^{13}C \rightarrow n + \cdots$	3.2
4. $p + {}^{14}N \rightarrow n + \cdots$	6.3
5. $p + {}^{16}O \rightarrow n + \cdots$	16.6
$p + {}^{18}O \rightarrow n + \cdots$	2.5
6. $p + {}^{20}Ne \rightarrow n + \cdots$	15.9
7. $p + {}^{56}Fe \rightarrow n + \cdots$	5.5
8. $\alpha + {}^{4}He \rightarrow {}^{7}Be + n$	9.5
9. $\alpha + {}^{12}C \rightarrow n + \cdots$	2.8
$\alpha + {}^{13}C \rightarrow n + \cdots$	–
10. $\alpha + {}^{14}N \rightarrow n + \cdots$	1.5
11. $\alpha + {}^{16}O \rightarrow n + \cdots$	3.8
$\alpha + {}^{18}O \rightarrow n + \cdots$	0.21
12. $\alpha + {}^{20}Ne \rightarrow n + \cdots$	2.16
$\alpha + {}^{22}Ne \rightarrow n + \cdots$	0.15
13. $\alpha + {}^{56}Fe \rightarrow n + \cdots$	1.37
14. $\alpha + {}^{25}Mg \rightarrow n + \cdots$	–
$\alpha + {}^{26}Mg \rightarrow n + \cdots$	–
15. $\alpha + {}^{28}Si \rightarrow n + \cdots$	0.43

– exoergic

Prompt gamma ray lines

Photon energy (MeV)	Origin	Production modes	Approximate relative intensity
0.431	$^{7}Be^{*0.431} \rightarrow$ g.s.	$^{4}He(\alpha, n)^{7}Be^{*0.431}$	1
0.478	$^{7}Li^{*0.478} \rightarrow$ g.s.	$^{4}He(\alpha, p)^{7}Li^{*0.478}$	
		$^{4}He(\alpha, n)^{7}Be \xrightarrow{12\%} {}^{7}Li^{*0.478}$	1
0.72	$^{10}B^{*0.72} \rightarrow$ g.s.	$^{12}C(p, 2p\alpha)^{10}B^{*0.72}$	
		$^{16}O(p,)^{10}B^{*0.72}$	0.1
0.845	$^{56}Fe^{*0.845} \rightarrow$ g.s.	$^{56}Fe(p, p')^{56}Fe^{*0.845}$	0.2
1.24	$^{56}Fe^{*2.08} \rightarrow {}^{56}Fe^{*0.845}$	$^{56}Fe(p, p')^{56}Fe^{*2.08}$	0.2
1.38	$^{24}Mg^{*1.37} \rightarrow$ g.s.	$^{24}Mg(p, p')^{24}Mg^{*1.37}$	0.1
1.63	$^{20}Ne^{*1.63} \rightarrow$ g.s.	$^{20}Ne(p, p')^{20}Ne^{*1.63}$	
	$^{14}N^{*2.94} \rightarrow N^{*2.31}$	$^{14}N(p, p')^{14}N^{*2.94}$	
		$^{16}O(p, 2p\alpha)^{14}N^{*2.94}$	0.2
	$^{25}Na^{*2.07} \rightarrow {}^{25}Na^{*0.44}$	$^{24}Mg(p, 2p)^{25}Na^{*2.07}$	
1.78	$^{28}Si^{*1.78} \rightarrow$ g.s.	$^{28}Si(p, p')^{28}Si^{*1.78}$	0.2
1.99	$^{11}C^{*1.99} \rightarrow$ g.s.	$^{12}C(p, p\alpha)^{11}C^{*1.99}$	0.07
2.31	$^{14}N^{*2.31} \rightarrow$ g.s.	$^{14}N(p, p')^{14}N^{*2.94} \rightarrow {}^{14}N^{*2.31}$	
		$^{14}N(p, n)^{14}O \rightarrow {}^{14}N^{*2.31}$	0.3
		$^{16}O(p, 2p\alpha)^{14}N^{*2.31}$	
2.75	$^{16}O^{*6.88} \rightarrow {}^{16}O^{*6.14}$	$^{16}O(p, p')^{16}O^{*6.88}$	0.07
~3.62	$^{12}C^{*3.68} \rightarrow$ g.s.	$^{16}O(p, 3p\alpha)^{12}C^{*3.68}$	0.07
	$^{6}Li^{*3.56} \rightarrow$ g.s.	$^{12}C(p,)^{6}Li^{*3.56}$	
3.84	$^{12}C^{*3.68} \rightarrow$ g.s.	$^{16}O(p, 3p\alpha)^{12}C^{*3.84}$	0.07
4.43	$^{12}C^{*4.43} \rightarrow$ g.s.	$^{12}C(p, p')^{12}C^{*4.43}$	
		$^{12}C(\alpha, \alpha')^{12}C^{*4.43}$	1
		$^{16}O(p, p\alpha)^{12}C^{*4.43}$	
~5.3	$^{15}O^{*5.24} \rightarrow$ g.s.	$^{16}O(p, p n)^{15}O^{*5.24}$	
	$^{15}N^{*5.28} \rightarrow$ g.s.	$^{16}O(p, 2p)^{15}N^{*5.28}$	0.3
	$^{15}N^{*5.31} \rightarrow$ g.s.	$^{16}O(p, 2p)^{15}N^{*5.31}$	
6.14	$^{16}O^{*6.14} \rightarrow$ g.s.	$^{16}O(p, p')^{16}O^{*6.14}$	0.5
		$^{16}O(\alpha, \alpha')^{16}O^{*6.14}$	
6.33	$^{15}N^{*6.33} \rightarrow$ g.s.	$^{16}O(p, 2p)^{15}N^{*6.33}$	0.5
~6.7	$^{11}B^{*6.79} \rightarrow$ g.s.	$^{12}C(p, 2p)^{11}B^{*6.79}$	
		$^{12}C(p, 2p)^{11}B^{*6.81}$	0.07
		$^{12}C(p, p n)^{11}C^{*6.89}$	
7.12	$^{16}O^{*7.12} \rightarrow$ g.s.	$^{16}O(p, p')^{16}O^{*7.12}$	0.2

TABLE 1 Flare-site reactions which produce neutrons and prompt gamma rays.

A FEW NEUTRON PRODUCTION REACTIONS

REACTION	THRESHOLD ENERGY (MeV/NUCLEON)
H^1 (p, n π^+) H^1	292.3
He^4 (p, pn) He^3	25.9
He^4 (p, 2pn) H^2	32.8
He^4 (p, 2p 2n) H^1	35.6
He^4 (α, n) Be^7	9.5
C^{12} (p, n . . .	19.8
N^{14} (p, n . . .	6.3
O^{16} (p, pn . . .	16.5

MAJOR PROMPT GAMMA RAY LINES

ORIGIN	PHOTON ENERGY (MeV)	PRODUCTION ENERGY (MeV)
C^{12*}	4.43	~5 - 100
O^{16*}	6.14	~7 - 100

DELAYED GAMMA RAY LINES

ORIGIN	PHOTON ENERGY (MeV)
$e^+ + e^- \rightarrow 2\gamma$	0.511
$n + H^1 \rightarrow H^2 + \gamma$	2.223

TABLE 2 Some flare-site reactions and their products.

yielding a 2.2 MeV gamma-ray, or $n + He^3 \rightarrow H^3 + p$, which yields no photon.

The ratio of the cross-sections of these reactions is

$$\sigma_{He^3} / \sigma_{H^1} \simeq 2 \times 10^4$$

If the He^3/H^1 abundance ratio in the photosphere, where the capture process takes place, is very small (10^{-6} or less) the yield of 2.2 MeV gamma-rays relative to neutrons will be larger than if the He^3/H^1 ratio is comparable to that observed in the solar wind (5×10^{-5}), in which case He^3 will capture nearly half of the neutrons. A measurement of the relative yield of neutrons and 2.2 MeV gamma-rays may thus provide information on the He^3/H ratio in the ambient material at the flare site.

2) Another example, kinematical, can be drawn from the fact that at the flare site of the nuclear reactions the excited gamma-rays will be emitted isotropically whereas the neutrons in the observer's frame of reference will have two separate components:

   a) A low-energy isotropic component arising from nuclear disintegration "evaporation" neutrons (at energies below 30 MeV), and

   b) High energy neutrons from collisions of high energy accelerated nuclei with solar matter. These neutrons will travel mainly in the direction of the incident flare-accelerated nucleons. (See also Kanbach et al. 1975, for additional "kinematical" examples).

It is clear, therefore, that the ratios of gamma-ray and neutron fluxes at high energies will be dependent on the direction of the accelerated nuclei. If the high energy neutrons are moving downward towards the photosphere they will not be observed in space.

3) The energy dependence for neutron production may also become an analytical tool. For example, whereas a gamma-ray production cross-section may drop off with increasing incident nucleon energy, the neutron production cross-section for important reactions remains high or may increase with energy. For some cases, such as the p-p interaction in Table 2, above the threshold the neutron yield is far above measurable gamma-ray intensities

at these energies.

Working out these and other illustrations will show that solar neutrons provide unique information crucial to our understanding of a wide variety of solar particle acceleration phenomena.

Since the neutron decays into a low energy proton plus electron and neutrino in the coronal and interplanetary magnetic field, the existence of a thermal neutron flux whose source is close to the photosphere may possibly be verified by determining the intensity and pitch angle of low energy decay protons as a function of distance along the coronal and interplanetary magnetic fields. (Simpson, 1963; Roelof, 1966).

Neutron Yields From Solar Flares

The neutron yields from different mechanisms in the Sun are difficult to estimate. However, in the Appendix we have attempted to show how estimates of the absolute neutron flux can be obtained from flares in which gamma-ray observations have been made (cf. Reppin, et al., 1973; Wang and Ramaty, 1974). Our example is the flare of August 4, 1972. The calculated integral flux of solar neutrons is shown in Figure 3.

The Measurement of Solar Neutrons

As we mentioned in the Introduction, the main emphasis for initial observations would be placed on determining the fluxes within broad energy ranges. A wide variety of classical techniques exists for detecting neutrons, but no single technique can cover effectively all the energy ranges proposed for this investigation. Some examples of detection methods are as follows:

1) Thermal and superthermal neutrons with velocity v at low energies may be detected using the $1/v$ cross-section dependence of neutron-capture isotopes such as B^{10}, Li^{6}, etc. (These $1/v$ dependent detectors are especially efficient after slowing down the neutrons in moderating material surrounding the detector but can be used only when no SIG is present. See following Section.)

2) In combination with the above detector scheme, various shields made of materials having epi-thermal resonance cross-sections, may be used (e.g., cadmium) to select higher energies.

3) For separating slow neutrons from fast neutrons, fission ion

chambers can be used effectively; e.g., U^{235} for slow neutrons; U^{238} or Th^{232} for > 1 MeV neutrons.

4) For high energy neutrons, n-p recoil reactions may be utilized.

5) Double-scattering and time-of-flight for neutrons within the instrument (and between source and detector, see below) is a well-developed technique for high energy neutrons.

6) For the study of impulsive events such as solar flares, the time-of-flight between the time of initial neutron production in the flare and the time of arrival at the spacecraft can be determined with reasonable accuracy for those events which simultaneously produce γ-rays. The average neutron velocity can then be determined from the difference in onset times of γ-ray and neutron fluxes at the spacecraft and the relative positions of the flare site and the spacecraft.

Constraints for Solar Neutron Measurements on the Solar Polar Mission

The SIG power sources proposed for a Solar Probe would produce a local neutron spectrum extending in energy up to ~12 MeV as shown in Figure 4 (from data taken with an RTG of the type used for Pioneer 10). Since the nuclear "fuel" emits α-particles and some light element impurities are present in the fuel, α-n reactions produce a significant local neutron background below 12 MeV. Thus, if SIGs are used, sensitive solar neutron measurements would have to be confined to neutron energies above 12 MeV. The experiment would still be of high interest and would achieve many of the scientific goals. However, the experiment would be greatly enhanced at low energies by using non-nuclear power sources, such as solar cells as proposed in conjunction with the ion propulsion option. Sensitive gamma-ray observations below 3 MeV would also encounter serious difficulties with SIGs.

For a mission employing a Jupiter-gravity assist, it would be interesting to investigate the possibility of jettisoning the SIGs on the inward leg of the trajectory at ~0.5 AU. The required power could then be supplied first by solar cells and, close to the Sun, by thermocouples embedded in the heat shield.

Since the analysis of solar flare events is enhanced by multiple spacecraft observations approximately in the plane of the probe trajectory, a trajectory whose plane is near the equatorial plane of the Sun is preferable to a plane out of the ecliptic since there are likely to be many more

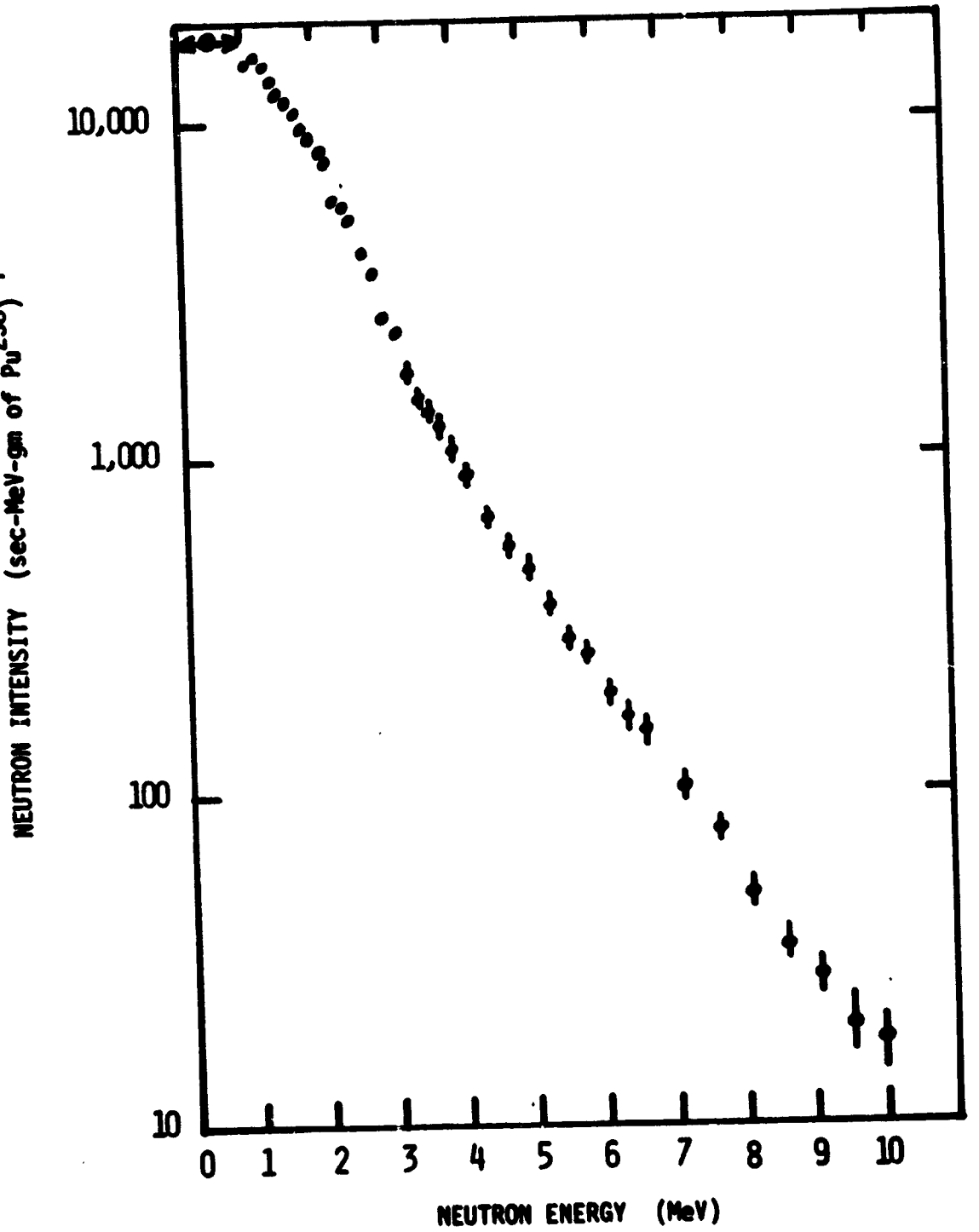


Figure 4 Neutron energy spectrum for Pioneer 10 type RTG capsule. Total neutron emission for this capsule: 4.3×10^7 neutrons/sec. Specific yield: 3.7×10^4 neutrons/sec-gm of Pu^{238}.

spacecraft in the solar equatorial plane than at high solar latitudes in the era of 1990.

Measurements to complement the solar neutron and gamma-ray observations should include the isotopic composition of nuclei (especially the H and He isotopes) over a wide energy range, spectra of electrons, cosmic ray background intensity, magnetic field, and plasma measurements.

Alternatives to the Solar Probe Mission

The only alternative to the Solar Probe Mission appears to be a Mercury fly-by (Simpson, 1970) or a Mercury orbiter (Simpson, 1976; Balogh, 1976). Other than not requiring SIGs, the advantages of a low latitude Mercury orbiter for neutron and gamma-ray observations are:

1) The planet may be used as an opaque "shutter" to remove periodically the solar neutron and gamma-ray fluxes for the purpose of determining neutron and gamma-ray backgrounds derived from the spacecraft (and planet).
2) The exposure time could extend to more than year. This is to be compared with the time of ~20 days inside 0.5 AU for the Solar Probe.

The Mercury orbiter probably could not investigate the lowest energy neutron fluxes. The Solar Probe and the Mercury orbiter missions are obviously complementary.

Concluding Remarks

The Solar Probe Mission is a unique mission for the discovery of solar neutrons, even if SIGs must be used on the spacecraft thus limiting the study of neutrons to energies above 12 MeV.

In the past, the discovery of each new component of radiation, opening a new "window to the universe", such as cosmic rays, x-rays, infrared etc., has led to dramatic advances in our understanding of astrophysical phenomena. It is likely that solar neutron spectroscopy will be no exception.

Several attempts have already been made by the author to include solar neutron studies on earlier NASA missions extending inward from 1 AU, but so far no experiment has been realized. The Solar Probe Mission appears to be the next best opportunity for opening the field of solar neutron spectroscopy. The Mercury orbiter would also be important for a program of solar neutron investigations.

Appendix

The Solar Flare Neutron Flux for the Proposed Solar Probe Mission

We present here an estimate of the neutron flux from the solar flare of August 4, 1972. Two approaches to estimating neutron fluxes from flares are possible:

1) The upper limit for production of secondary nuclei produced in flares could be used for large flares (H^1, H^2, H^3, He^3, etc.), or

2) Evidence from nuclear gamma rays could be invoked. In this note we follow the latter approach which was worked out in principle for the large flare of August 4, 1972 by Ramaty *et al.* (1975), (referred to as RKL). They calculated both:

   a) the neutron production spectrum, and

   b) the yield of gamma rays due to neutron capture on protons (2.2 MeV gamma ray line) and excited states of nuclei in the solar atmosphere (e.g., the 4.4 MeV carbon line),

   assuming unit values for the "emission measure" for the flare, that is, the product of the density in the interaction region and the intensity of high energy primary protons. The observations by Chupp *et al.* (1973) of the gamma ray flux provide the normalization factors necessary to calculate the expected neutron flux for the August 4, 1972 solar flare.

For the calculation of the neutron flux as a function of distance from the Sun, a neutron emission flux is used which is of the form:

$$q(E_n) = \begin{cases} q_0 \text{ for } E_n < 30 \text{ MeV} \\ q_0\left(\frac{E_n}{30}\right)^{-2} \text{ for } E_n > 30 \text{ MeV} \end{cases}$$

i.e. independent of energy up to 30 MeV and then falling as E^{-2} (after Figure 7 of RKL).

For the August 4, 1972 solar flare the normalization factor q_o is:

$$q_o = q(E=0)\ n_H N_p\ (>30\ \text{MeV})\ P_E = 6 \times 10^{25}\ \text{neutrons/MeV sec}$$

where

$q(E=0)$ is the neutron production rate for unit emission measure, taken to be 2×10^{-18} MeV^{-1} sec^{-1}, (after Figure 7 of RKL)

$n_H N_p$ (>30 MeV) is the emission measure, calculated by RKL on the basis of the gamma-ray observations to be 5×10^{43} for the solar flare of August 4, 1972, and

P_E is the probability of neutron escape from the production region, taken to be 0.6 independent of neutron energy (Wang and Ramaty, 1974).

These results were taken from the RKL model calculation for the thin target model with a power law primary spectrum but the results are not changed significantly (factor of $\sim$5) for other models calculated by RKL.

The solar neutron flux as a function of neutron energy and distance from the Sun, is then calculated from

$$J(E_n,R) = q(E_n)\ P_s(E_n,R)\ R^{-2}$$

where

$P_s(E_n,R)$ is the probability that a neutron of energy E_n will survive to travel a distance R. This function, divided by R^2 is plotted in Figure 2 for neutron energies of 10 keV, 1 MeV, and 100 MeV.

In Figure A1 is plotted our calculation of the expected neutron energy spectrum for the August 4, 1972 flare at distances from the flare site of 5, 10 and 20 solar radii and 0.2, 0.5, and 1.0 AU. In Figures A2 and 3, the integral neutron flux (integrated over neutron energy to 10^3 MeV) is plotted for R = 0.01 to 1.0 AU (Figure A2) and R = 2 to 50 solar radii (Figure 3). (The major contributions to this integral are from neutrons of energy 1 MeV to 100 MeV).

The extrapolation of these results to flares other than the August 4, 1972 flare is, at best, tenuous. With neither neutron nor gamma ray observations the only other normalization factor available is the high energy proton flux.

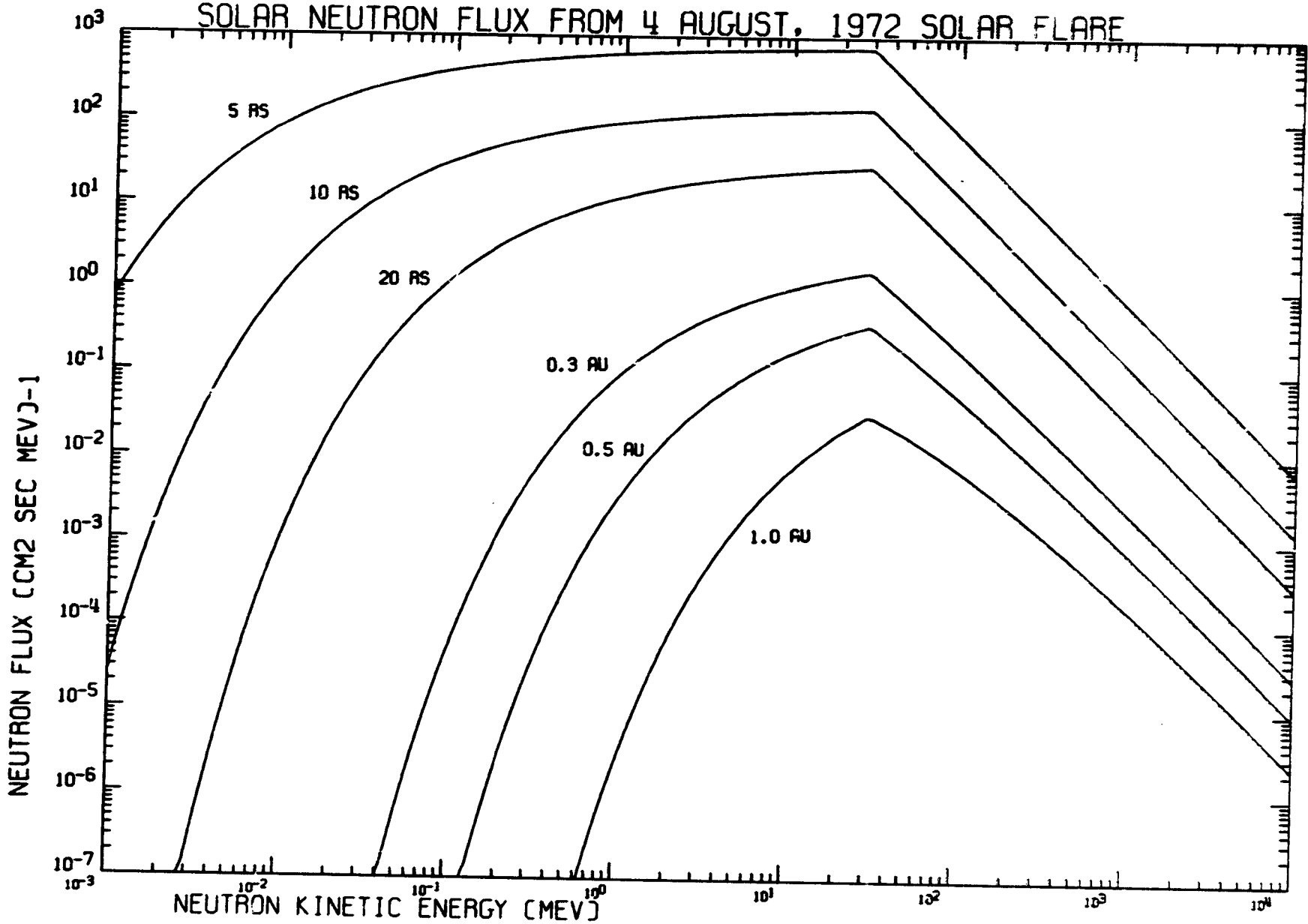


Figure A1 Calculated differential solar neutron spectrum expected from the August 4, 1972 solar flare at 5, 10, and 20 solar radii and 0.3, 0.5, and 1.0 AU.

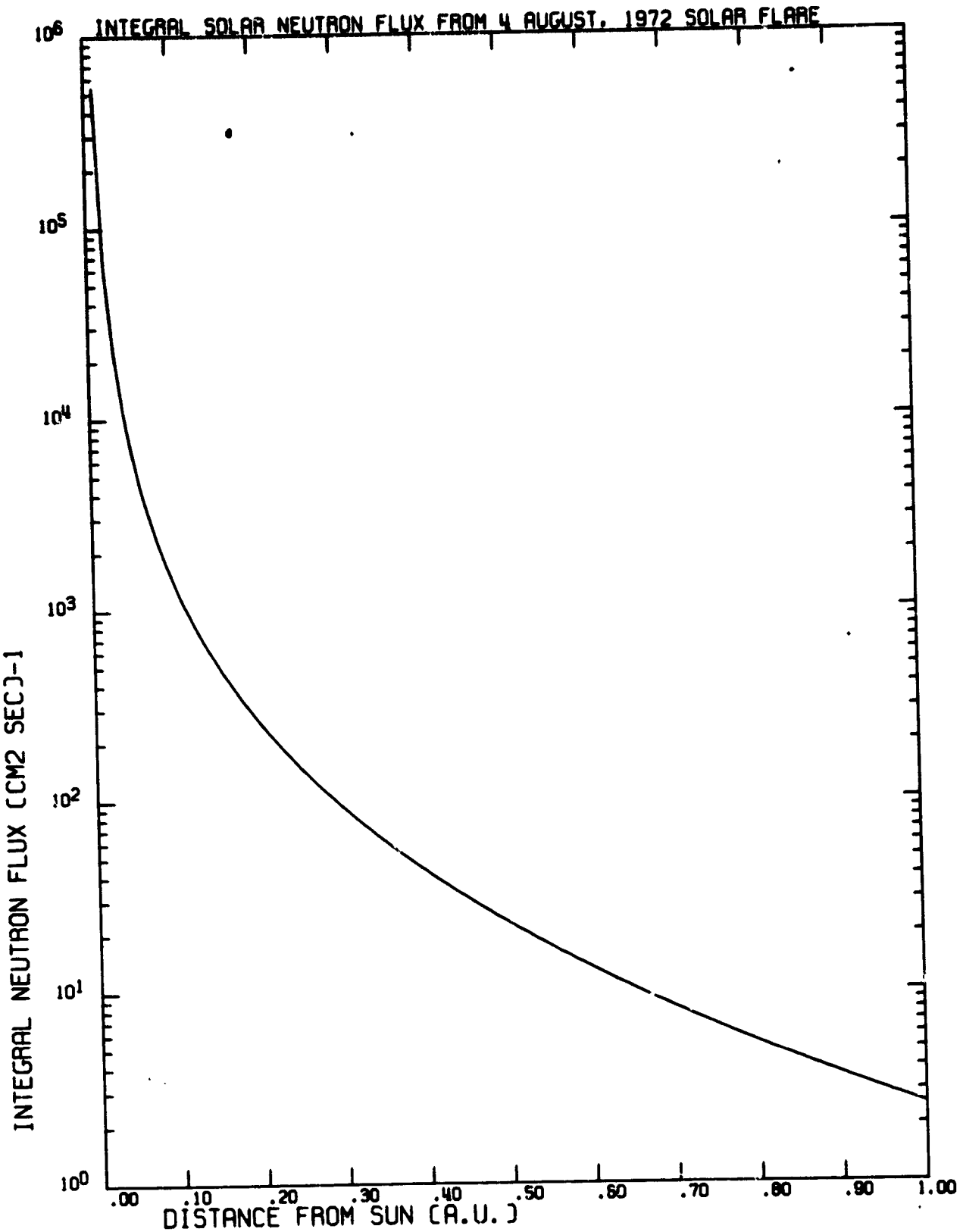


Figure A2 Calculated integral solar neutron flux expected from the August 4, 1972 solar flare from 0.01 to 1.0 AU.

Of 25 solar flares defined as "major" by King (1974) which occurred during solar cycle 20 from 1966 to 1972, the August 4, 1972 flare was the largest. Total fluences of protons with energy above 10 MeV ranged from ∿10% to 0.1% of the August flare fluence. Presumably the expected neutron flux may be scaled from flare to flare in proportion to the observed proton fluence. Thus one may naively expect an average "major" flare to contribute a neutron flux of a few percent of the values calculated here. Finally, due to the fact that the neutron pulse from a flare is so short (on the order of an hour) any observer will either see the flare and be subject to the full neutron fluence or will be on the wrong side of the Sun and see nothing. Assuming an isotropic neutron angular emission distribution, any observer will see ∿ half of all neutron events.

In summary, for the spacecraft on a trajectory which passes within 10 solar radii, the neutron fluence for a flare on the "visible" side of the Sun of the magnitude of the August 4, 1972 event, and with neutron emission occurring for $\sim 10^3$ sec, will be $10^6 - 10^7$ neutrons/cm^2 in the energy range 1-100 MeV.

The writer is indebted to David Chenette for his assistance, especially for the calculations in the Appendix relating to the August 4, 1972 event. This work was supported in part by NSF Grant ATM 77-24494 and NASA Grant NGL 14-001-006.

References

Ahluwalia, H.S., in The Proceedings of the 14th Int'l. Cosmic Ray Conf., 5, 1655, 1975.

Anglin, J.D., W.F. Dietrich, and J.A. Simpson, in High Energy Phenomena on the Sun Symposium Proceedings, R. Ramaty and R.G. Stone eds., NASA SP 342, 315, 1973.

Balogh, A., European Space Agency Document, 1976.

Biermann, Von L., O. Haxel, and A. Schluster, Z. Naturforschung, 6a, 47, 1951.

Chupp, E.L., D.J. Forrest, and A.N. Suri, in High Energy Phenomena on the Sun Symposium Proceedings, R. Ramaty and R. G. Stone eds., NASA SP 342, 285, 1973.

Ibragimov, I.A. and G.E. Kocharov, in The Proceedings of the 14th Int'l. Cosmic Ray Conference, 5, 1620, 1975.

Kanbach, G., C. Reppin, D.J. Forrest, and E.L. Chupp, in The Proceedings of the 14th International Cosmic Ray Conference, 5, 1644, 1975

King, Joseph H., J. Spacecraft and Rockets, 11, 401, 1974.

Kirsch, E., Solar Physics, 28, 233, 1973.

Lockwood, John A., Space Sci. Rev., 14, 663, 1973

Lockwood, J.A., S.O. Ifedili, and R.W. Jenkins, Solar Physics, 30, 183, 1973.

Ramaty, R., B. Kozlovsky, and R.E. Lingenfelter, Space Sci. Rev., 18, 341, 1975.

Reppin, C., E.L. Chupp, D.J. Forrest, and A.N. Suri, in The Proceedings of the 13th International Cosmic Ray Conference, 2, 1577, 1973.

Roelof, E.C., J. Geophys. Res., 71, 1305, 1966.

Simpson, J.A., Semaine d'Etude sur le Probleme du Rayonnement Cosmique dans l'Espace Interplanetaire, Vatican, Pontificia Academic Scientiarum, 323, 1963.

Simpson, J.A., NASA Document No. SL039 (unpublished), 1969.

Simpson, J.A., J.P.L. Document, JPL 615-5, 1970.

Simpson, J.A., Memo to J.P.L., Attention: V. Clarke, 1976.

Wang, H.T., and R. Ramaty, Solar Physics, 36, 129, 1974.

N78-32987

ON NEUTRONS FROM THE SUN

H.S. Ahluwalia

Department of Physics and Astronomy
The University of New Mexico
Albuquerque, New Mexico 87131

1. Introduction. Biermann et al. (1951) were the first to predict that neutrons of solar origin may sometimes be found in the neighborhood of the earth. They pointed out that solar protons which give rise to the ground level enhancement (GLE) following a large solar flare must undergo nuclear interactions in the solar atmosphere. Neutrons are one of the by-products of these reactions. Magnetic fields on the sun and in the interplanetary medium do not significantly restrain them. Solar neutrons should therefore be observed near earth at the time of GLE, provided that (a) the neutrons are emitted in the direction away from the sun, (b) there is no sink for neutrons in the solar atmosphere, and (c) the neutrons have large enough kinetic energy to survive the journey to the orbit of the earth; neutrons have a mean life of about 1000 seconds.

Refined calculations made by Hess (1962), Chupp (1963, 1971), and Lingenfelter et al. (1965a,b; 1967) indicate that a measurable flux of high energy solar neutrons should be obtainable near the orbit of the earth following a large solar flare, perhaps of the type observed on 23 February 1956.

2. Solar Neutron Searches. In 1975 I made a survey of all the experiments carried out over the previous decade to verify the theoretical predictions mentioned above (Ahluwalia, 1975a,b). The survey is given in

Table 1. One can see that with the exception of one measurement, all others only define upper limits for the flux of solar neutrons. The claim of definite discovery by Apparao et al (1966) has been seriously questioned by Hess and Kaifer (1967). Moreover the large solar neutron flux reported by Apparao et al. has not been confirmed by later experiments with much better statistics. For example, compare the result of Apparao et al. with the flux reported by Leavitt et al. (1972), my colleagues at the University of New Mexico. Leavitt et al. covered an almost identical energy range as Apparao et al., yet their reported upper limit of 4 neutrons/(m^2 sec) is the lowest in Table 1. Similar numbers have been reported recently by Moon et al. (1976). It should be pointed out that no GLE was observed either immediately before or during the times of observation listed in Table 1. The null results therefore allow only a restricted inference to be drawn. They perhaps imply that there is no steady state flux of high energy solar neutrons near the orbit of the earth, above the noise level of the detectors used. One need not conclude however that neutrons are not produced in the solar atmosphere.

As a matter of fact, Chupp et al. (1973) have reported the discovery of the 2.224 MeV γ-ray line characteristically associated with the formation of a neutron. The line was observed during the large solar flares of August 4 and 7, 1972. The clear implication is that at least low energy neutrons must have been present in the solar atmosphere during these times. Interestingly enough, GLE's were observed on these two dates. Unfortunately, no neutron detectors were aloft on these days, so it is not clear whether any high energy neutrons were also produced. However, Reppin et al. (1973) estimate that a neutron detector covering a kinetic energy (T_n) range

TABLE 1

Quiet-Time Flux of Solar Neutrons

Literature Reference	Detector Response Characteristics: Detector type	Energy Range MeV	Energy of Maximum Response MeV	Detection Efficiency at Max. %	Average A_p	Location of Experiment	Upper Limit on Solar Neutron Flux m^{-2} sec^{-1}
Haymes (1964)	Proton recoil (Balloon)	1-14	3	18	25	Brownwood, Texas	$\lesssim$200
Apparao et al (1966)	Nuclear Emulsions (Balloon)	20-160	?	?	4	Hyderabad, India	450±60
Bame & Asbridge (1966)	Proton recoil (Vela 1,2,3)	0.1-10	5	?	?	outer space	$\lesssim$100
Hess & Kaifer (1967)	BF_3 Counter (OSO-1)	0.01-10	?	?	?	outer space	$\lesssim$20
Webber & Ormes (1967)	Proton recoil (Balloon)	100	?	10	5	Tucuman, Argentina	$\lesssim$24
Wolcott (1968)	Proton recoil (Balloon)	30-130	65	$\lesssim$1	5	Alamogordo, New Mexico	$\lesssim$28
Heidbreder et al (1970)	Proton recoil (Balloon)	100-400	100	1	19	Palestine, Texas	$\lesssim$12.5
Daniel et al (1971)	Proton recoil (Balloon)	50-500	?	?	4	Hyderabad, India	$\lesssim$1500
Eyles et al (1971)	Proton recoil (Balloon)	50-350	200	1.5	10	Kampala, Uganda	$\lesssim$30
Cortellesa et al (1971)	Proton recoil (Balloon)	10-200	75	14	6	Gap & Aire-sur-L'Adour, France	$\lesssim$55
Leavitt et al (1972)	Proton recoil (OGO-6)	20-200	100	0.01	?	outer space	$\lesssim$4
Lockwood et al (1973)	Proton recoil (OGO-6)	1-20	1	?	?	outer space	$\lesssim$18

10 MeV $\leq T_n \leq$ 100 MeV would have detected a peak flux of about 20 neutrons/(m^2 sec) on August 4 and about 30 to 50 neutrons/(m^2 sec) on August 7, following the flares.

3. A Case for Low Energy Solar Neutrons. Our sun is known to emit protons in the energy range, 0.5 MeV $\leq T_p \leq$ 100 MeV, quite frequently (McCracken and Rao, 1970). These protons are prevented from reaching the ground level by absorption in the atmosphere of the earth. They can only be observed directly by detectors placed on high flying balloons, on satellites and on space probes or indirectly by observing the so-called polar cap absorption (PCA) events. At this workshop I wish to restate my case regarding the strong possibility that low energy neutrons are likely to be far more frequently produced in the solar atmosphere. A large fraction of such neutrons decay on their way to earth; e.g., a 2 MeV solar neutron has less than 1 in 1000 chances to arrive at the orbit of the earth. So even if we have a spectrum of neutrons produced on the sun, the lower energy neutrons are lost through decay in the interplanetary medium. Also the higher energy neutron flux suffers a $1/r^2$ attenuation with distance. So their presence at the orbit of the earth is completely masked by the noise level of the detectors, at least of the types listed in Table 1. So the null result of direct neutron searches can indeed

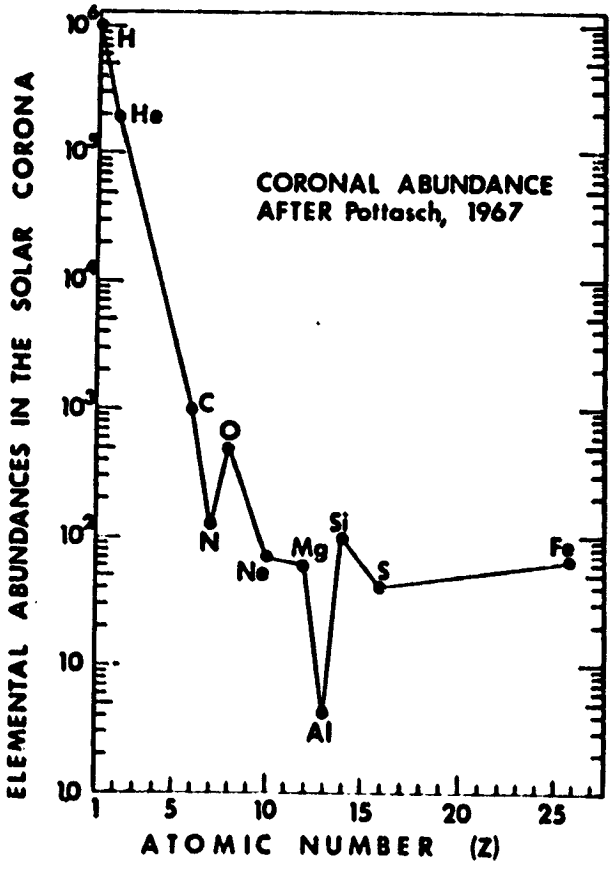


Figure 1

be reconciled with the positive inference demanded by the Chupp et al. discovery.

Figure 1 shows the elemental abundance in the solar corona relative to the hydrogen abundance of 10^6, given by Pottasch (1967). Note that the heavy elements Fe, S, Si and Mg are about as abundant as are the light elements Ne and N. Together the four heavy elements are about 1/6 times as abundant as are the light elements C, N, O, and Ne together. Lingenfelter et al. (1965a,b; 1967) have made estimates of neutron fluxes in which protons with $T_p > 100$ MeV undergo (p,nx) reactions with light nucleii, He, C, N, O, and Ne. Unfortunately the energy range and the fluxes of the energetic protons used by them are based on the numbers observed during the flare of 23 February 1956. Their calculations therefore make very optimistic predictions for the flux of solar neutrons with mean energy of about 100 MeV. In fact searches listed in Table 1 were in part inspired by these calculations. We now know that the flare of 23 February 1956 was quite a spectacular event. No other such event has been observed in the last forty years.

Figure 2 gives neutron separation energy for stable nuclides up to iron. Their natural abundance is indicated by vertical lines from the top of the diagram. The following features are easily noted.

(a) For the most abundant natural isotopes of elements C, O, Ne, Mg, Si, S, Ar and Ca, the neutron separation energies are in excess of 10 MeV.

(b) For the most abundant natural isotopes of

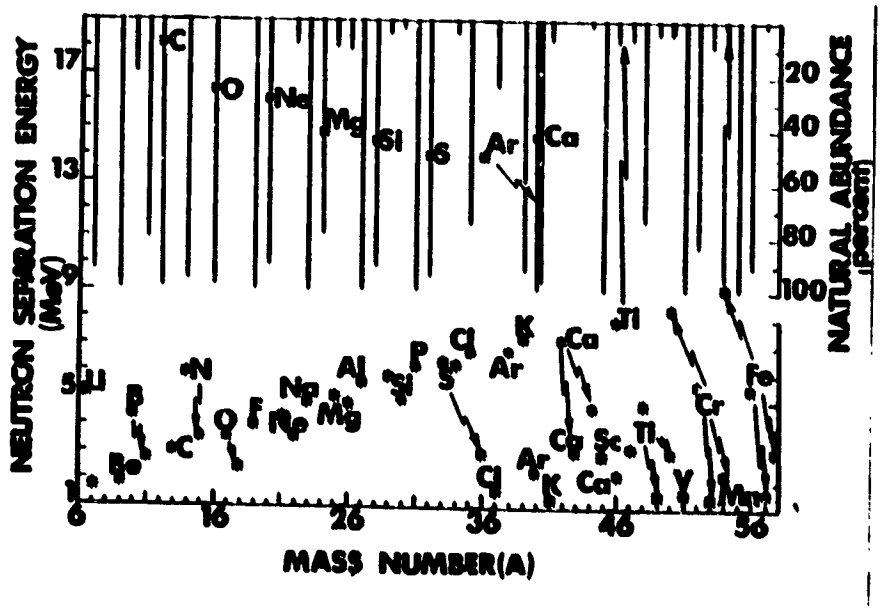


Figure 2

elements Li, Be, B, N, F, Na, Al, P, Cl, K, Sc, Ti, Cr, and Fe, the neutron separation energies are below 10 MeV. For some nuclides like Be, V, and Mn, the neutron separation energies are less than 2 MeV. Some of these nuclides are not very abundant in the solar corona, however.

It is clear therefore that protons with energies $T_p \leq 10$ MeV can undergo (p,n) reactions with a large number of nuclides in the solar atmosphere. This expectation is confirmed by a study of the cross-sections for (p,n) reactions involving some of the nuclides. Figure 3 shows the plot of some of the available cross-sections as a function of proton kinetic energy (T_p). The cross-sections may be classified into two groups. One group of cross-sections applies to protons with $T_p \sim 20$ MeV and the other group applies to protons with $T_p \leq 10$ MeV. Note that the first group contains the light elements. Lingenfelter et al.'s (1965a,b; 1967) neutron yield calculations refer to this group. The second group contains heavy elements. A comprehensive calculation involving these nuclides is yet to be made, although some of the nuclides have been included in a more recent calculation of Ramaty et al (1975).

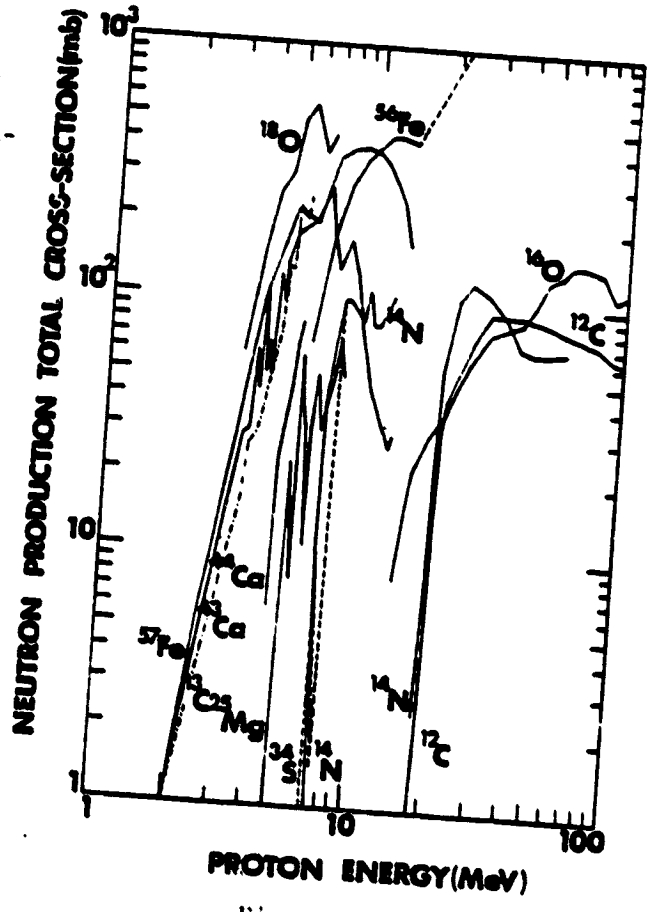


Figure 3

(i) The "kinkiness" in the second group of cross-sections arises from the existence of narrow resonances. So relatively minor constituents in the solar atmosphere such as ^{57}Fe (~1%) and ^{25}Mg (~20%) should produce a

significant number of daughter neutrons through (p,n) reactions for $T_p \leq 10$ MeV. Also relatively more abundant species like ^{56}Fe (~90%) and ^{14}N (~100%) should contribute a significant proportion of neutrons for $T_p < 10$ MeV.

(ii) It should be borne in mind that the solar proton spectrum is known to be very steep (Ahluwalia, 1969). It has the form of a power-law, with exponent ≈ -5. This indicates that for each proton with $T_p = 100$ MeV, there are 10^5 protons with $T_p = 10$ MeV. This large flux of lower energy protons can turn even a minor solar constituent into a prolific contributor of neutrons involving (p,n) reactions.

(iii) Elsewhere (Ahluwalia, 1969, 1971) we have argued that on some occasions protons with $T_p \leq 10$ MeV continue to be produced on the sun for several tens of hours after the visible traces of the optical solar flare have vanished. Such protons may even be stored at suitable locations in the solar atmosphere. It is clear therefore that during the period between the acceleration and the escape of solar protons into the interplanetary medium neutrons must be produced through (p,n) reactions, among other things. One might then expect to see the 2.224 MeV γ-ray line for an appropriate period of time. Clearly the intensity of the line at a given time would depend upon the reaction rate available in the region where the protons are accelerated and/or are stored.

4. Conclusion. From the above discussion it is clear that the solar probe mission discussed at this workshop for the last two days is just the right thing to do in order to discover the low energy neutrons of solar origin. If they are detected, a faster than r^2-rise in the counting rate of the detector towards the sun would of course confirm the discovery.

These neutrons are likely to carry very useful information about the nature, the locale, the lifetime and the mode of operation of the acceleration processes in the solar atmosphere. My modest hope is that the information that we gather on this mission will lead to a rapid development of the science of solar meteorology which in turn might one day make it possible to forecast solar activity more reliably, on a long-term basis.

This afternoon I have avoided talking about the desirable detector characteristics that might be used on the exploratory mission. We propose to do this at a later date. My aim was more limited, in that I wanted to convince this audience that if and when solar probe mission is undertaken, search for low energy solar neutrons must be one of its prime objectives.

Acknowledgements. My travel to this Workshop is made possible in part by the generous support received from the University of New Mexico.

References.

Ahluwalia, H.S., 1969. J. Geophys. Res., 74, 1230.

Ahluwalia, H.S., 1971. Twelfth Intern. Conf. Cosmic Rays, Hobart. Conference Papers (University of Tasmania), 2, 468.

Ahluwalia, H.S., 1975a. Proc. Conf. Nuclear Cross Sections and Technology, Washington D.C., NBS SP 425, 2, 512. Eds. Schrack and Bowman.

Ahluwalia, H.S., 1975b. Fourteenth Intern. Conf. Cosmic Rays, Munich. Conference Papers (Max-Planck-Institut fur Extraterrestrische Physik), 12, 4251.

Apparao, M.V.K., Daniel, R.R., Vijayalakshmi, B., and Bhatt, V.L., 1966. J. Geophys. Res., 71, 1781.

Bame, S.J., and Asbridge, J.R., 1966. J. Geophys. Res., 71, 4605.

Biermann, L., Haxel, O., and Schulter, A.Z., 1951. Naturforsch., 6a, 47.

Chupp, E.L., 1963. AAS-NASA Symposium on the Physics of Solar Flares, Greenbelt. NASA SP-50, p. 445. Ed. W.N. Hess.

Chupp, E.L., 1971. Space Sci. Rev., 12, 486.

Chupp, E.L., Forrest, D.J., Higbie, R.R., Suri, A.N., Tsai, C., and Dunphy, P., 1973. Nature, 241, 333.

Cortellessa, P., Benedetto, P., and Palzis, C., 1971. Solar Phys., 20, 474.

Daniel, R.R., Joseph, G., Lavakare, P.J., and Sunderrajan, R., 1971. J. Geophys. Res., 76, 3152.

Eyles, C.J., Linney, A.D., and Rochester, G.K., 1971. Twelfth Intern. Conf. Cosmic Rays, Hobart. Conference Papers (University of Tasmania), 2, 462.

Haymes, R.C., 1964. J. Geophys. Res., 69, 853.

Heidbreder, E., Pinkau, K., Reppin, C. and Schonfelder, V., 1970. J. Geophys. Res., 75, 6347.

Hess, W.N., 1962. Proc. Fifth InterAmerican Seminar on Cosmic Rays, La Paz, 1, 17.

Hess, W.N., and Kaifer, R.C., 1967. Solar Phy ., 2, 202.

Leavitt, C.P., Robb, D.S., and Young, F., 1972. Bull. Am. Phys. Soc., 17, 687.

Lingenfelter, R.E., Flamm, E.J., Canfield, E.H., and Kellman, S., 1965a. J. Geophys. Res., 70, 4077.

_____, 1965b. J. Geophys. Res., 70, 4087.

Lingenfelter, R.E., and Ramaty, R., 1967. High Energy Nuclear Reactions in Astrophysics. Ed. B.S.P. Shen. W.A. Benjamin Press, New York, p. 99.

Lockwood, J.A., Ifedili, S.O., and Jenkins, R.W., 1973. Solar Phys., 30, 183.

McCracken, K.G., and Rao, U.R., 1970. Space Sci. Rev. 11, 155.

Moon, S., Simnett, G.M., and White, R.S., 1976. Astrophys. J., 207, 630.

Pottasch, S.R., 1967. Bull. Astron. Inst. Netherlands, 19, 113.

Ramaty, R., Kozlovski, B., and Lingenfelter, R.E., 1975. Space Sci. Rev., 25, 461.

Reppin, C., Chupp, E.L., Forrest, D.J., and Suri, A.N., 1973. Thirteenth Intern. Conf. Cosmic Rays, Denver. Conference Papers (University of Denver), 2, 1577.

Webber, W.R., and Ormes, J.F., 1967. J. Geophys. Res., 72, 3387.

Wolcott, J.H., 1968. Ph.D. Dissertation submitted to the University of New Mexico.

SECTION VI.

SOLAR WIND

SOLAR WIND SCIENCE FOR A SOLAR PROBE MISSION

William C. Feldman
University of California, Los Alamos Scientific Laboratory
Los Alamos, New Mexico 87545

ABSTRACT

Scientific objectives for solar-wind investigations relevant to a Solar Probe mission are outlined and put in perspective. Information gained from the various possible measurements are also listed.

C-4

Over the past 20 years, research workers in the field of interplanetary physics have made substantial progress toward understanding the solar wind phenomenon. Many experiments conducted in space since and including the Mariner-II mission to Venus have provided a fairly complete observational description of solar wind characteristics in the ecliptic plane between about 0.3 and 10 AU (see, for example, reviews by Neugebauer, 1975; Montgomery, 1976; Feldman et al., 1977; Behannon, 1978). Plans are presently being made to extend this description to high heliographic latitudes.

Concurrently, theoretical tools were developed to understand the physics of the coronal expansion and to organize the mass of observations into a few simple models (see, for example, Parker, 1963; Hundhausen, 1972; Holzer, 1976, 1978). For example, given the fact that a hot corona surrounds the Sun, the necessity for a supersonic expansion is presently understood from both fluid dynamic and exospheric points of view. Additionally, simple fluid models have been developed which are capable of describing the evolution of large scale plasma structures injected into the ambient solar wind flow by virtue of temporal or spatial variations in coronal conditions.

Despite all of this effort, several basic aspects of the solar wind flow remain unresolved. Foremost is that neither the state of the coronal plasma nor the microphysical processes which affect the flow in interplanetary space are known well enough to define a unique model for the acceleration of the solar wind.

The primary gap in our understanding of this acceleration is knowledge of the evolution of the various forms of solar energy flux leading to a supersonic expansion. Several such forms have been suggested previously and are schematically illustrated in Figure 1. Distance and vector length scales are not indicated because details of the coronal expansion have not yet been experimentally verified. It is thought, but not yet directly observed, that waves from the Sun penetrate the chromosphere-corona transition region to heat

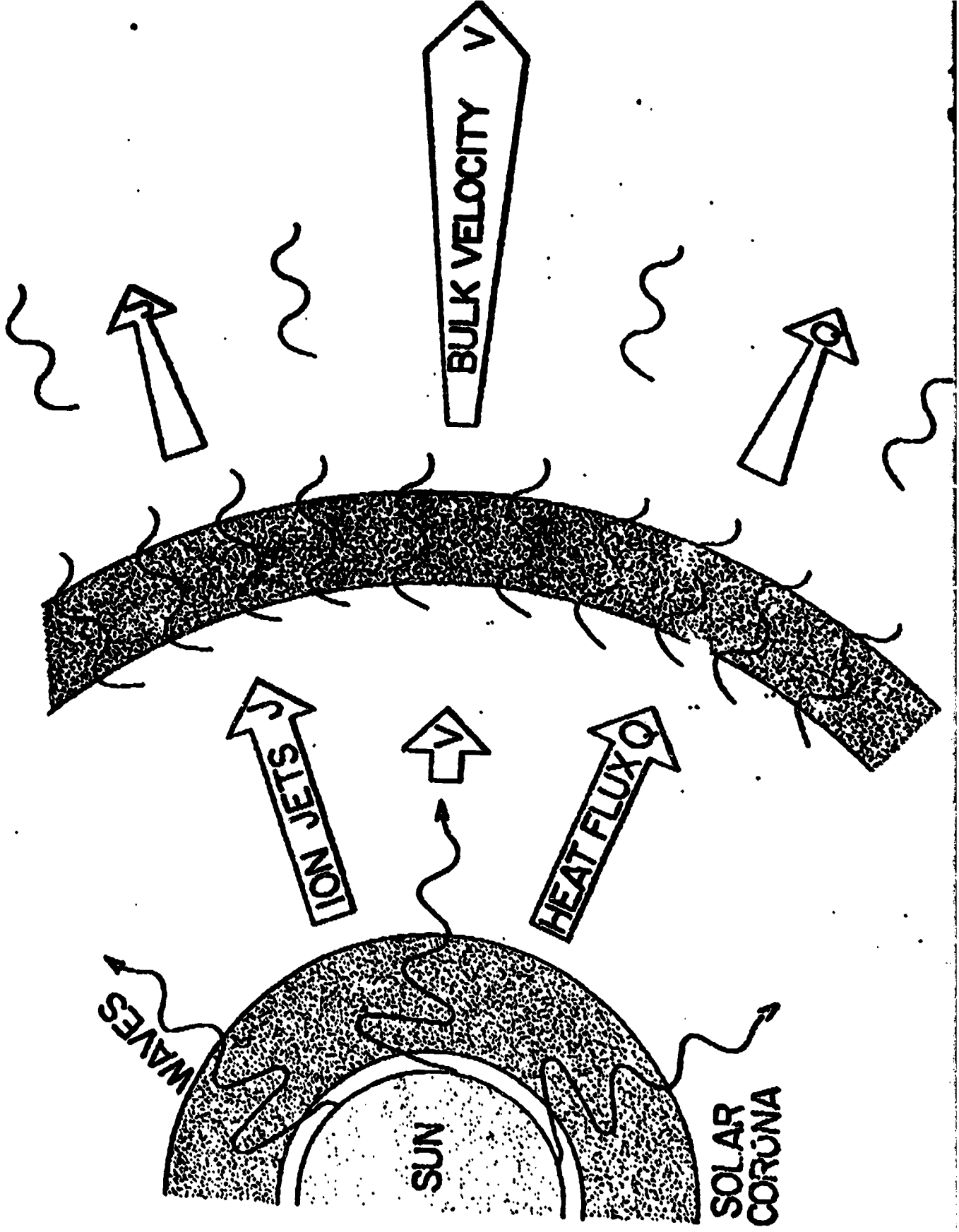


Fig. 1. A schematic illustration of the various forms of solar energy flux leading to s supersonic expansion

the solar corona. Energy carried by the waves into the low corona flows outward in the form of other waves, electron heat flux, convected enthalpy flux, convected kinetic and gravitational potential energy fluxes, and perhaps ion jets. This situation is indicated in the left hand portion of the figure. The narrow white ring between the Sun and the solar corona is meant to represent the chromosphere-corona transition region. The decreasing amplitude of the wavy vectors emanating from the Sun represents wave damping and a consequent heating of the coronal gas. Part of this wave energy may escape into interplanetary space as indicated by extensions of the wavy vectors beyond the coronal base (the stipled ring just outside the disc of the Sun) (see, for example, Barnes, 1974; Hollweg, 1975). The waves may then either damp and thereby heat an extended volume of the inner solar wind (Barnes, 1974) or directly accelerate the flow through their pressure gradient (Hollweg, 1978). The hot coronal gas drives an outward heat flux (arrow labeled Q), a bulk convection (represented by the arrow labeled V) which is slow at first but increases with increasing distance from the Sun, and perhaps transient jets of ions (Brueckner et al., 1977) (represented by the arrow labeled J) flowing faster than, and through, the relatively slowly expanding ambient plasma.

The internal state of the ambient plasma changes slowly with distance from the Sun. At some distance (indicated by the outer stipled ring) this state may have evolved to the point where waves are driven unstable by the electron heat flux and relative ion streaming motion (see, for example, Dobrowolny and Moreno, 1977; Feldman, 1978). The resulting enhancement of oscillating wave fields should add to the turbulence already present in the form of waves propagating outward from the Sun thereby increasing particle scattering rates and enhancing the conversion of electron heat conduction and relative ion streaming energy fluxes to random internal energy and ultimately to convection energy flux. This possibility is indicated schematically by the increased number of wave symbols superimposed on the outer stipled ring beyond which the ion jet

and heat flux vectors decrease and the velocity vector increases in breadth and length. In actual fact, this ring may overlap the region where waves of solar origin are still significant and damping.

Theories of the steady state, structure-free solar wind start above the coronal temperature maximum and explore the consequences of: 1) a thermal pressure gradient, 2) electron heat flux under the assumption of various conductivity laws, 3) a wave pressure gradient leading to direct acceleration, 4) wave damping contributing to a thermal pressure gradient and acceleration, and 5) various degrees of magnetic field line spreading (see, for example, reviews by Holzer, 1978, and Hollweg, 1978). Despite much effort, no unified and unique theory of the outer solar atmosphere is generally accepted. The primary lesson here is that sufficiently many steady-state components may possibly cooperate to accelerate coronal plasma that observation of the flow state at 1 AU is not sufficient to define the precise mix of mechanisms operating at any one time.

Causes for observed variations in this flow state are consequently not known. The problem is compounded by the fact that although high speed flows are known to evolve from coronal holes, the origin of low speed flows is uncertain. Suggested coronal footpoints of low-speed solar wind include: 1) the edges of coronal holes, 2) quiet coronal regions, 3) active regions, and 4) streamers. An additional complication concerns the role of transient phenomena in determining the structure of the solar atmosphere through their contribution of coronal heating and solar wind acceleration.

Most of the present ambiguities can only be resolved by complete measurements of the plasma flow and wave field characteristics for heliocentric distances inside about 20 solar radii. For example, radial profiles of the following parameters should be measured: 1) the solar wind density and bulk speed to determine the radial extent of the region of bulk acceleration, 2) the electron velocity distribution to determine a conductivity law including the

role of supra-thermal electrons in carrying energy flux, the gradient of the interplanetary electrostatic potential and the electron thermal pressure gradient, 3) the proton velocity distribution to determine the proton contribution to the thermal pressure gradient as well as to provide a signature of wave damping and ion jets, 4) the alpha-particle velocity distribution to provide some information concerning the nature and damping rate of the high frequency wave field as well as the degree to which alpha particles load the bulk expansion, 5) the hydromagnetic wave field to determine the extent and magnitude of direct wave acceleration and wave damping leading to acceleration through the intermediary of a thermal pressure gradient, 6) the D.C. magnetic field to allow estimatesof the Alfven speed and the position of the critical point as well as to facilitate an identification of the coronal footpoints of the flow, and 7) the heavy ion ionization state densities to provide an estimate of the electron temperature and temperature gradient inside of 4 solar radii.

All these measurements can and should be made aboard one and preferably two solar probe spacecraft with perihelion as close to the Sun as possible. Two spacecraft will be very important for separating spatial and time variations for defining the wave field close to the Sun.

Although the problem of the acceleration of the solar wind is fundamental to progress in the field of coronal and interplanetary physics, several other problems are also important and can be resolved with Solar-Probe instrumentation. These include a determination of: 1) the solar wind angular momentum flux and consequent solar spin down rate, 2) the character of the expansion within streamers and above active regions, 3) the sharpness of boundaries between flows close to the Sun which evolve from different coronal regions, 4) the origin of various slow speed flows including those density enhanced flows observed at 1 AU which do not appear to result from a compression in interplanetary space, 5) the extent of heavy element fractionation as well as the mix of physical processes causing spatial and/or temporal abundance varia-

tions, 6) the origin of hydromagnetic waves in the solar wind, and 7) the mechanisms of hydromagnetic wave damping when wave amplitudes are large.

In summary, many of the fundamental questions concerning solar wind acceleration have not been answered to date but can be successfully resolved with state of the art plasma and field measurements made from one, and preferably two, Solar-Probe spacecraft. Consequently this mission should receive the highest priority in current NASA planning.

Acknowledgements

I am grateful to J. T. Gosling for useful conversations. This work was performed under the auspices of the U. S. Department of Energy.

References

Barnes, A., Theoretical studies of the large scale behavior of the solar wind, advan. Electron, Electron Phys., 36, 1, 1974.

Behannon, K. W., Heliocentric distance dependence of the interplanetary magnetic field, Rev. of Geophys. Space Phys., 16, 125, 1978.

Brueckner, G. E., J. D. F. Bartoe, and M. E. Van Hoosier, High spatial resolution observations of the solar EUV spectrum, in "Proceedings of the OSO-8 Workshop," E. Hansen and S. Schoffner, eds., Boulder, CO, 1977.

Dobrowolny, M. and G. Moreno, Plasma kinetics in the solar wind, Space Sci. Rev., 20, 577, 1977.

Feldman, W. C., Kinetic processes in the solar wind, in Solar System Plasma Physics: A Twentieth Anniversary Overview, C. F. Kennel, L. J. Lanzerotti, and E. N. Parker, eds., North Holland, 1978.

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, O. R. White, ed., Col. Assoc. Univ. Press, Boulder, CO, 1977, p. 351.

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

Hollweg, J. V., Some physical processes in the solar wind, to be publ. in Rev. of Geophys. Space Phys., 1978.

Holzer, T. E., Our quantitative understanding of the coronal expansion, in Physics of Solar Planetary Environments, D. J. Williams, ed., AGU, 1976, p. 366.

Holzer, T. E., The solar wind and related astrophysical phenomena, to be publ. in Solar System Plasma Physics: A Twentieth Anniversary Overview, C. F. Kennel, L. J. Lanzerotti, and E. N. Parker, eds., 1978.

Hundhausen, A. J., Solar Wind and Coronal Expansion, Springer-Verlag, N. Y. Heidelberg, Berlin, 1972.

Montgomery, M. D., Solar wind observations throughout the solar system, in Physics of Solar Planetary Environments, D. J. Williams, ed., AGU, 1976, p. 304.

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

Parker, E. N., Interplanetary Dynamical Processes, Interscience Publ., New York-London, 1963.

Plasma Waves Near the Sun: Advances Possible with a Solar Probe

by

Donald A. Gurnett

Department of Physics and Astronomy
The University of Iowa
Iowa City, Iowa 52242

ABSTRACT

A review is presented of current experimental and theoretical knowledge of plasma waves in the solar wind, with comments on the scientific importance of obtaining plasma wave measurements in the region near the sun with the solar probe. The waves discussed include the acoustic waves and shocks which are thought to be the primary source for heating the solar corona, Alfvén waves, ion-cyclotron waves, whistler-mode turbulence, ion-acoustic waves, and electron plasma oscillations associated with solar radio emissions. A discussion is presented of the types of measurements which would be needed to study these waves on the solar probe, the constraints imposed on the spacecraft and the research and development which would be needed to provide the necessary instrumentation.

I. INTRODUCTION

For many years it has been widely accepted that waves play a fundamental role in the energy transport and microscopic structure of the solar corona and the solar wind. To explain the very high temperatures of the solar corona, Biermann [1948] and Schwarzchild [1948] proposed that waves propagating outward from the photosphere provide the primary mechanism for heating the solar corona. Later Parker [1958] in his model of the solar wind assumed the presence of waves to provide a thermalizing process which would maintain the fluid-like properties of the collisionless solar wind plasma. Early radio observations by Wild [1950] and others also provided direct evidence of various complex plasma wave and radio emission processes occurring at high frequencies following solar flares and other transient events at the sun. With the advent of direct in situ spacecraft measurements, our knowledge of waves in the solar wind has increased rapidly. It is now known that an extraordinarily wide variety of plasma waves occur in the solar wind near the orbit of the earth, extending from very long wavelength Alfvén waves, with periods of minutes to hours, to short wave length electrostatic waves with frequencies near the local electron plasma frequency. These waves have many important effects on the physics of the interplanetary medium.

Although much is known concerning the waves which exist in the solar wind near the orbit of the earth, many questions remain concerning the waves which occur much closer to the sun. At the present time the closest measurements to the sun have been obtained from the Helios 1 and 2 spacecraft at approximately 0.3 AU. As will be discussed in the next section most of the waves detected by Helios show a strong radial dependence, increasing rapidly in intensity with decreasing radial distance from the sun. Presumably these waves have increasingly important effects on the solar wind plasma in the unexplored region inside of 0.3 AU. Furthermore, some of the most important waves, which play an important role in heating the solar wind can only be observable well inside of 0.3 AU. The solar probe mission which will provide measurements in to 4 solar radii (0.019 AU) represents an ideal opportunity to investigate the role of waves in the solar wind and solar corona, particularly in the very important region near about 7 solar radii where the solar wind becomes supersonic. In this paper we will briefly review the types of waves which are expected in the solar wind on the basis of present theoretical and experimental ideas, and comment on the scientific relevance of providing measurements of these waves on the solar probe. In addition, we will comment on the type of measurements which need to be made, the constraints imposed on the spacecraft and the research and development which needs to be done to provide the necessary instrumentation.

II. A SURVEY OF SOLAR WIND PLASMA WAVES AND THE CONTRIBUTION A SOLAR PROBE MISSION COULD MAKE TO THE INVESTIGATION OF THESE WAVES

In Figure 1 we attempt to summarize the various types of plasma waves which are either known to occur or are theoretically predicted to occur in the solar corona and solar wind. The frequency range of interest extends over a very large range, from well below ion cyclotron frequency to above the electron plasma frequency. The approximate radial variation of the various characteristic frequencies of the plasma are indicated in Figure 1. The wave phenomena associated with each frequency range will be discussed in order of increasing frequency, starting with the lowest frequencies closest to the sun.

A. Acoustic and Shock Wave Heating of the Solar Corona

The basic energy source for heating the solar corona is believed to be acoustic waves which are propagating upward into the chromosphere and corona from the photosphere. (For a review of this process see Brandt, 1970, and Bray and Loughhead, 1974.) The existence of these waves was first proposed by Biermann [1948] and Schwarzschild [1948] to explain the rapid rise in the temperature from the photosphere to the solar corona (see Figure 2). Since the temperature in this transition region increases with increasing radial distance thermal conduction and radiation transfer cannot explain the observed temperature rise.

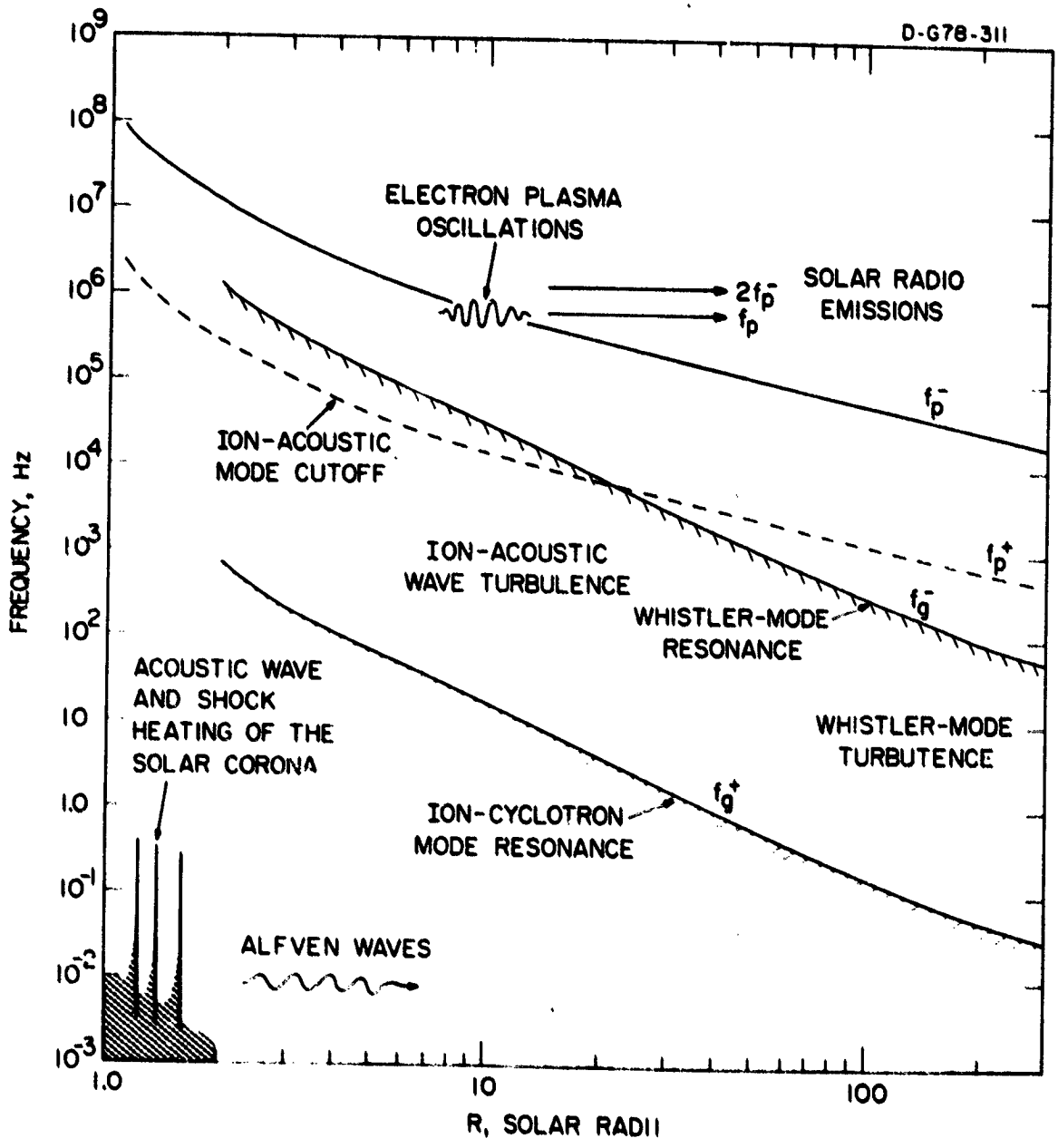


Figure 1 The heliocentric radial variation of the various characteristic frequencies of the plasma in the solar wind and solar corona.

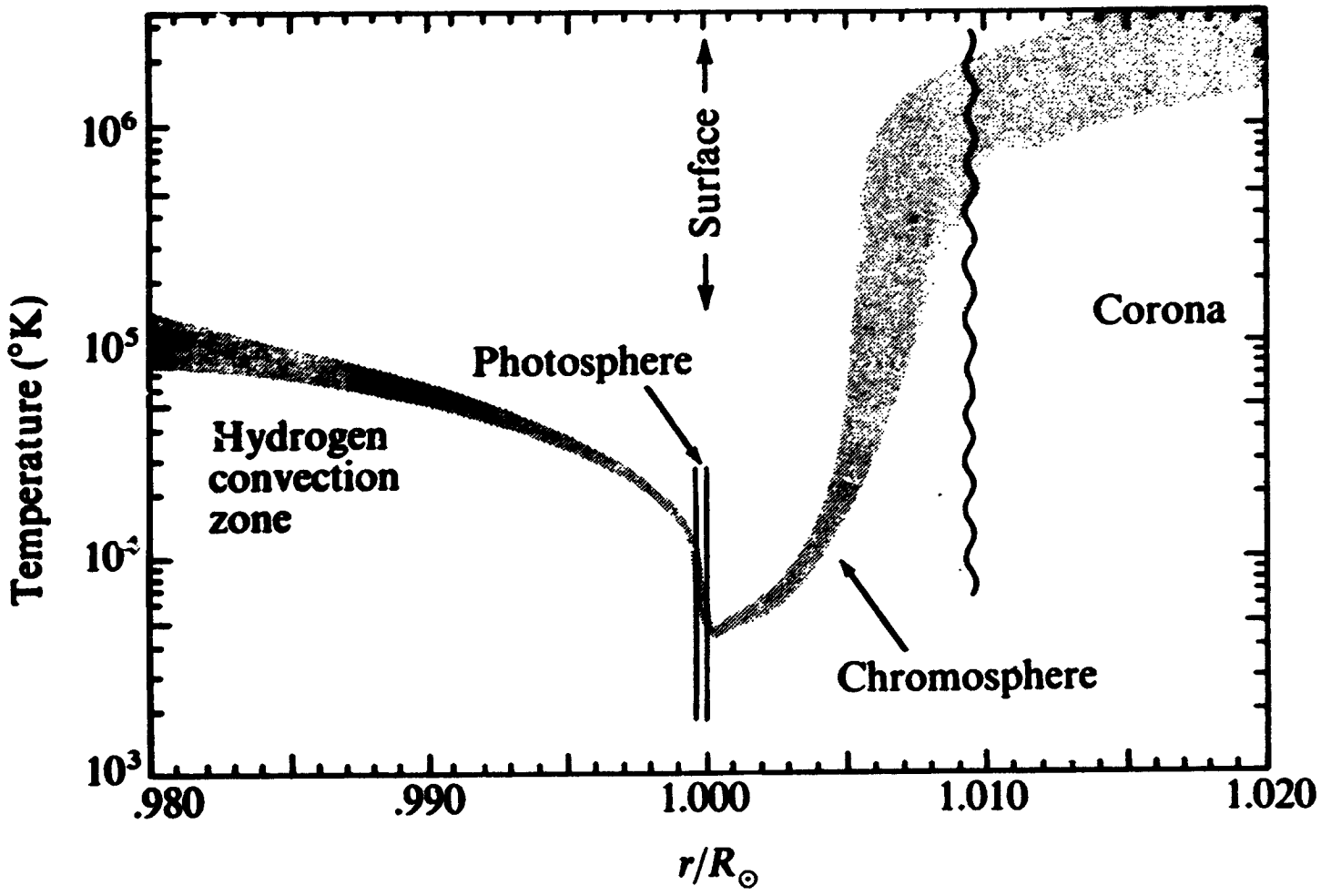


Figure 2 The radial variation of the plasma temperature near the sun from Brandt [1970]. The steep rise in the temperature in the chromosphere cannot be explained by thermal conduction or radiation transfer and is widely attributed to heating by waves propagating upward from the photosphere.

Some other energy transport mechanism, such as waves, is therefore needed to account for the high temperature of the corona. The waves responsible for the coronal heating are thought to be long wavelength acoustic waves generated by turbulence convection cells below the photosphere (see Figure 3). The periods of these waves are greater than about 100 sec. As currently understood, the acoustic waves generated by this turbulence propagate upward into the chormosphere and are absorbed, thereby heating the plasma. Details of the absorption process are uncertain. In one model, the waves grow in amplitude as they propagate upward into regions of decreasing density and are eventually converted to shocks which dissipate their energy in the lower regions of the corona. Just how far these waves propagate into the solar corona is unknown. According to the classical hydrodynamic picture the waves are strongly damped in the lower corona and would not be observed by the solar probe at 4.0 solar radii. However, when magnetohydrodynamic effects are considered it is possible that the acoustic waves couple to a relatively undamped mode, such as the transverse Alfvén wave which could propagate over long distances. Confirmation of the basic wave heating mechanism for the solar corona is extremely important, since at the present time no evidence exists which directly demonstrates the presence of large amplitude acoustic waves at the base of the solar corona. Although the solar probe does not pass close enough to the sun to obtain direct in situ measurements in the region where the maximum heating takes place, the possibilities for detecting some associated wave activity propagating out to ~ 4 solar radii are probably very high.

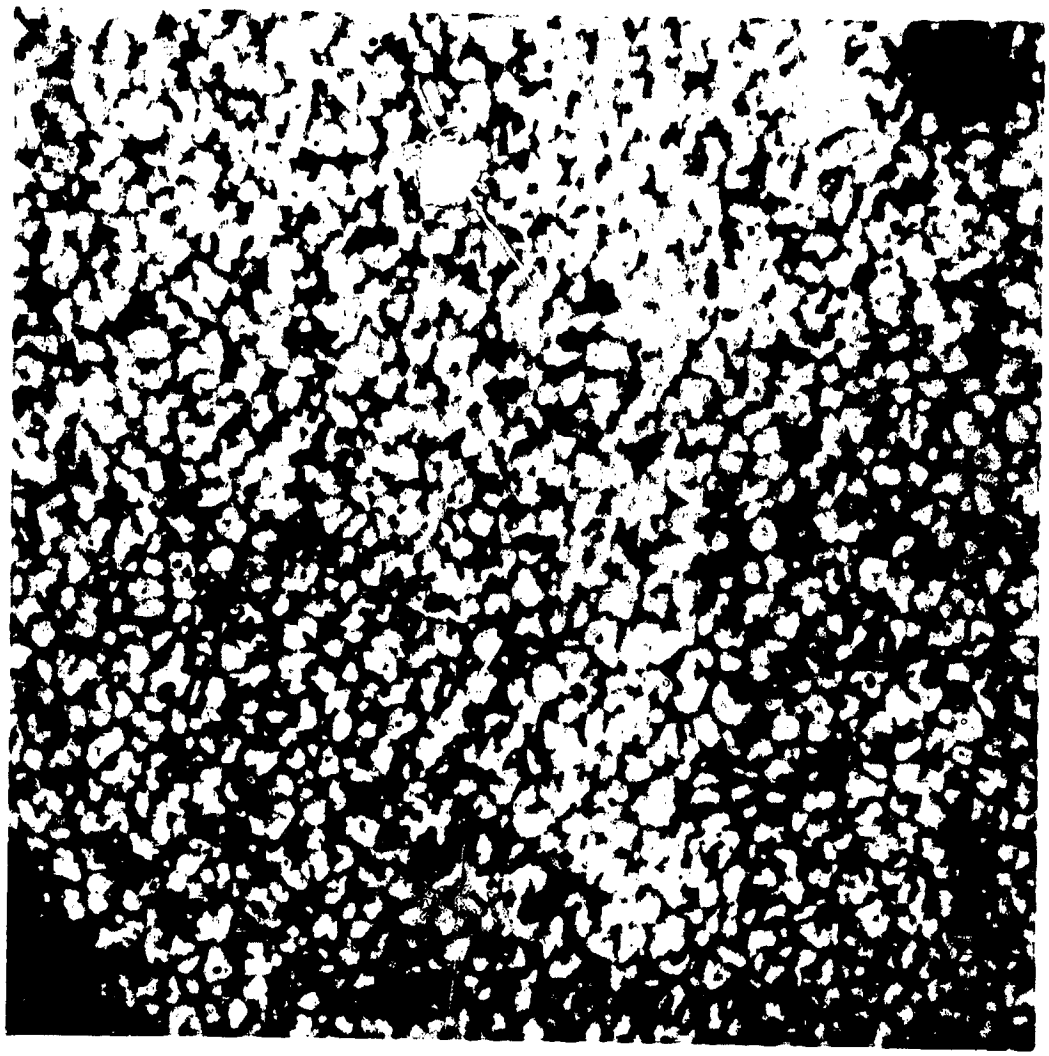


Figure 3 A photograph of the turbulent convection cells which are believed to produce the waves which heat the solar corona. The waves produced by this turbulence are believed to be acoustic waves with periods of about 100 sec. Photograph from Brandt [1970].

B. Alfvén Waves

It has been known for many years that large amplitude Alfvén waves are commonly observed in the solar wind near 1 AU [Coleman, 1966]. A typical example of these waves, showing the close correlation of the magnetic field $\vec{b}$ and plasma flow velocity $\vec{v}$, is shown in Figure 4, from Belcher and Davis [1971]. For a review of these and other waves in the solar wind, see Hollweg [1975]. As can be seen the waves often have very large amplitudes at 1 AU, with magnetic field amplitudes in some cases comparable to the ambient solar wind magnetic field strength. Comparisons of the phase of the magnetic field and plasma flow velocity perturbations show that the waves are always propagating outward from the sun, with no component propagating back toward the sun. This observation is used by Belcher and Davis to argue that the waves must be generated inside of the Alfvén critical point at ~ 7 solar radii. If the source of the waves is located beyond the Alfvén critical point, then the sunward propagating components should be observed at 1 AU, since all waves generated beyond the critical point are transported outward from the sun by the solar wind flow. However, if the source is located inside of the critical point, the sunward propagating waves would propagate in to the photosphere and be lost by collisional damping, thereby explaining their absence at 1 AU. Since the Alfvén waves are weakly damped and are apparently generated very close to the sun it seems quite likely that they may be a by-product of the intense acoustic wave heating which occurs at the base of the solar corona.

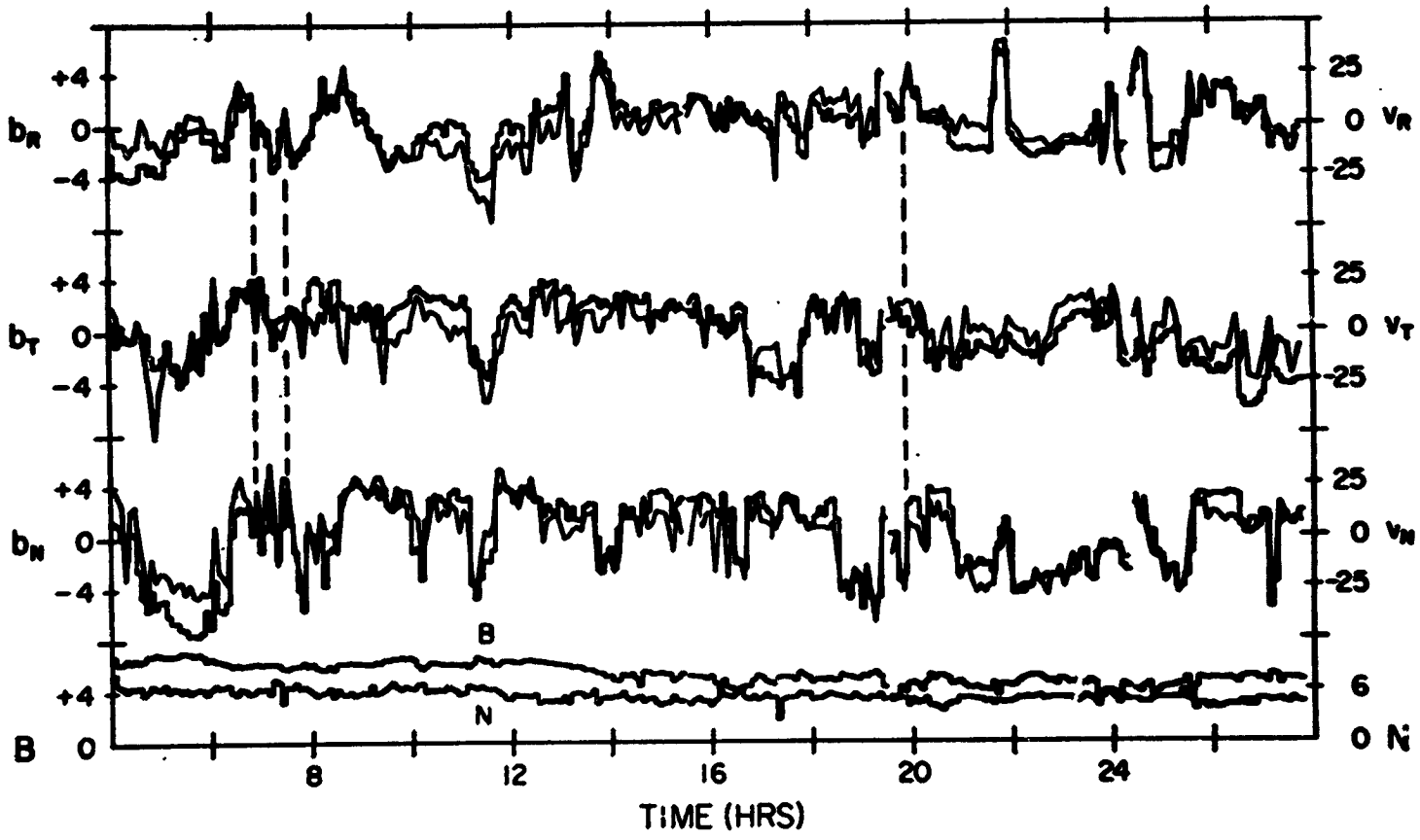


Figure 4 An example of large amplitude Alfvén waves observed in the solar wind from Belcher and Davis [1971]. The phase between the magnetic field $\vec{b}$ and plasma velocity $\vec{v}$ fluctuations indicates that the Alfvén waves are propagating outward from the sun.

Because of the large amplitude of the Alfvén waves observed at 1 AU, several authors have suggested that the damping of these waves, by various mode conversion processes, may contribute significantly to the heating of the solar corona out to large distances from the sun, ~ 10 to 20 solar radii [Hartle and Barnes, 1970]. This possibility is important for understanding the overall structure of the solar wind since various models of the solar wind flow show disagreements with observations which could be reconciled by a relatively small 10% heat input into the solar wind at radial distances out to about 20 solar radii [see the discussion by Hundhausen, 1972]. At the present time very little is known about the radial dependence of the Aflvén wave amplitudes inside of 1 AU. Appropriate wave measurements on the solar probe mission would provide a very important contribution to our understanding of the solar wind by determining the relative contribution of magnetohydrodynamic waves to the solar wind heating in the region close to the sun.

C. Ion-Cyclotron Waves

Extending upward in frequency from the Alfvén wave portion of the spectrum the first plasma resonance effect to be encountered is at the proton cyclotron frequency, f_g^+. As discussed by Stix [1962], the slow (shear) Alfvén mode has a resonance at f_g^+, and does not propagate at frequencies greater than f_g^+. Near the proton cyclotron frequency this mode is called the ion cyclotron mode. Considerable theoretical interest has been placed on the ion cyclotron mode because of possible resonant wave-particle interactions with ions in the solar wind.

Observationally, it is difficult to establish the presence of ion-cyclotron waves in the solar wind because the large Doppler shifts tend to spread out the resonance effects at f_g^+. However, solar wind magnetic field spectrums, such as in Figure 5 from Behannon [1976], consistently show a break in the slope of the spectrum near f_g^+, which strongly suggests a propagation cutoff at the proton cyclotron frequency. Behannon [1976] also presents arguments for the occurrence of ion cyclotron waves in the solar wind on the basis of polarization measurements.

Because of the decreasing proton gyrofrequency with increasing radial distance from the sun ion cyclotron waves are generated as a natural consequence of the outward propagation of Alfvén waves from the sun, similar to the "magnetic beach" discussed by Stix [1962]. Ion cyclotron waves can also be produced by ion cyclotron mode instability when $T_{\|} < T_{\perp}$ [Gary et al., 1976]. Since ion cyclotron waves can interact resonantly with alpha particles and other heavy ions as well as with protons, it is thought that these waves may play an important role in the acceleration of alpha particles and other heavy ions [Hollweg and Turner, 1978], which are observed to have anomalously high flow speeds at 1 AU. To fully evaluate these ion-cyclotron wave interaction effects magnetic field spectrums similar to those in Figure 5 need to be obtained at radial distances much closer to the sun.

D. Whistler-Mode Turbulence

At frequencies above the proton gyrofrequency, the next higher resonance frequency to be encountered is the electron gyrofrequency, f_g^-.

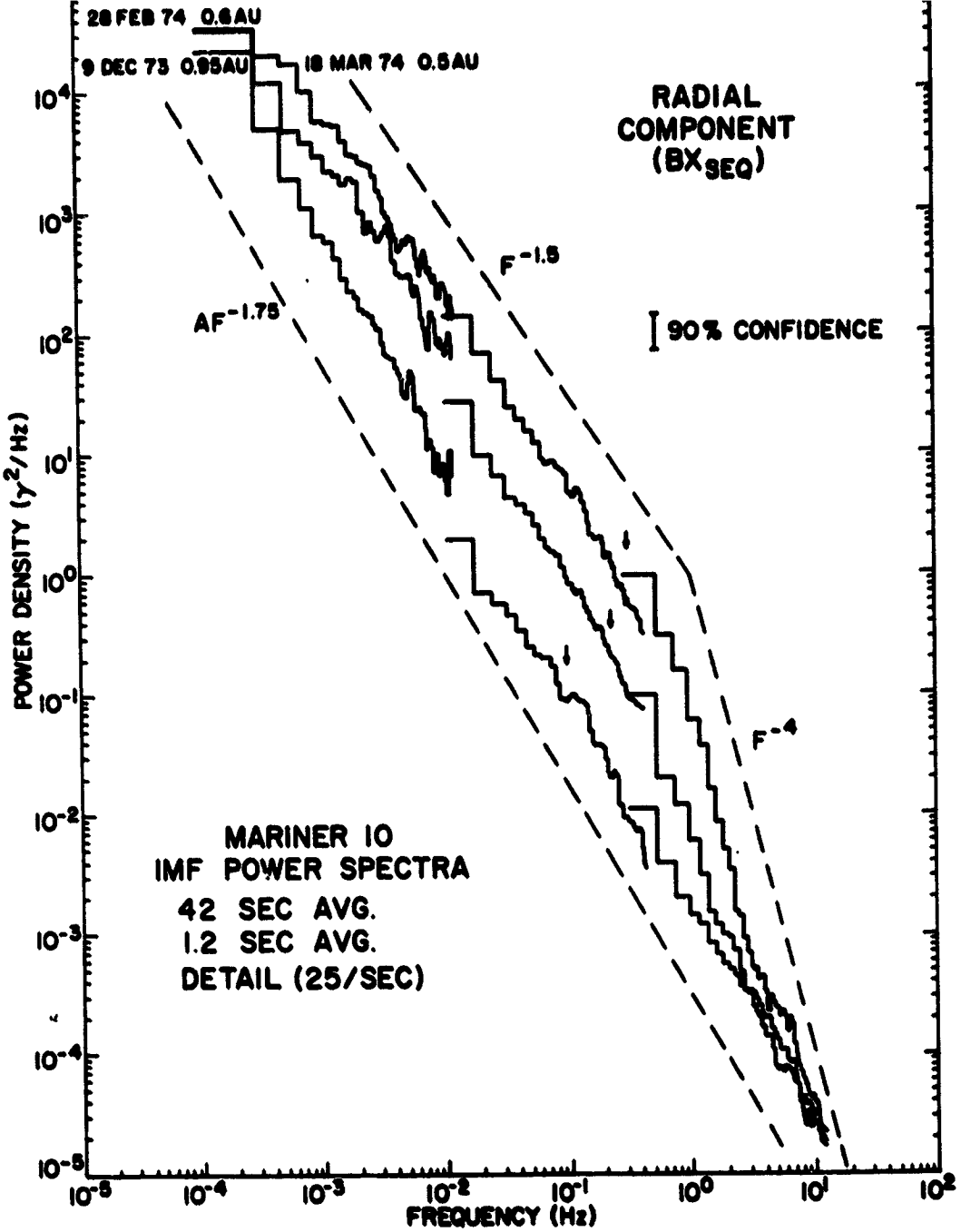


Figure 5 Magnetic field spectrums in the solar wind given by Behannon [1976] from the Mariner 10 spacecraft. The break in the slope of the spectrum near ~ 1.0 Hz is believed to be due to the ion cyclotron resonance at f_g^+ (see Figure 1), which represents the upper limit of the ion cyclotron mode. Because of Doppler shifts the cyclotron resonance is not sharply defined.

For frequencies between f_g^+ and f_g^- the only electromagnetic mode which can propagate is the right-hand polarized whistler mode. This mode represents the upward extension of the fast (compressional) Alfvén mode. Recent measurements from the Helios spacecraft have shown that a weak nearly steady level of magnetic field turbulence is present in the solar wind at frequencies up to near the electron gyrofrequency [Neubauer et al., 1977]. Considerations of the probable Doppler shift and frequency range, between the electron and ion gyrofrequencies, strongly suggest that this turbulence is caused by whistler-mode waves. A typical example of the magnetic field intensities associated with this whistler-mode turbulence is shown in Figure 6, from Neubauer et al. [1977]. As can be seen, the intensity of this turbulence increases rapidly with decreasing frequency and with decreasing radial distance from the sun.

Several possible mechanisms for generating unstable whistler-mode waves in the solar wind have been proposed, including thermal anisotropies in the ion distribution function [Scarf et al., 1967] and instabilities driven by the electron heat flux [Gary et al., 1976]. At the present time it seems most likely that the whistler-mode turbulence detected in the solar wind by Helios is produced by an electron heat flux instability, although other possibilities such as the nonlinear cascading of long wavelength high amplitude Alfvén waves cannot be entirely ruled out. Since whistler-mode waves can interact resonantly with electrons it is possible that this turbulence could play an important role in the pitch angle scattering of solar wind electrons or in controlling the thermal conductivity of the solar

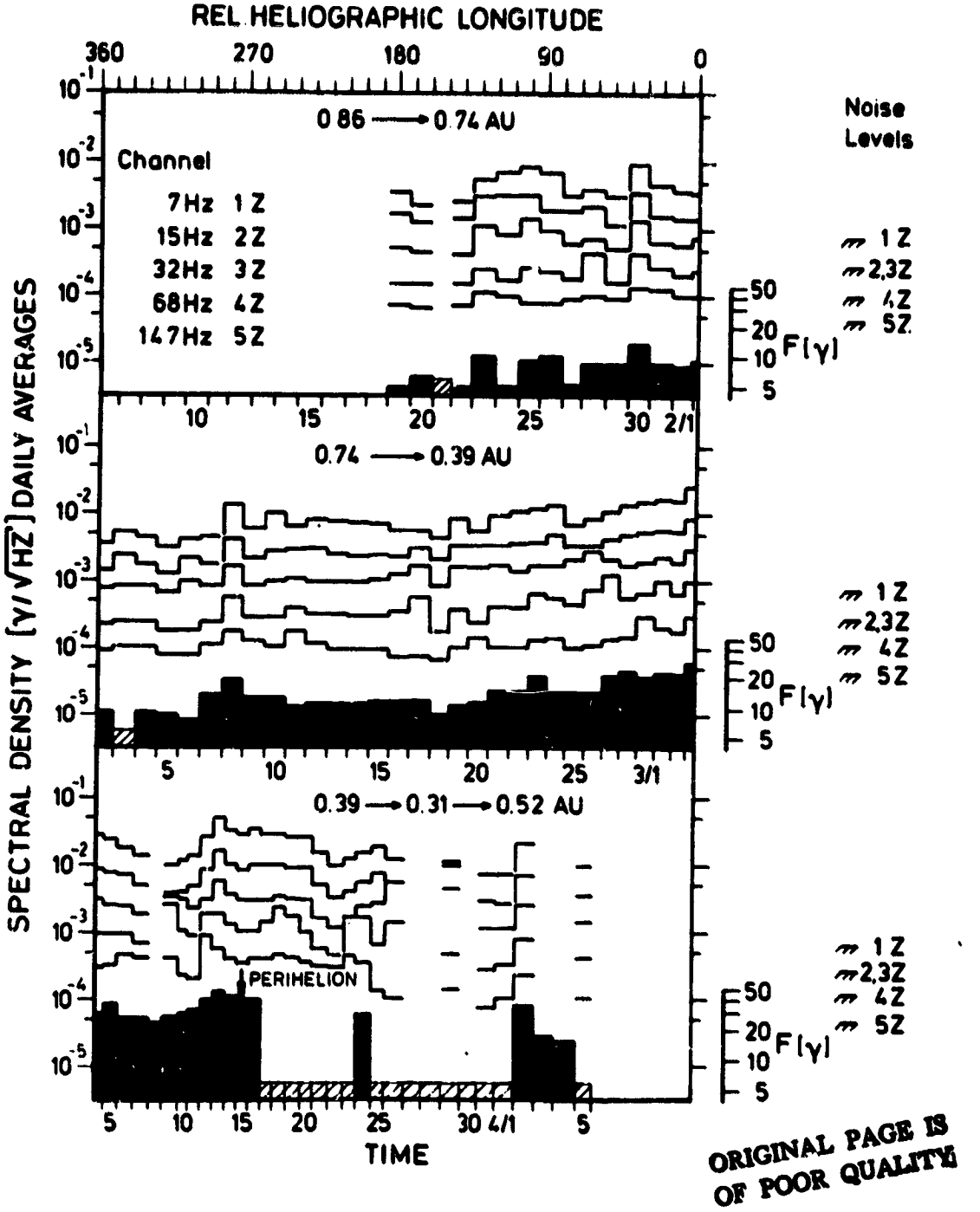


Figure 6 Magnetic field fluctuations in the frequency range between the proton and electron gyrofrequencies observed by the Helios-1 spacecraft, from Neubauer et al. [1977].

wind [Gary and Feldman, 1977]. For a review of these and other kinetic processes in the solar wind see Feldman [1978]. Since the intensity of the whistler-mode turbulence detected by Helios is found to increase rapidly with decreasing radial distance from the sun, it is important to determine the intensity of these waves and their effect on the solar wind electron distribution inside of 0.3 AU. These questions can only be answered by obtaining whistler-mode measurements closer to the sun.

E. Ion-Acoustic Waves

In laboratory plasmas one of the most commonly observed instabilities is the electrostatic ion-acoustic mode, which is driven unstable by relatively weak currents when $T_e \gg T_i$. Since T_e is usually only slightly larger than T_i in the solar wind, it was initially somewhat uncertain whether ion-acoustic waves would be generated in the solar wind. Fredricks [1969] predicted that ion-acoustic waves should be observed in association with shocks and other discontinuities where large currents exist, and Forslund [1970] proposed that ion-acoustic waves should be generated by an electron heat flux instability whenever T_e/T_i is sufficiently large. Early investigations of plasma wave electric fields by Scarf [1970] and Siscoe *et al.* [1971] suggested the occasional occurrence of ion-acoustic waves in association with interplanetary shocks and other discontinuities. The interpretation of these electric fields was, however, complicated by Doppler shifts and poor frequency resolution which made a precise identification of the plasma wave mode impossible. Later, more sensitive electric field measurements with the Helios and Voyager spacecraft revealed the occurrence of

relatively weak electrostatic waves in the solar wind at frequencies between the electron and ion plasma frequencies, f_p^- and f_p^+, which were interpreted as being ion-acoustic (or ion-acoustic-like) waves by Gurnett and Anderson [1977]. Also, see the discussions by Gurnett and Frank [1978] and Kurth et al. [1978]. Typical electric field spectrums of these waves at various radial distances from the sun are shown in Figure 7 and a high resolution frequency-time spectrogram obtained from the Voyager-1 spacecraft is shown in Figure 8. Although the ion-acoustic mode only propagates at frequencies below the ion plasma frequencies the wavelengths are so short that the wave frequency is mainly determined by the Doppler shift which is much larger than the ion plasma frequency. The systematic increase in the upper cutoff frequency with decreasing radial distance from the sun, evident in Figure 8, is caused by the radial variation in the Debye length, which controls the maximum possible Doppler shift for these waves.

At the present time the exact mechanism by which these ion-acoustic waves are generated has not been definitely established. The most likely possibility is that these waves are produced by double-ion streams in the solar wind [Gary, 1978] or by the electron heat flux instability proposed by Forslund [1970]. Since the exact mechanism by which these waves are produced is not known, it is difficult to determine what effect these waves may have on the macroscopic structure of the solar wind. If the waves are produced by double-ion streams one interesting possibility is that they may play an important role in coupling or thermalizing the double ion streams. Beyond 0.3 AU the

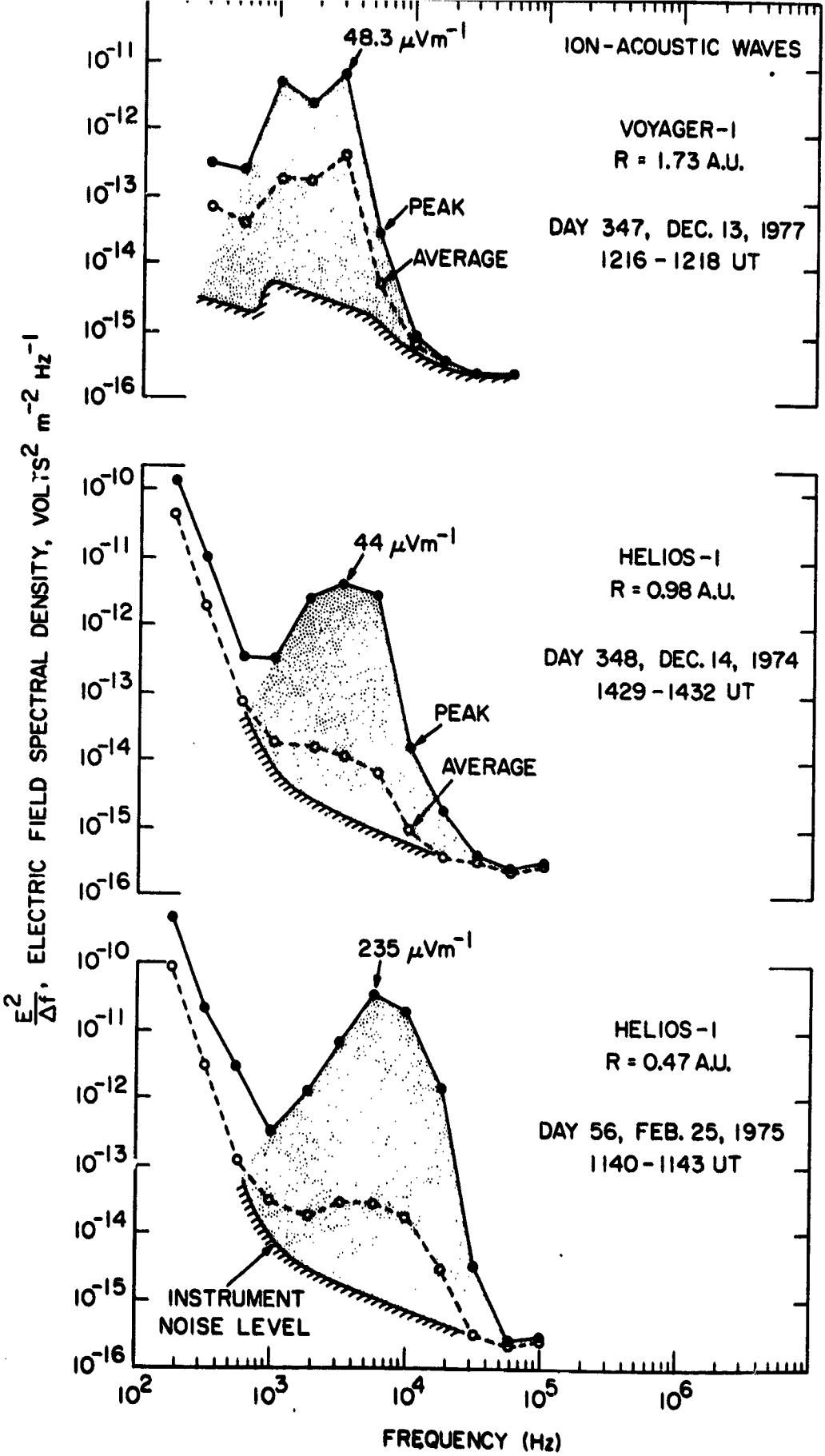


Figure 7 Representative spectrums of ion-acoustic waves observed from the Helios-1 and Voyager-1 spacecraft. The short wavelength ion-acoustic waves are strongly Doppler shifted, so that the peak in the spectrum occurs well above the ion-acoustic wave cutoff at f_p^+ (see Figure 1).

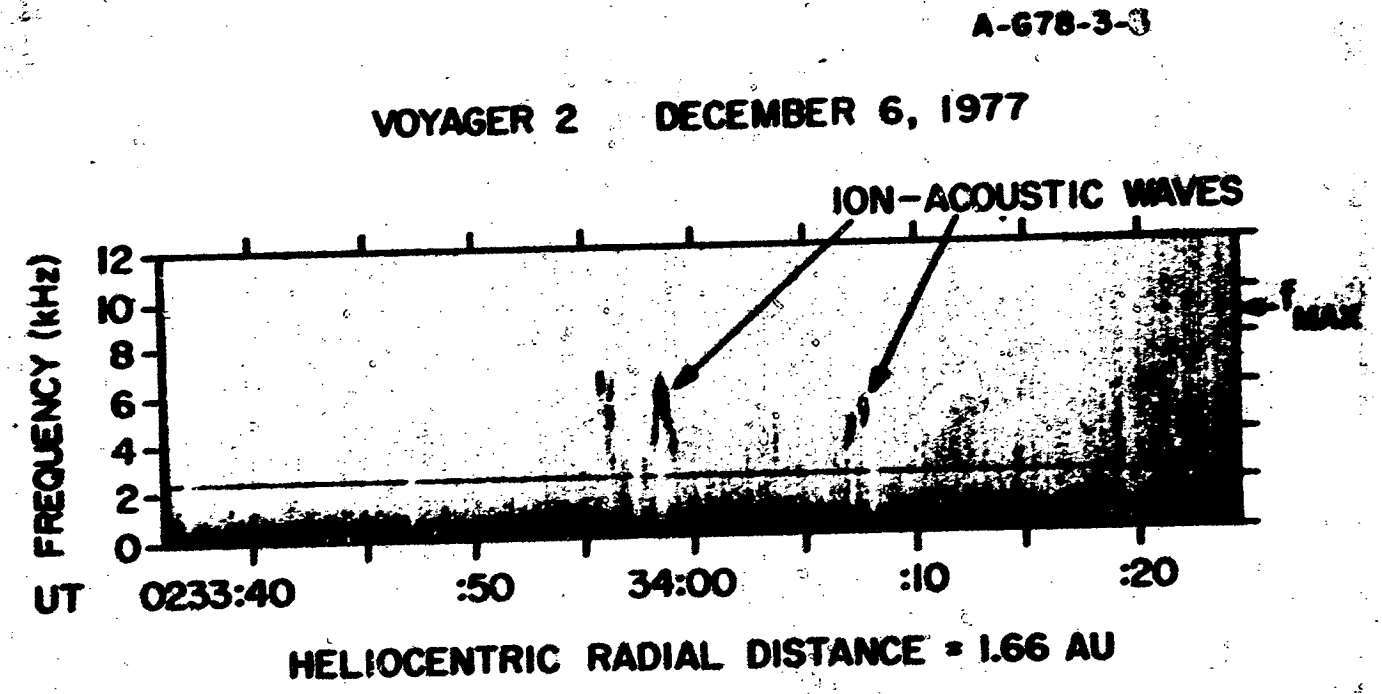


Figure 8 A high resolution spectrogram of some ion-acoustic wave bursts observed by Voyager-2.

energy density of these waves, $E^2/8\pi\ nkT \simeq 10^{-5}$, is sufficiently small that it seems unlikely that they play a major role in processes of this type. However, as shown in Figure 7, the electric field strength increases rapidly with decreasing radial distance from the sun, so that closer to the sun the intensity of these waves may be large enough to play a significant role in controlling the macroscopic properties of the solar wind.

F. Electron Plasma Oscillations and Solar Radio Bursts

The currently accepted model for the generation of several types of solar radio bursts is that the radio emissions are produced by non-linear processes involving electron plasma oscillations excited by solar flare electrons and other transient effects in the solar corona. The electron plasma oscillation mechanism, first proposed by Ginzburg and Zheleznyakov [1958], has become the basic element of essentially all theories of solar radio bursts, with suitable refinements to account for various types of non-linear interactions. Although the plasma oscillation mechanism has been accepted for many years only in the last two years have measurements been obtained from the Helios spacecraft which definitely established the existence of these electron plasma oscillations [Gurnett and Anderson, 1976]. A typical example of the electron plasma oscillations detected by Helios, and the resulting type III radio burst are shown in Figure 9. As indicated in Figure 1, current theories of the non-linear interactions predict radio emission at both the electron plasma frequency, f_p^-, and the second harmonic, $2f_p^-$. At

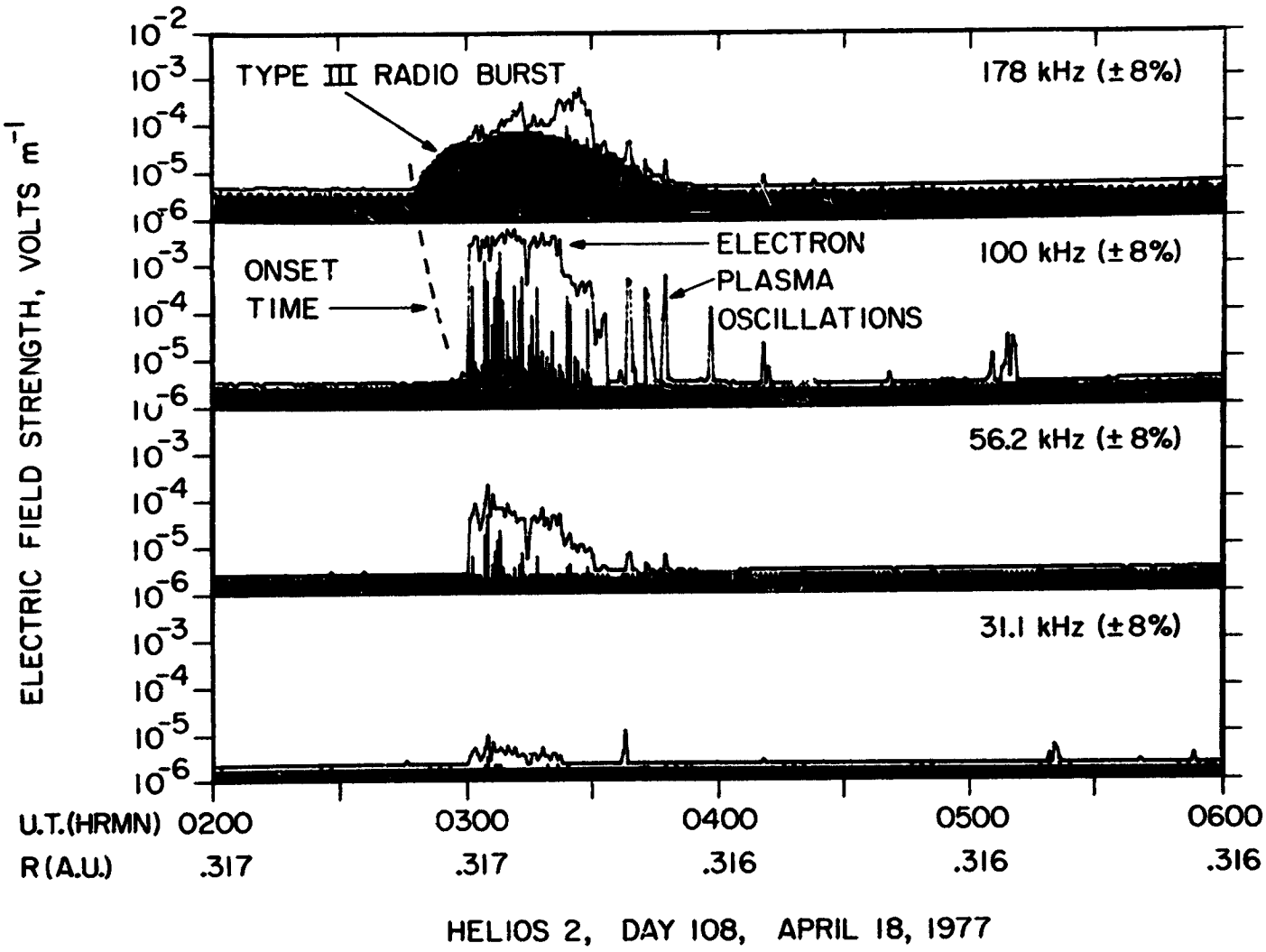


Figure 9 An intense burst of electron plasma oscillations observed in association with a type III solar radio burst. The electron plasma frequency, f_p, in this case was approximately 100 kHz.

low frequencies, $\leq$ 100 kHz, it is generally believed that the second harmonic emission dominates, whereas at high frequencies, $\geq$ 1 MHz, both the fundamental and the second harmonic are generated. This transition in the emission characteristics apparently takes place in the radial distance range from about 5 to 20 solar radii because of changes in the characteristics of the plasma oscillations. Measurements of the plasma oscillation intensities associated with type III bursts obtained from the Helios and Voyager spacecraft show a very pronounced variation with radial distance from the sun, increasing very rapidly in intensity with decreasing radial distance from the sun. This radial dependence is illustrated in Figure 10, from Gurnett et al. [1978]. Based on these results it appears quite likely that the plasma oscillation intensities continue to increase inside of 0.3 AU, possibly reaching amplitudes which are a substantial fraction, 10^{-2} to 10^{-3}, of the beam energy density in the region close to the sun. The expected increase in the plasma oscillation intensities near the sun is also consistent with the general trend for the radio emission intensities to increase at higher frequencies, corresponding to source positions closer to the sun. At these high intensities, if they occur, wave-particle interactions are expected to play a significant role in the evolution and dynamics of the electron beam responsible for the plasma oscillations. See for example the discussions by Papadopoulos et al. [1974] and Magelssen and Smith [1977]. To completely understand the generation of solar radio bursts and the role electron plasma oscillations play in controlling the evolution of energetic electron streams emitted from

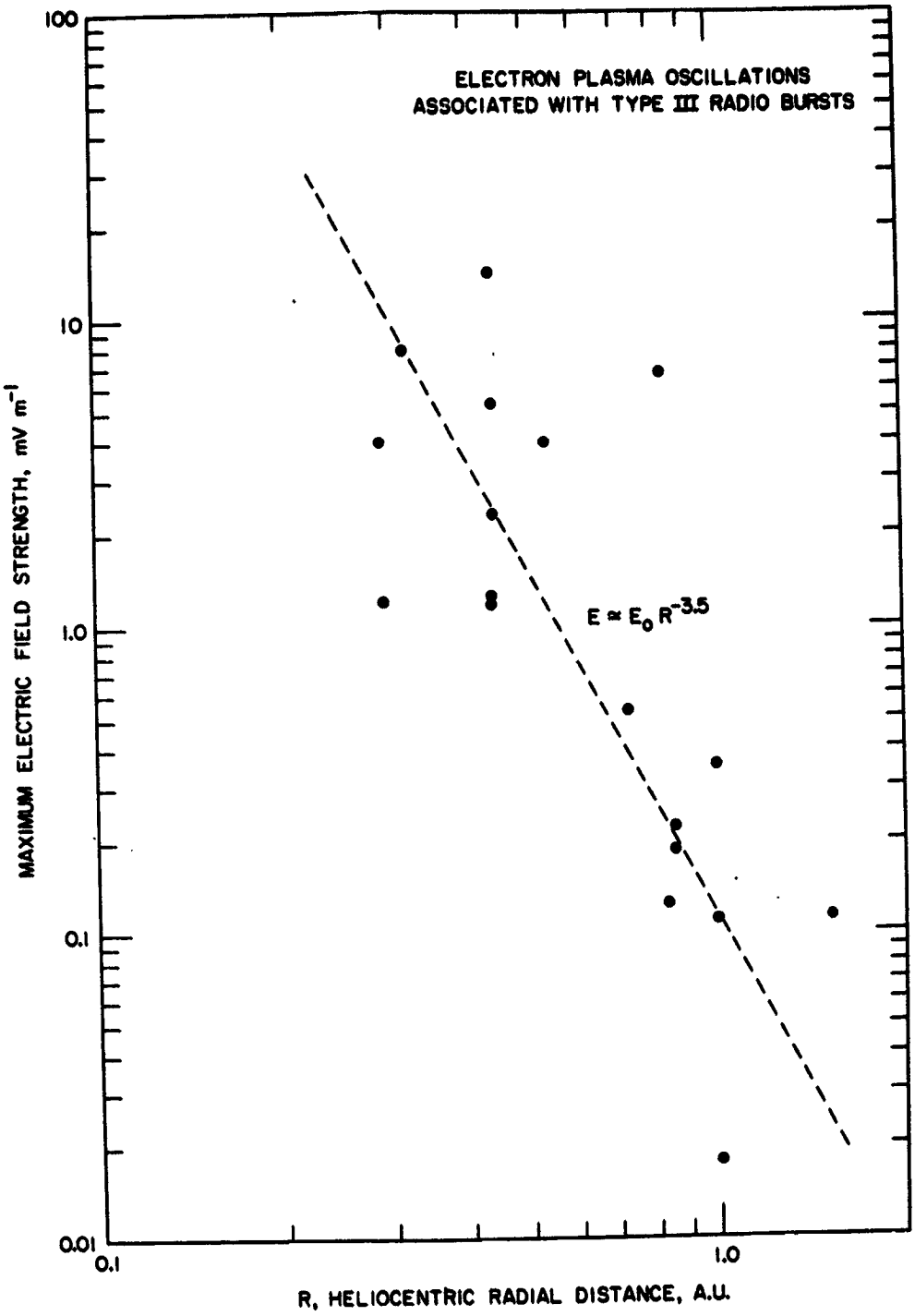


Figure 10 The radial variation of electron plasma oscillation intensities observed in association with type III solar radio bursts. The plasma oscillation intensities increase rapidly with decreasing radial distance from the sun.

solar flares it is essential that direct in situ measurements be obtained closer to the sun, well inside 0.3 AU. The solar probe mission provides an ideal opportunity to extend our knowledge of these emission processes. Since the investigation of these waves requires the occurrence of a suitable transient event near the sun (solar electron event or interplanetary shock) the likelihood of obtaining a suitable event during the few hours when the spacecraft is in the close vicinity of the sun, $\leq$ 10 solar radii, is not very high. However, during 1989 the sun will be approaching solar maximum and one would expect to observe several type III radio bursts on the solar disk every day. Thus, over a period of several days there is a good chance of obtaining observations of the plasma oscillations associated with a type III radio burst at radial distances inside of $\sim$ 30 to 50 solar radii. These investigations are clearly enhanced by observations for more than one perihelion pass.

III. WHAT TYPES OF WAVE MEASUREMENTS SHOULD BE INCLUDED ON THE SOLAR PROBE

A brief look at Figure 1 shows that the most demanding feature of a solar probe plasma wave investigation is the extremely large frequency range which must be covered, from less than 10^{-2} Hz up to several MHz. It is also important that both electric and magnetic fields be measured over this entire frequency range. Since both electrostatic and electromagnetic modes of propagation exist in the solar wind, simultaneous electric and magnetic field measurements are needed to distinguish these two fundamentally different modes of propagation. Electric field measurements can also be used to distinguish convective magnetic field structures from electromagnetic waves on the basis of the observed electric to magnetic field ratios, and comparisons of the phase of the electric and magnetic fields can answer the important question of whether the waves are propagating away from or toward the sun.

Because of the extremely large frequency range which must be covered it is extremely unlikely that all of the wave measurements can be provided by a single instrument. Practically, a natural break in the techniques for making magnetic field measurements occurs at about 5 Hz, with search coil magnetometers providing better sensitivity above about 5 Hz, and other types of DC magnetometers providing better sensitivity below 5 Hz. A similar break also occurs for the electric field measurements, with frequencies above ~ 5 Hz favoring on-board

spectrum analysis and frequencies below ~ 5 Hz favoring A-D conversion and direct transmission of the waveforms to the ground. Over the frequency range from about 5 Hz to a few MHz both the electric and magnetic field measurements can, and probably should, be performed in the same instrument.

The electric field measurements pose a difficult problem in the antenna design. For high sensitivity and protection from spacecraft generated interference a relatively long dipole antenna is needed, with a tip-to-tip length of about 30 meters, comparable to the electric antennas used on Helios and Voyager. However, close to the sun electrostatic waves with wavelengths less than 30 meters may be present. The shortest wavelength which can occur in a plasma is approximately $2\pi \lambda_D$, where λ_D is the Debye shielding length. A plot of the minimum wavelength, $2\pi \lambda_D$, is shown in Figure 11. As can be seen electrostatic waves with wavelengths as short as 1 meter could be encountered near perihelion at 4 solar radii. This minimum wavelength implies an antenna with a tip-to-tip length of less than 1 meter. Although the minimum wavelength condition provides a strong case for a relatively short antenna, experience with the sensitivity of such short antennas and the typical electric field intensities observed by Helios at 0.3 AU indicate that it would be very unwise to rely solely on such a short antenna for all the electric field measurements. Many of the most important electric fields occurring in the solar wind, such as the whistler-mode noise and ion-acoustic waves, probably could not be detected with an antenna length of 1 meter, given the interference

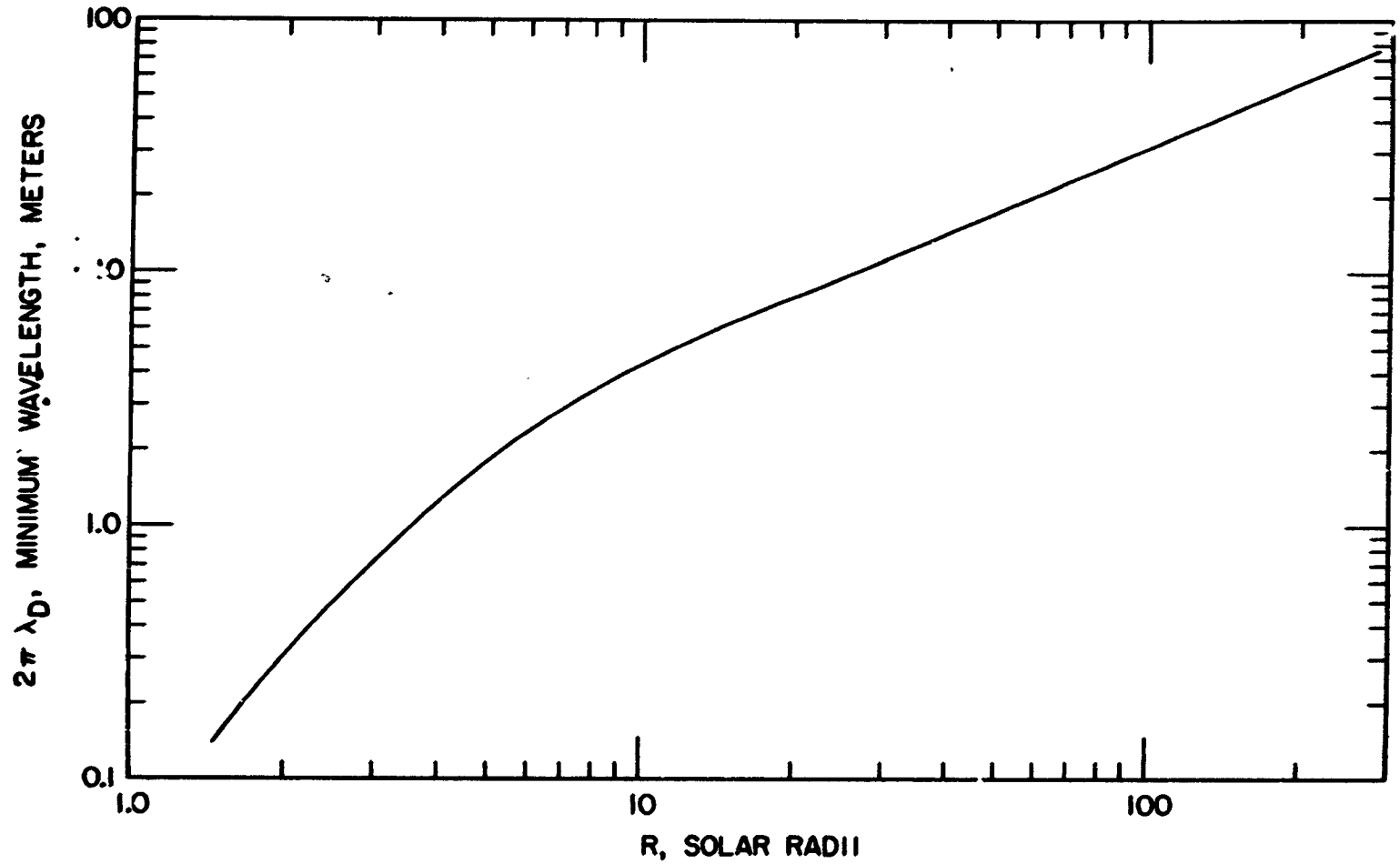


Figure 11 The minimum wavelength, $2\pi \lambda_D$, of plasma waves in the solar wind as a function of heliocentric radial distance. To detect all possible wavelengths in to 4 solar radii the electric antenna should not be longer than about 1 meter. Considerations of sensitivity and protection from spacecraft generated interference suggest that a somewhat longer antenna, $\sim$ 30 meters tip-to-tip, should also be used. This antenna would be used for detecting low intensity waves such as whistler-mode turbulence, ion-acoustic waves and solar radio bursts which probably could not be detected with a short 1-meter antenna.

levels expected on a spacecraft such as the solar probe. On the basis of these considerations, two different electric antennas, one with a tip-to-tip length of about 30 meters and the other with a tip-to-tip length somewhat less than 1 meter (0.5 meters nominal) are recommended for the baseline design considerations. Two antennas of different lengths have the additional advantage that the occurrence of short wavelength ($\lambda < 30$ meters) could be detected by comparing the field strengths observed by two antennas. For this purpose the axes of the long and short antennas should be parallel.

To provide wave magnetic field measurements over the extremely large frequency range from 5 Hz to a few MHz, it is virtually certain that at least two different types of antennas must be used, probably a search coil with a μ-metal core for frequencies less than 10 kHz, and either a search coil with a ferrite core or a loop antenna above 10 kHz. To reduce the magnetic field interference from the spacecraft to an acceptable level the magnetic antennas must be mounted on a boom at least 2 and preferably 4 meters from the main spacecraft body. Because of the limited temperature range of the preamplifiers for the magnetic field sensors and the short electric antenna these elements must be mounted in the shade behind the heat shield. A possible mounting configuration is shown in Figure 12.

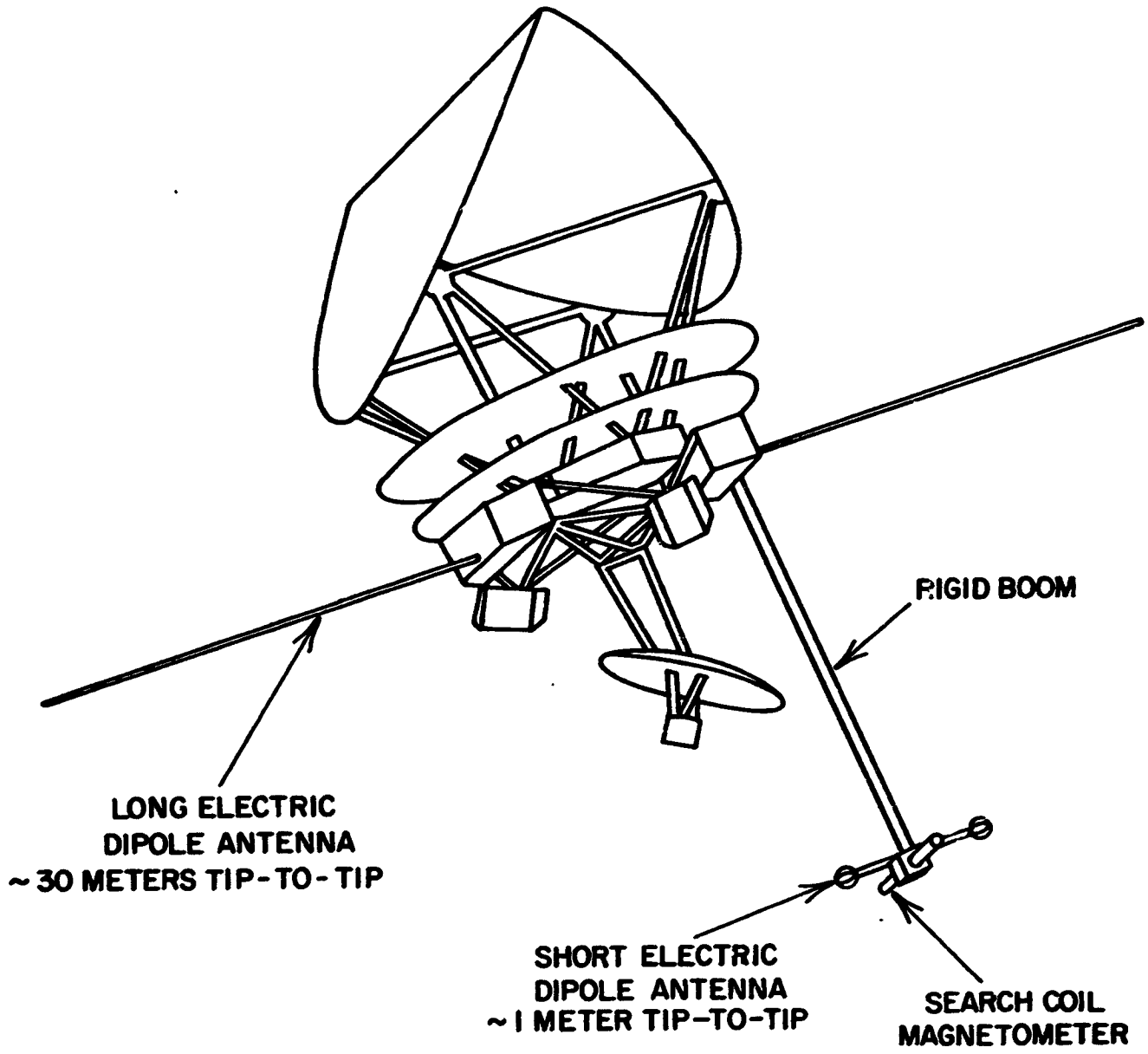


Figure 12 Suggested locations for mounting the plasma wave electric and magnetic field antennas.

IV. WHAT CONSTRAINTS WOULD A PLASMA WAVE EXPERIMENT PLACE ON THE SPACECRAFT

A. Electromagnetic Interference

Because of the susceptibility of the plasma wave experiment to electromagnetic interference special attention must be given to electromagnetic interference control from both the spacecraft and the experiments. These controls need not be exceptionally stringent, since very low interference levels were achieved on the Voyager spacecraft with only minimal constraints. The primary objective should be to shield the time varying electric fields generated by the spacecraft electronics. At low frequencies there is no way to effectively shield the magnetic field, so the only possibility for the magnetic field measurements is to mount the magnetic field sensors on a boom as far as possible from the spacecraft and to reduce the area of current loops in the spacecraft by using twisted leads for power lines and other similar techniques. For the electric field excellent shielding can be obtained by enclosing all electronics inside of a grounded "watertight" electrostatic shield, with special attention given to any cables which extend outside of this shield. A grounded electrostatic shield will probably be needed in any case to protect the spacecraft from differential charging effects at Jupiter.

B. High Telemetry Rates and Data Storage

The obtain very high resolution frequency-time spectrograms, such as in Figure 8, it is desirable to have the capability of obtaining very high data rate, 100 to 500 kb/s, snapshots of electric and magnetic field waveforms for short periods of time. This waveform data could be obtained either by real time telemetry or by use of an on-board memory. If the spacecraft bubble memory is used for this purpose the scientific return from the waveform measurements could be increased substantially if the memory is organized so that it can be used to find the best event (largest field strength) over the interval since the previous readout. This could be accomplished for example by using one of the bubble memories to store the previous best (largest) waveform while the other memory searches for the next larger event. Since the primary purpose of the spacecraft bubble memory is to store data during gaps between ground station coverage, it is anticipated that high-rate waveform readins of this type would only be performed on a non-interference basis during periods when real time telemetry is available.

C. Weight, Power and Telemetry

1. Weight

a. Main Instrument Package	3 kg
b. Boom Mounted Sensors	1 kg
c. Extendible Electric Antenna	2 kg (?)
	6 kg

2. Power 4 watts

3. Telemetry

a. Low Rate Telemetry
(continuous for the entire mission) 120 b/s.

b. High Rate Telemetry
(in short bursts, for approximately
100 sec.) $\geq$ 100 kb/s.

The primary developmental problem is the extendible 30-meter electric antenna, which would be required to survive at temperatures much higher than on any previous mission.

ACKNOWLEDGMENTS

This research was supported by the National Aeronautics and Space Administration under Grant NGL-16-001-043.

REFERENCES

Behannon, K., Observations of the interplanetary magnetic field between 0.46 and 1 AU by the Mariner 10 spacecraft, GSFC Preprint X-692-76-2, Greenbelt, MD, 1976.

Belcher, J. W. and L. Davis, Large-amplitude Alfvén waves in the interplanetary medium, 2, J. Geophys. Res., 76, 3534, 1971.

Biermann, L., Über die Ursache der Chromosphärischen Turbulenz und des UV-Exzesses der Sonnenstrahlung, Zs. f. Astrophys., 25, 11, 1948.

Brandt, John C., Introduction to the Solar Wind, W. H. Freemand and Co., San Francisco, CA, 1970.

Bray, R. J. and Loughhead, The Solar Chromosphere, Chapman and Hall, London, 1974.

Coleman, P. J., Jr., Hydromagnetic waves in the interplanetary plasma, Phys. Rev. Lett., 17, 207, 1966.

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

Feldman, W. C., Kinetic Processes in the Solar Wind, Solar System Plasma Physics -- A Twentieth Anniversary Review, C. F. Kennel and L. J. Lanzerotti, Eds., North-Holland, (in press), 1978.

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

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

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.

Gary, S. P. and W. C. Feldman, Solar wind heat flux regulation by the whistler instability, J. Geophys. Res., 82, 1087, 1977.

Gary, S. P., Electrostatic heat flux instabilities in the solar wind, J. Geophys. Res. (submitted for publication), 1978.

Ginzburg, V. L. and V. V. Zheleznyakov, On the possible mechanism of sporadic radio emission (radiation in an isotropic plasma), Sov. Astron. A. J., 2, 653, 1958.

Gurnett, D. A. and R. R. Anderson, Electron plasma oscillations associated with type III radio bursts, Science, 194, 1159, 1976.

Gurnett, D. A. and R. R. Anderson, Plasma wave electric fields in the solar wind: Initial results from Helios 1, J. Geophys. Res., 82, 632, 1977.

Gurnett, D. A. and L. A. Frank, Ion-acoustic waves in the solar wind, J. Geophys. Res., 83, 58, 1978.

Gurnett, D. A., R. R. Anderson, F. L. Scarf, and W. S. Kurth, The heliocentric radial variation of plasma oscillations associated with type III radio bursts, J. Geophys. Res., (submitted for publication), 1978.

Hartle, R. E. and A. Barnes, Nonthermal heating in the two-fluid solar wind model, J. Geophys. Res., 75, 6915, 1975.

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

Hollweg, J. V., and J. M. Turner, Acceleration of solar wind He^{++}. 3. Effects of resonant and nonresonant interactions with transverse waves, J. Geophys. Res., 83, 97, 1978.

Hundhausen, A. J., Solar Wind and Coronal Expansion, Springer-Verlag, N. Y., Heidelberg, Berlin, 1972.

Kurth, W. S., D. A. Gurnett, and F. L. Scarf, High resolution spectrograms of ion-acoustic waves in the solar wind, J. Geophys. Res. (submitted for publication), 1978.

Magelssen, G. R. and D. F. Smith, Nonrelativistic electron stream propagation in the solar atmosphere and type III radio bursts, Solar Physics, 55, 211, 1977.

Neubauer, F. M., H. J. Beinroth, H. Barnstorf, and G. Dehmel, Initial results from the Helios-1 search-coil magnetometer experiment, J. Geophys. Res., 42, 599, 1977.

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

Parker, E N., Dynamics of the interplanetary gas and magnetic fields, Ap. J., 128, 664, 1958.

Scarf, F. L., J. H. Wolfe, and R. W. Silva, A plasma instability associated with thermal anisotropies in the solar wind, J. Geophys. Res., 72, 993, 1967.

Scarf, F. L., Microscopic structure of the solar wind, Space Sci. Rev., 11, 234, 1970.

Schwarzschild, M., On noise arising from the solar granulation, Ap. J., 107, 1, 1948.

Siscoe, G. L., F. L. Scarf, I. M. Green, J. H. Binsack, and H. S. Bridge, Very low frequency electric fields in the interplanetary medium: Pioneer 8, J. Geophys. Res., 76, 828, 1971.

Stix, T. H., The Theory of Plasma Waves, McGraw-Hill, New York, 1962.

Wild, J. P., Observation of the spectrum of high-intensity solar radiation at metre wavelengths. III. Isolated bursts, Australian J. Sci. Res., Ser. A, 3, 541, 1950.

THE SOLAR PROBE AND CORONAL DYNAMICS

by

J. Belcher[+], M. Heinemann[*], C. Goodrich[+]

+ Department of Physics and Center for Space Research
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

* Department of Physics
Boston College
Chestnut Hill, Massachusetts

Abstract

The discovery of coronal holes led to basic changes in our ideas about the structure of the low corona and its expansion into the solar wind. The nature of the energy flux which provides the $\sim 5 \times 10^5$ ergs/cm^{2}sec required to drive the wind from coronal holes is not understood. Current ideas include enhanced thermal conductivities, extended MHD wave heating, and wave momentum transfer, all in rapidly diverging geometries. There is little feel for the relative importance of these processes. The Solar Probe, with its penetration deep into the solar corona, could lead to observational constraints on their relative importance, and thus to an understanding of the origin of the solar wind. Observations from the Solar Probe will also bear on such questions as to whether small scale "intrastream" structure is common close to the Sun in open field-line regions, whether the properties of the wind are pronouncedly different over closed and open field-line regions at five solar radii, and many others.

Experimentally, the resolution of these questions requires measurements of the magnetic field and of the proton and electron distribution functions. Obvious difficulties in these measurements include, for example: 1) limited coverage of phase space due to thermal constraints; 2) the transonic nature

of the proton flow close to the Sun; 3) the high velocity of the spacecraft; 4) the short measurement time near the Sun, especially when compared to the periods of low frequency MHD waves; and 5) possible high "turbulent" fluid velocities near the Sun, which may approach hundreds of kilometers per second.

Introduction

Observations from the Solar Prove will have a major impact on our understanding of the dynamics of coronal expansion. The recent discovery of coronal holes has led to a re-examination of our ideas about the basic structure of the lower corona and its expansion into the solar wind. Coronal hole observations have underscored our lack of understanding of the thermal and mechanical (wave) energy transport mechanisms in the corona, and how they lead to the solar wind velocities observed near Earth. The only way we can hope to understand the balance between various acceleration mechanisms is through in situ observations near the Sun, i.e. from the Solar Probe.

Before addressing the potential of the Solar Probe mission, it is useful to review the energies involved. The energy sources and sinks which determine the structure of the corona are usually taken to be: (1) the mechanical (wave) energy, ($\sim 5 \times 10^5$ ergs/cm^{2}sec) which is the basic energy source; (2) radiation; (3) thermal conduction; and (4) convection of energy by mass motion. In the classical picture, the majority of the mechanical energy that manages to make it through the transition region into the corona was conducted back down into the chromosphere and only a small amount was thought to be radiated by the corona ($\sim 5 \times 10^4$ ergs/cm^{2}sec) or conducted outward beyond the temperature maximum to ultimately become the solar wind ($\sim 5 \times 10^4$ ergs/cm^{2}sec). Thus, in classical solar wind models, the energy required to drive the solar wind was only $\sim$10% of the energy dissipated in the corona and the solar wind had little influence on the structure of the transition region and lower corona. The lower regions could be modeled by assuming hydrostatic equilibrium, with the solar wind tacked on beyond temperature maximum, with only modest energy requirements. Moreover, reasonable estimates of the conductive flux plus some extended heating beyond the temperature maximum gave approximately what was needed to explain the solar wind.

The discovery that the most energetic solar wind appears to come from

coronal holes has led to a critical re-examination of this model. In coronal holes, the maximum coronal temperature is lower by $\sim 6 \times 10^5$K, the temperature gradient in the transition region is less steep by a factor of ~ 4, and, therefore, the conductive flux downward is less. Most crucially, the rapid divergence of flow from the holes, a factor of ~ 7 faster than r^2, requires that the energy flux which ultimately drives the solar wind be on the order of 5×10^5 ergs/cm^2 sec at the Sun in order to explain what we see at 1 AU. Thus, in coronal holes, the solar wind is no longer a small perturbation to a static lower corona and the expansion may have a non-negligible effect on the structure below the temperature maximum. The need for a self-consistent treatment seems inescapable. Furthermore, the nature of the energy flux which provides this 5×10^5 ergs/cm^2-sec to the expansion is obscure. Can thermal conductivity in rapidly diverging geometries still provide the majority of that flux? If not, can extended wave heating and momentum transfer make up the difference, or perhaps even be dominant? If so, what are the wave modes and their dissipation lengths? How does the process of extended wave heating and subsequent thermal acceleration compare to direct momentum transfer from the waves to the wind? These are the central questions to our understanding of wind acceleration from coronal holes. There appears to be little hope of answering them exclusive of the Solar Probe.

The Potential of the Solar Probe

As far as the solar wind is concerned, the energy balance appears to be of prime interest for the Solar Probe. The Solar Probe will, of course, also bear on questions concerning spatial structure near the Sun. For example, small scale "intra-stream" structures may be common close to the Sun, and it may be possible to relate observations at 4 $R_\odot$ to features lower in the corona. Also, the flows from the edges of coronal holes and over active regions are of great interest. However, even for these spatial structures, the energetics of the flow is still of central interest.

The Solar Probe will certainly measure the proton bulk velocity to 4 $R_\odot$. That measurement alone provides significant information about the acceleration mechanisms (e.g., the scale length over which they operate). In addition, electron measurements will determine the thermal conduction energy flux, just as proton and magnetic field measurements combined will determine MHD wave

modes and fluxes in to 4 $R_\odot$. Overall, measurements in to 4 $R_\odot$ will help to define the complex interplay between acceleration, thermal conduction, wave dissipation, and wave "pressures" described above. Of course, by 4 $R_\odot$ the solar wind is already moving at respectable speeds, and the initial acceleration region inside 4 $R_\odot$ is inaccessible. Even so, observations at 4 $R_\odot$ may still be sufficient to understand in large measure what is happening within 4 $R_\odot$. This point is most easily illustrated in the context of MHD waves, as we discuss below.

Lower Limits on the Solar Alfven Wave Flux

MHD wave fluxes of solar origin are commonly observed at 1 AU. These observations can be combined with the theory of propagation and convection of undamped waves to estimate the MHD wave fluxes near the Sun and thus the percentage of the solar wind energy flux which derives from momentum transfer from a particular wave made, in this case Alfven waves. These estimates will be lower limits because they do not include the effects of wave damping. The procedure described below is based on 1 AU data, but it could be carried out at any point along the Solar Probe trajectory. The estimates improve with decreasing radial distance, because damping effects become less important.

A lower limit to the ratio, R, of the Alfven wave flux near the Sun (not at 1 AU) to the total energy flux needed to drive the coronal expansion (including the energy needed to lift the plasma out of the solar gravitational field) is

$$R = \frac{\delta V^2 + \delta B^2/4\pi\rho}{V\,V_A} \quad \frac{(1 + V_A/V)^2}{1 + V_E^2/V^2} \qquad (1)$$

where δV is the wave strength, V is the wind speed, V_E is the solar escape velicity, and V_A is the radial Alfven speed. This ratio is clearly the dynamically important one. It is valid for arbitrary wave amplitudes and for non-spherical expansion. All quantities in this expression, with the exception of the escape velocity from the solar surface, are to be evaluated at 1 AU, or any other point of observation.

Table 1 gives average values of some relevant quantities for five high-velocity streames in 1974, with the averages taken only in the interior of the stream well away from the compression regions at the leading edges.

Table 1

IMP-8 Plasma and Field Data for Five High Velocity Streams

Wind Speed	622 km/sec
Alfven Speed	65 km/sec
$(\delta V^2 + \delta B^2/4\pi\rho)^{\frac{1}{2}}$	57 km/sec
$\frac{\text{Wave Energy Density}}{\text{Magnetic energy density}}$	82%
$\frac{\text{Wave Flux at Sun}}{\text{Wind energy flux}}$	7%

(The magnetic field data were provided by R. Lepping of Goddard Space Flight Center.) Wave amplitudes are determined using six-hour intervals, with only the Alfvenic (correlated) fluctuation level shown. From Equation (1), we find that on the average direct acceleration by Alfven waves contributes at least 7% of the total energy flux needed to drive the coronal expansion in these streams. This is not a large fraction, but it is not negligible, especially since our estimate may be nonlinearly saturated. That is, Table 1 shows that the average energy density of Alfven waves at 1 AU is 82% of the background field energy density. This may be more than a coincidence, since it can be plausibly argued that the wave energy density cannot exceed the background field energy density due to the onset of nonlinear damping effects. If the waves are saturated, the Alfven wave flux at the Sun could be much larger than the 7% shown in Table 1. On the other hand, it could be that the Alfven waves have just reached nonlinearity at 1 AU with little damping beforehand, in which case our 7% estimate is realistic. From observations at or near 1 AU, we cannot differentiate between these two possibilities.

In contrast, the Solar Probe can easily provide a definitive measurement. If Alfven wave energy densities remain comparable to background field energy densities (i.e., the saturation case), the lower limit given in Table 1 for 1 AU will rise as 1/r (for $r > \sim 30\ R_\odot$). In the non-saturated, undamped case, the ratio will stay constant at 7%. Either of these alternatives is of considerable interest, in that it settles a question of basic energetics which is answerable in no other way. Early reports from Helios and Mariner 10 seem to favor the "saturation" alternative, although no systematic study is yet available. If this is the case, the "saturated" 7% at 1 AU would rise

to 23% at 0.3 AU. The Solar Probe, with perihelion at 0.02 AU, could settle the Alfven wave pressure question once and for all.

This example is presented as only one of many. The moral is that the closer one gets to the Sun, the easier it is to discover the important dynamical processes, both by direct observations and by theoretical extrapolations thereof.

Measurement Difficulties

There are, of course, difficulties in making complete electron and proton measurements on the Solar Probe, especially in the regions of greatest dynamical interest. Obvious difficulties include: (1) the limited coverage of phase space due to thermal constraints; (2) the transonic nature of the proton flow close to the Sun; (3) the short time spent near the Sun, especially compared to the periods of low frequency MHD waves; (4) the high velocity of the spacecraft; (5) possible high turbulent fluid velocities near the Sun.

Let us consider points (4) and (5) in some detail. One way to determine the "turbulent" velocities is to calculate Alfven wave amplitudes inside 1 AU using the parameters given in Table 1 as a boundary condition at 1 AU and assuming no damping. This will again be a lower limit, because damping and other wave modes are neglected. However, the calculation is illustrative and gives an idea of what to expect.

To obtain wave amplitudes close to the Sun, we assume that the Alfven velocity and the solar wind velocity vary with radial distance as shown in Figure 1. This is the spherical model of Sittler (1978), with the velocity profiles determined from an empirical density profile and mass conservation. The velocity is 530 km/sec at the Alfven point (20 $R_\odot$) and 700 km/sec at 1 AU. The Alfven velocity is about 1500 km/sec near 4 $R_\odot$, which is somewhat high, but not inappropriate for conditions above coronal holes.

Figure 2 gives the amplitudes of transverse waves in the solar wind for wave periods from zero to 1.9 days, assuming a 40 km/sec amplitude at 1 AU. The curves are not limited to short wavelengths, but are solutions to the full MHD equations for small amplitude toroidal waves of finite wavelength (Heinemann and Olbert, 1978). The zero period curve is, of course, the WKB limit. Points of interest are that: (1) inside 1 AU, the WKB limit yields the largest amplitudes; (2) δV rises as r decreases until the Alfven

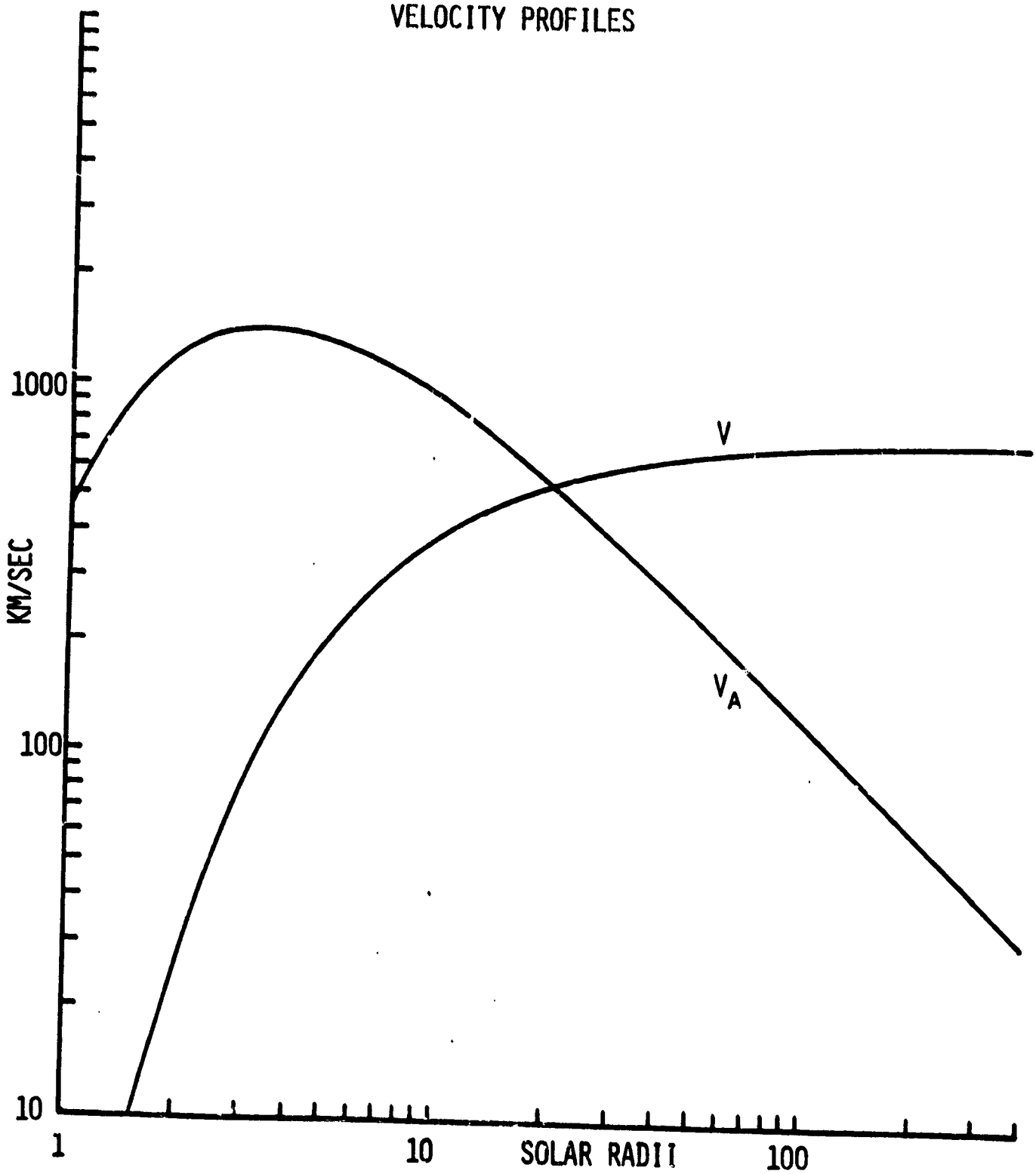


Figure 1: The variation of the solar wind and Alfven velocities as a function of distance from the Sun, from the spherical model of Sittler (1978).

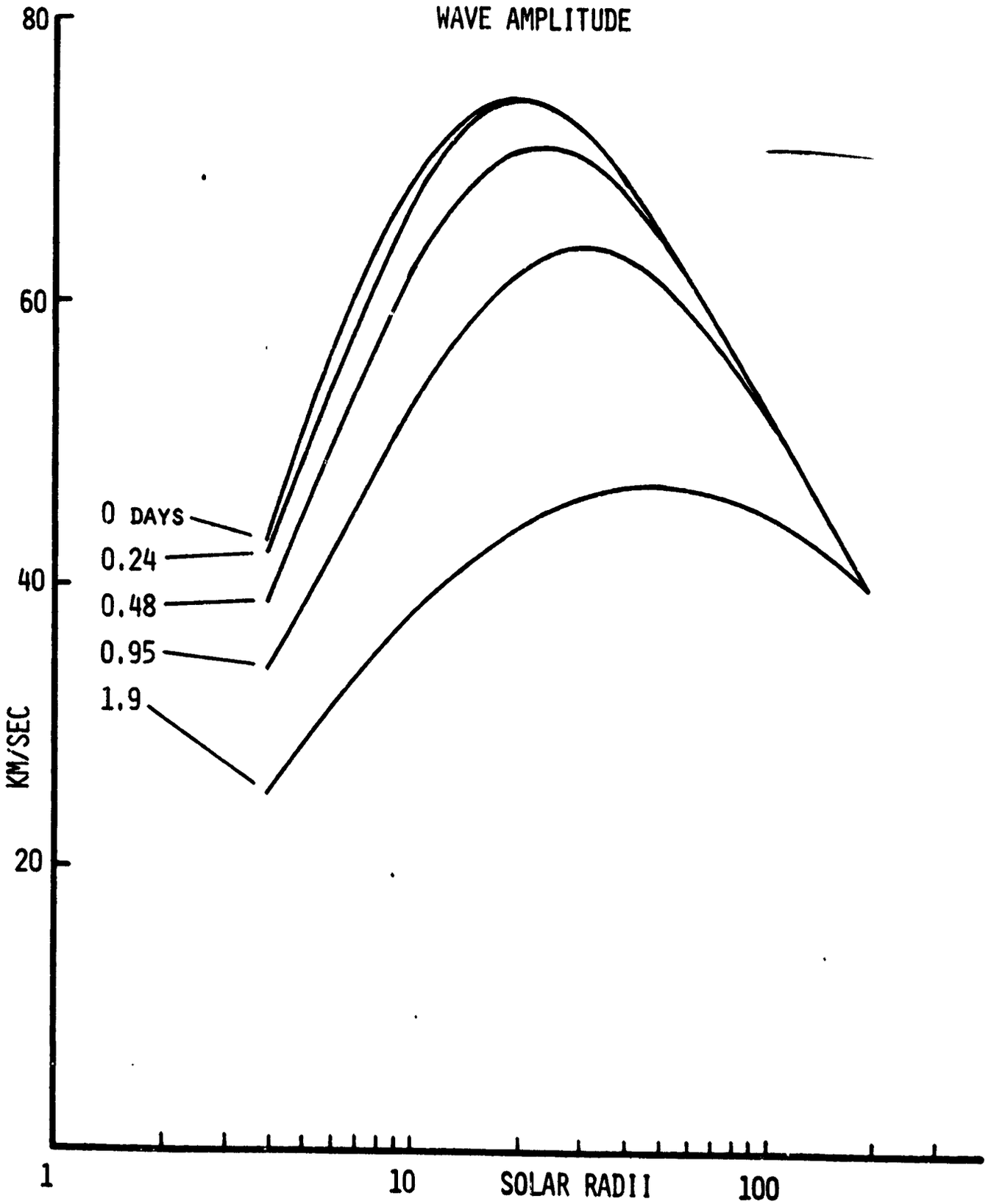


Figure 2: The predicted amplitudes of transverse waves in the solar wind for wave periods from zero to 1.9 days as a function of radius, assuming a 40 km/sec amplitude at 1 AU.

point is reached, after which δV decreases as we move closer to the Sun; (3) these curves are predictions of root mean square velocities--peak to peak values are larger.

Again, Figure 2 is only a lower limit for Alfven waves; other wave modes may totally alter this picture. If we consider only Alfven waves, an absolute upper limit on velocity amplitudes would be a factor of $\simeq 3$ higher than the curves shown; at higher velocities, Alfven wave fluxes at the Sun would exceed the total flux necessary to drive the expansion, which is an impossible situation.

It is also instructive to add the spacecraft velocity along the Solar Probe trajectory to the velocity profile shown in Figure 2 and consider the resulting aberration of the plasma flow. In Figure 3, we show various curves in velocity space. The vertical axis is the radial velocity (positive away from the Sun) and the horizontal axis is velocity transverse to the radial. The circle at the bottom is the probe trajectory in velocity space, with time in days and radial distance in solar radii, as indicated by tick marks. The curve labeled "700 km/sec stream" is the locus in velocity space of the radial velocity profile in Figure 1, as it would appear in the rest frame of the spacecraft. That is, the purely radial velocity profile in Figure 1 will appear to have a large transverse velocity in the spacecraft frame, as well as a varying radial velocity. The curve labeled "350 km/sec stream" is the velocity profile in Figure 1 divided by two and treated in the same manner. Note that after perihelion the probe overtakes the "350 km/sec" wind as it moves away from the Sun. The aberration of the "700 km/sec stream" is shown in Figure 4, where we plot the deviation of the apparent solar wind flow from radial. Clearly, large aberrations of the bulk flow will be the rule close to the Sun. Although this will present a challenging opportunity for experimentalists, especially on a non-spinning spacecraft, the large aberration means the instrument does not have to look directly at the Sun and can thus remain shaded near perihelion. Of course, there are also obvious problems in sampling the thermal spread of the flow, which for the transonic protons will resemble the traditional problems of electron measurements.

In Figures 3 and 4 we indicate the probable spread due to the "turbulent" velocities of Figure 1 (horizontal bars in Figure 3 and vertical bars in Figure 4). The aberration problems increase in direct proportion to the turbulent velocities.

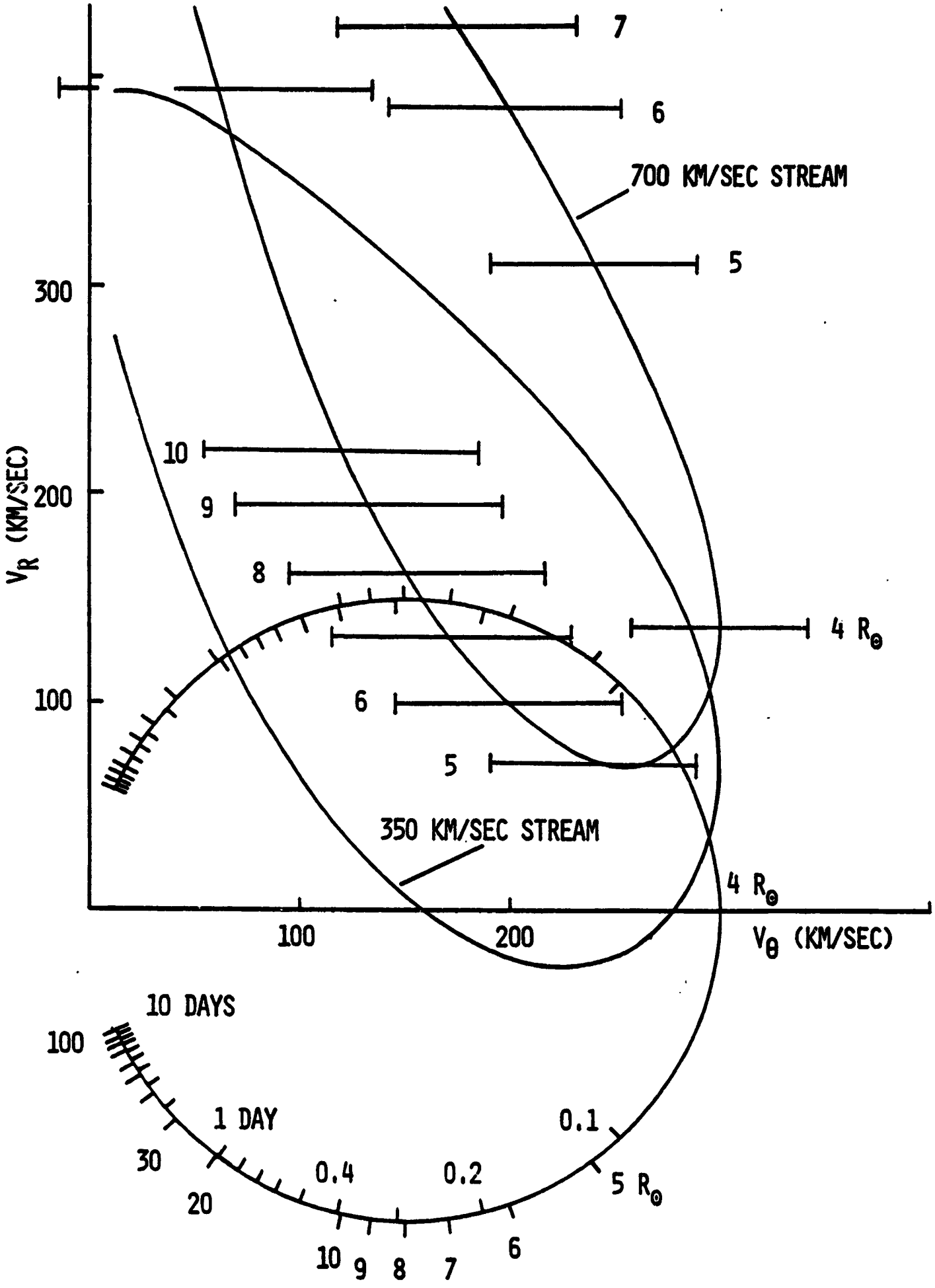


Figure 3. Probe trajectory and apparent solar wind flow in velocity space for a radial solar wind speed equal to or half that shown in Figure 1. See text for details.

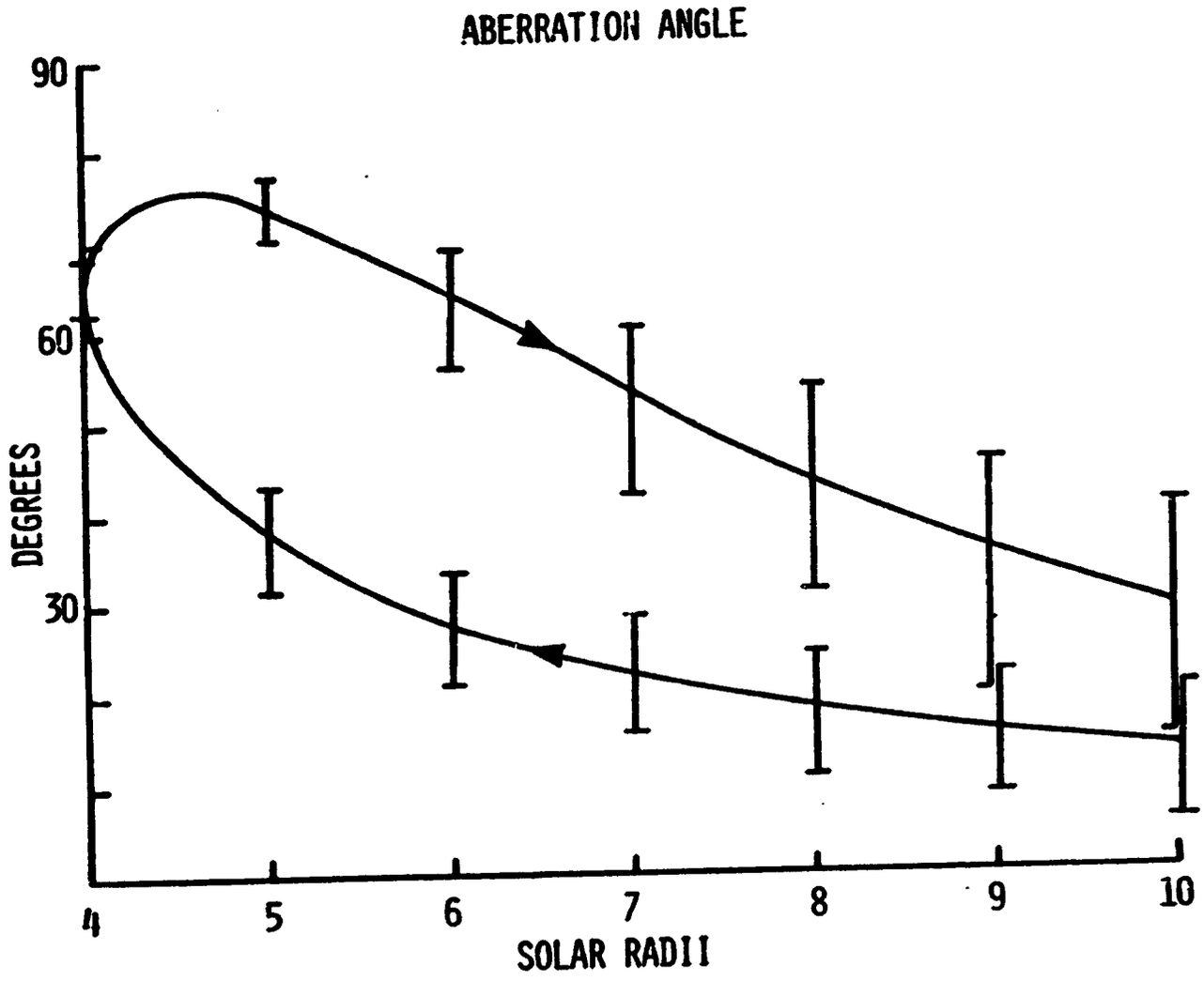


Figure 4: The aberration angle of the "700 km/sec stream", as a function of radius. The aberration angle is the angle between the radial direction and the apparent direction of flow as seen in the spacecraft frame.

Kinetic Effects of Alfven Wave Pressures

To end on a physically interesting note, we consider again the concept of Alfven wave pressure. It is well known that such pressure can accelerate the solar wind. With the exception of Hollweg (1978), however, all treatments of this effect have been fluid treatments, primarily because a proper kinetic (i. e. Vlasov) treatment is very complex. Recently, Goodrich (1978) has formulated a general quasilinear theory for an inhomgeneous plasma and obtained numerical solutions for the case of Alfven waves in the solar wind. Goodrich's numerical results for a radial magnetic field are shown in Figures 5a through 5c. The upper panel shows velocity space contours of the proton distribution function in the case of no Alfven waves. The vertical axis is velocity perpendicular to the radial (field) direction and the horizontal axis is velocity parallel to the radial direction. The lower panel shows contours of the proton distribution function if Alfven waves are present and accelerating the flow. The distribution function without waves is assumed isotropic at 40 $R_\odot$, and the evolution is shown at 60, 80, and 100 $R_\odot$ in Figures 5a, b, and c, respectively.

In the progression from 60 to 100 $R_\odot$, one can easily see the wave acceleration of the bulk of the distribution to higher velocities as compared to the no-wave case. The details of how that acceleration is actually accomplished are quite interesting. Those thermal particles on axis ($v_\perp = 0$) moving at a speed equal to the wind speed plus the Alfven speed feel no wave pressure. Particles on axis moving at less than this speed feel an effect which increases as their velocity separation increases. Thus the distributions in Figure 5 look as if they have been pushed along the axis from behind. The exact details of the effects for particles off axis ($v_\perp \neq 0$) are complicated, but the general tendency is clear. It is not clear whether such distributions will actually be observed in the solar wind. Most obviously, they may be unstable to the generation of other wave modes. However, the neighborhood of the Alfven point, where the wave pressure gradient is strongest, could be a good region to observe such kinetic effects.

This is only one of many physical processes which may best be studied at and inside of the Alfven critical point. Solar angular momentum loss is another. Our final point is that the Solar Probe offers an opportunity to study <u>in situ</u> the evolution of an astrophysical plasma flow from highly

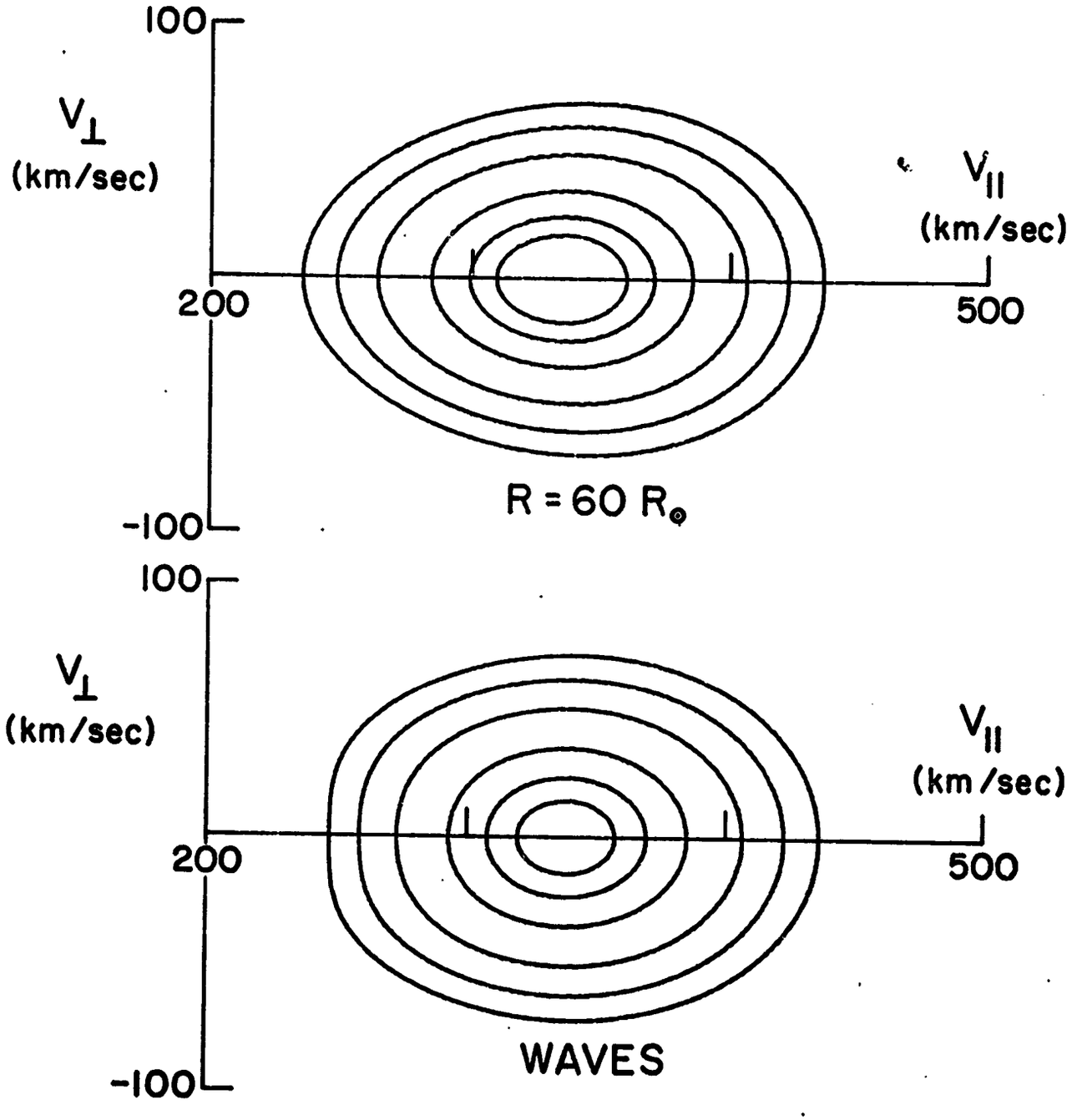


Figure 5a: Velocity space contours of the proton distribution function in the absence of Alfven wave acceleration (upper panel) and in the presence of Alfven wave acceleration (lower panel). The proton distribution is shown at 60 $R_\odot$, assuming isotropy in the no-wave case at 40 $R_\odot$.

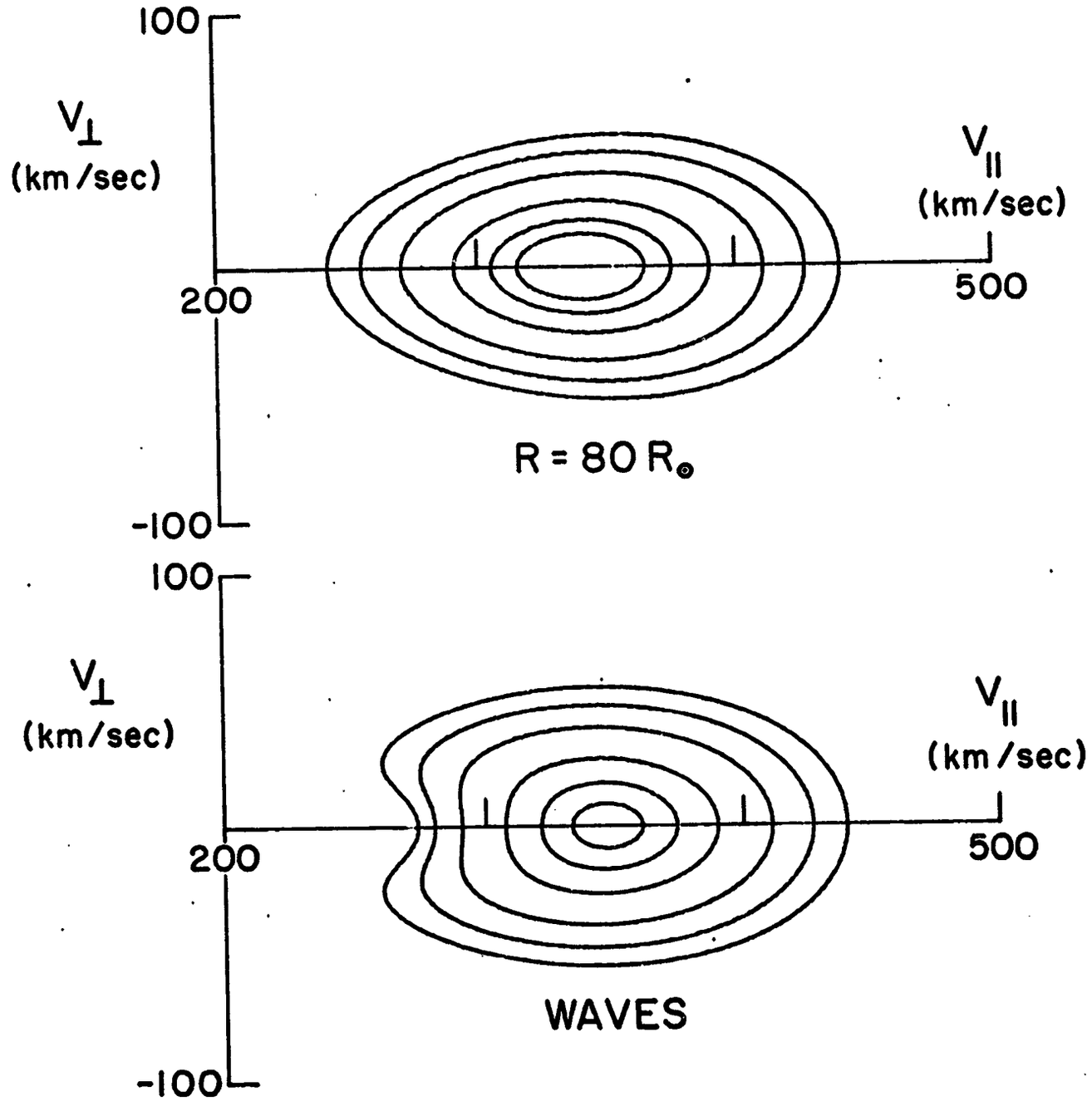


Figure 5b: Velocity space contours of the proton distribution function in the absence of Alfven wave acceleration (upper panel) and in the presence of Alfven wave acceleration (lower panel). The proton distribution is shown at 80 $R_\odot$, assuming isotropy in the no-wave case at 40 $R_\odot$.

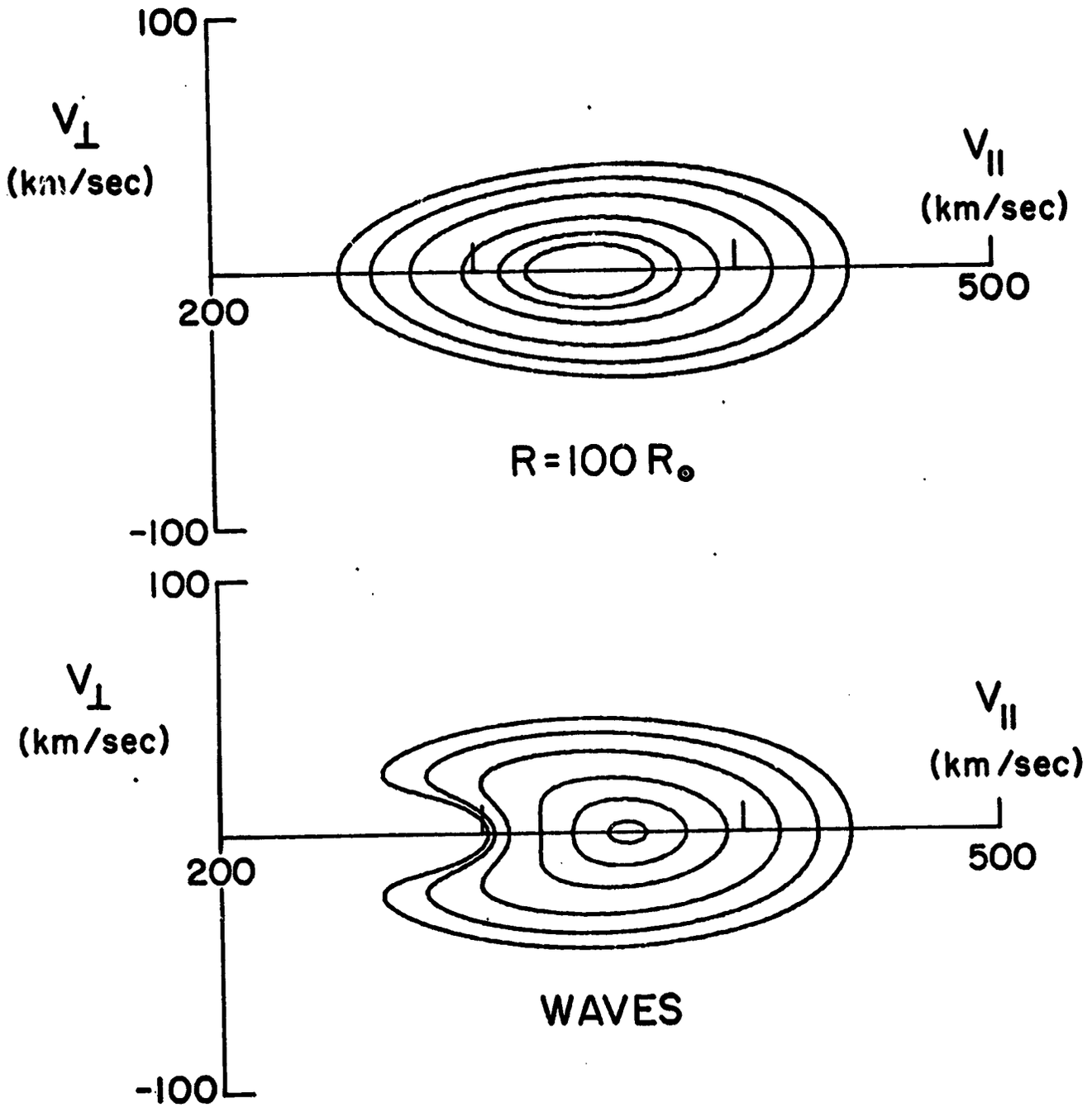


Figure 5c: Velocity space contours of the proton distribution function in the absence of Alfven wave acceleration (upper panel) and in the presence of Alfven wave acceleration (lower panel). The proton distribution is shown at 100 $R_\odot$, assuming isotropy in the no-wave case at 40 $R_\odot$.

sub-Alfvenic to highly super-Alfvenic velocities. This opportunity is unique, and should provide a wealth of astrophysically interesting observations which can be obtained in no other way.

References

Goodrich, C., "Wave Particle Interactions and the Dynamics of the Solar Wind", Ph.D. Thesis, M.I.T. Department of Physics and Center for Space Research, 1978.

Heinemann, M., and S. Olbert, "Finite Frequency Transverse Waves in the Solar Wind", EOS, 59, 367, 1978.

Hollweg, J., "A Quasi-Linear WKB Kinetic Theory for Nonplanar Waves in a Nonhomogeneous Warm Plasma,1, Transverse Waves Propagating Along Axisymmetric $\underset{\sim}{B}_o$", JGR, 83, 563,1978.

Sittler, E., "Study of the Electron Component of the Solar Wind and Magnetospheric Plasma", Ph.D. Thesis, M.I.T. Department of Physics and Center for Space Research, 1978.

D27 N78-32991

ACCELERATION AND HEATING OF THE SOLAR WIND

Aaron Barnes

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

Abstract

The study of the solar wind bears an importance beyond being an interesting topic of space science. The solar wind is the prototype of the escape of matter from a large class of stars, and perhaps from star clusters and galaxies. Moreover, an impressive array of evidence indicates that large variations in the solar wind-solar activity complex can lead to appreciable changes in the Earth's climate. In this light it is especially frustrating that we still do not understand which processes drive the solar wind, i.e., how it is accelerated and heated. This state of affairs arises primarily because we have observations of only two relevant regions: the inner corona and the asymptotic coasting region of the wind (mostly observed at 1 AU, never nearer the Sun than 0.3 AU). Hence theoretical models are constrained by data at essentially only two points; it is not surprising that models with quite different assumptions about the acceleration and heating mechanisms are equally consistent with observation.

We review briefly some of the competing theories of solar wind acceleration and heating, and discuss what observations are required to distinguish among them. In most cases what is required is measurement of plasma velocity and temperature and magnetic field, as near the Sun as possible and certainly inside 20 $R_\odot$; another critical aspect of this question is determining whether a turbulent "envelope" exists in this inner region, and if so, defining its properties. Plasma and magnetic observations from the proposed Solar Probe mission would thus yield a quantum jump in our understanding of the dynamics of the solar wind.

I am very pleased that there is a consensus at this meeting on the importance of the problem of solar wind acceleration and heating, and that there seems to be a widespread feeling that the solar probe has the potential of bringing about major progress in our understanding of this question.

Since a number of other speakers are addressing similar issues, I will not attempt a thorough review covering every facet of this problem, but will limit my comments to a few topics that I regard as especially interesting and important. An appropriate beginning is to ask ourselves why so many scientists attach major importance to the problem of solar wind acceleration and heating. It is of obvious interest to people who, like some of you in the audience and like me, have spent much of their scientific careers studying the solar wind. But this question is of greater scientific significance than just that. The solar wind is the paradigm for the escape of matter from a large class of stars, and probably other objects as well. I think that we have little hope of understanding the general problem of astrophysical mass loss if we don't know what causes the expansion of the solar wind. It now seems quite likely that the complex of the solar wind and solar activity is somehow related to variations in Earth's weather and climate. We think that solar activity was depressed during the Maunder minimum. What was the solar wind doing at that time? I have no idea, and a fair part of the reason why I have no idea is that we do not yet understand very well how the solar wind is accelerated.

The dominant mechanism of acceleration and heating has remained elusive because we have observations of only two relevant regions: the inner corona,

and the asymptotic coasting region of the wind (not nearer the Sun than 60 $R_\odot$). Theoretical models are therefore constrained by data at essentially only two points; it is not surprising that models with quite different assumptions about the acceleration mechanisms are equally consistent with observation. This problem is illustrated in Fig. 1, which gives profiles of flow velocity vs. heliocentric distance as predicted by a few sample models. I have used these particular models because I happen to have worked with them, and they are therefore convenient, but the same point could be made with any set of models. All models give velocity profiles that are rather flat between about 20 $R_\odot$ and 1 AU, and in fact the typical day-to-day variations in speed observed at 1 AU are larger than the predicted variation over the range 0.1-1 AU in any given model. Hence it is clear that velocity measurements taken outside of 15-20 $R_\odot$ are of little value in trying to distinguish among the theoretical alternatives.

In contrast, the solar probe near perihelion (presumably 4 $R_\odot$) could yield velocity data of great consequence. Clearly a single pass will not give us simple radial profiles like those shown in Fig. 1; instead it will give a profile along which distance, latitude, longitude and time are convolved. We may anticipate large variations with latitude and longitude, possibly including regions of static corona. Velocity information by itself is therefore not likely to yield a definitive picture of solar wind acceleration. It may, however, restrict the alternatives in an important way. Radial-flow models of the solar wind (e.g., Fig. 1) normally predict fairly low velocities at 4 $R_\odot$, of order 50-150 km/sec. On the other hand, Kopp and Holzer (1976) have raised the possibility that, if the magnetic field in solar-wind source regions diverges rapidly enough, the sonic point may lie very near the coronal base. In this case the flow speed at 4 $R_\odot$ would be rapid ($\geq$ 300 km/sec). We should be able to get a feeling for which of these two extremes is more typical of the Sun at the time of probe perihelion.

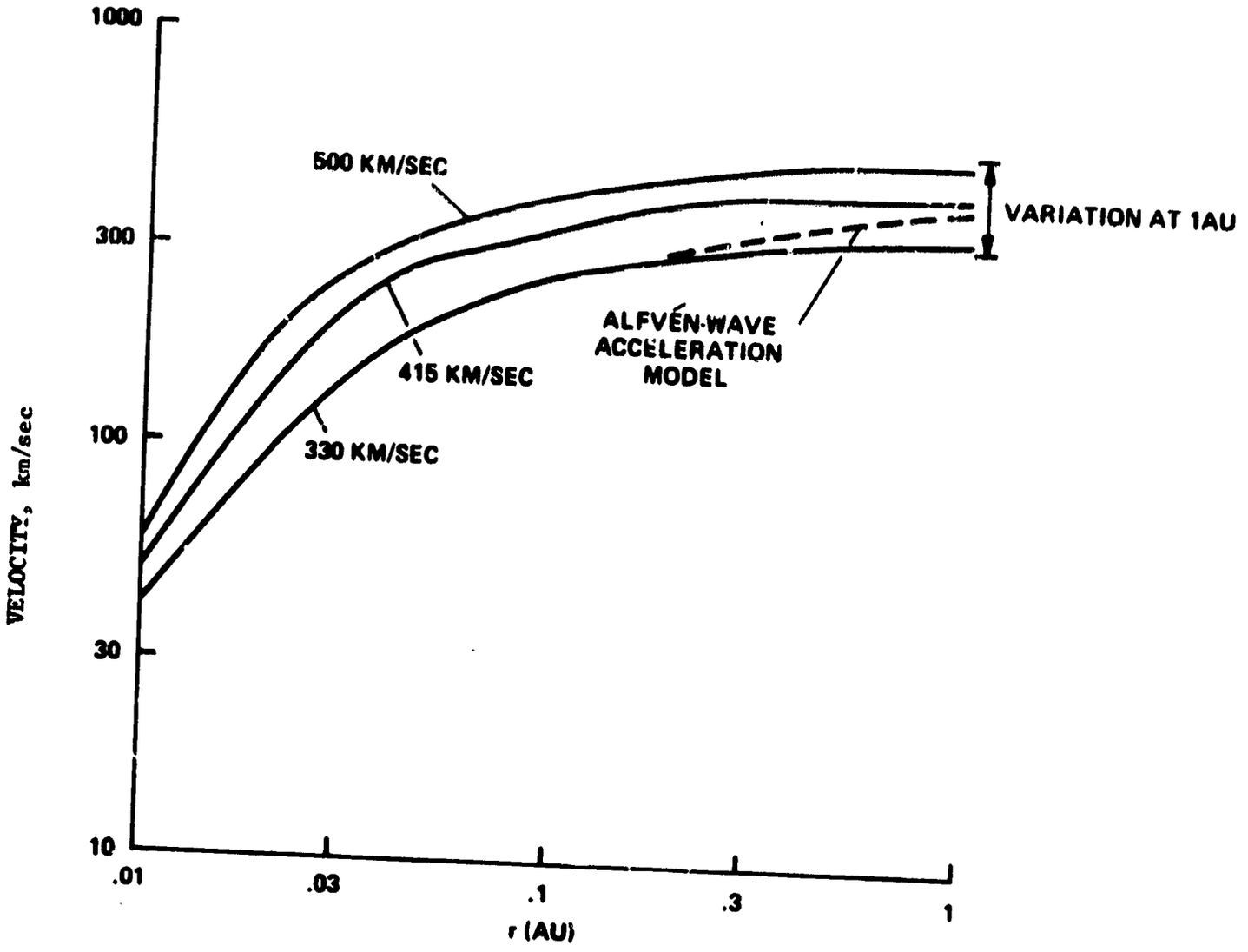


Figure 1: Sample velocity profiles illustrating the slow variation of v with r beyond 20 $R_\odot$. Solid curves are magnetoacoustic-wave driven models calculated by Hartle, Barnes and Bredekamp (unpublished), and a dashed curve corresponding to an Alfven-wave driven model of Hollweg (1973) is shown for comparison. All other published models show similarly flat velocity profiles.

There are a number of reasons to expect that the region inside 20 $R_\odot$ may be turbulent. This turbulence is an important topic by itself, which I shall discuss shortly, but for the present I wish to consider it merely as a potential problem for measuring the velocity profile. Fig. 2 shows the calculated profiles of velocity amplitude according to several different theories. In addition, the measure of randomness in velocity (conceivably due to turbulence) inferred by Ekers and Little (1971) from radio scintillation data has been indicated. Altogether, it would not be surprising to find a fluctuating component of the flow velocity of order 100-200 km/sec near probe perihelion. It would obviously be difficult to determine from a single spacecraft whether observed velocity variations represent true wave motions or were due to the sampling of a number of steady streams of different velocity. A second spacecraft close by would resolve much of this ambiguity.

The parameter whose spatial profile is most sensitively dependent on the mechanism of solar wind heating is the proton temperature. Fig. 3 shows the proton temperature profiles for three sample models that give about the same asymptotic flow speed for the wind. Two of the temperature profiles are rather alike, but a third, based on nonlinear dissipation of Alfven waves, is quite different because the dissipation occurs further from the Sun than in the other cases. The most striking differences in the profiles occur at intermediate distances, of order 20 $R_\odot$. Unfortunately, there are great difficulties in determining temperature profiles from spacecraft observations because of the large day-to-day variations in the flow. The typical 1 AU temperature variation is comparable to the maximum difference of the models at any distance.

Just as in the case of velocity measurements, temperature measurements from a single solar probe pass will not yield a definitive picture of how the solar wind is heated. Once again, however, the alternatives can be restricted. The extreme cases that can be imagined are: (a) extended heating out to 10 $R_\odot$

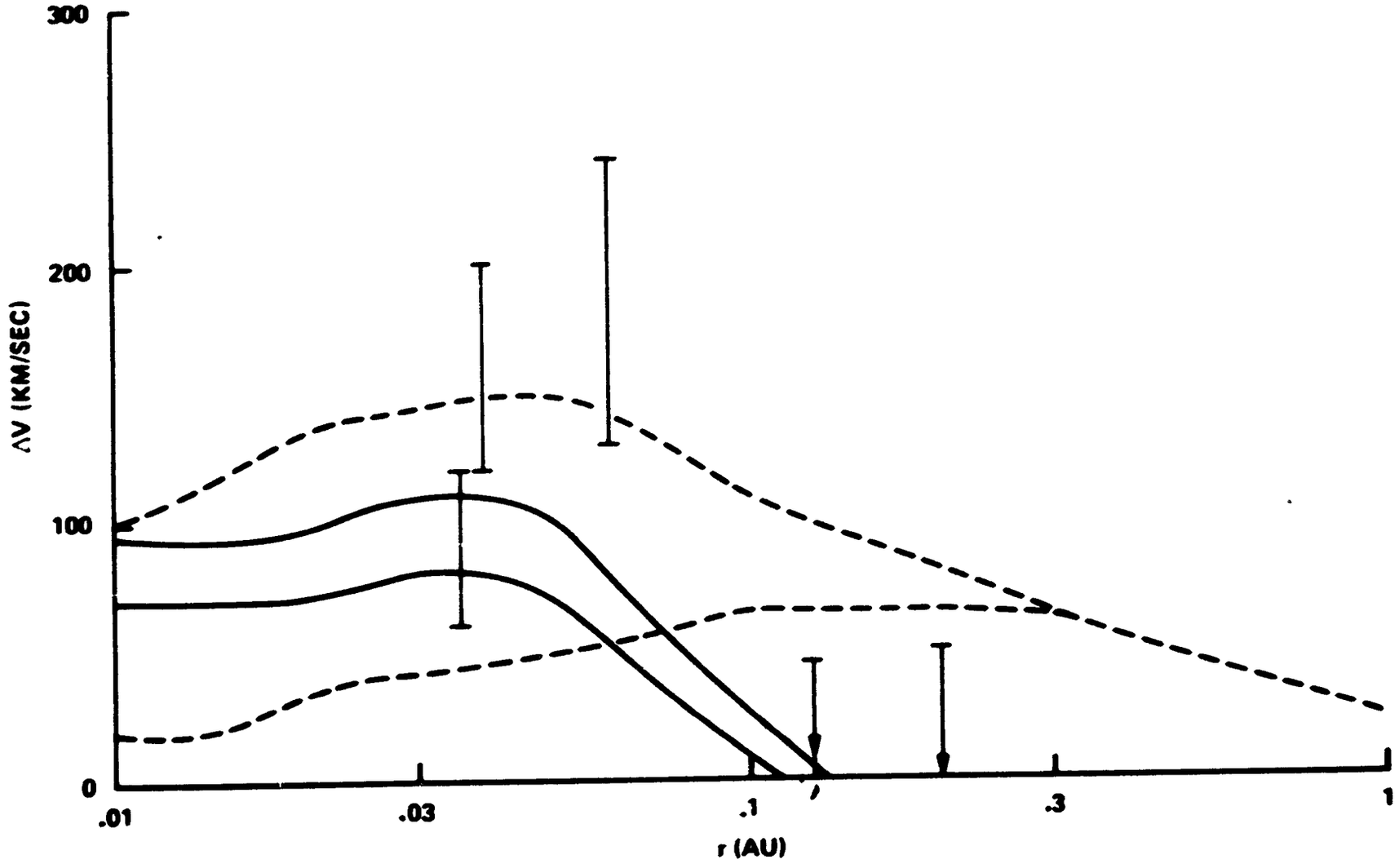


Figure 2: Theoretical profiles of the velocity amplitude of wave fields according to several models. Vertical bars represent randomness (possibly due to turbulence) in velocity inferred by Ekers and Little (1971) from radio scintillations. Solid curves represent magnetoacoustic turbulence as calculated by Hartle, Barnes, and Bredekamp (unpublished). Dashed curves represent Alfvenic turbulence calculated in the WKB approximation with and without nonlinear saturation (saturated profile after Hollweg, 1973).

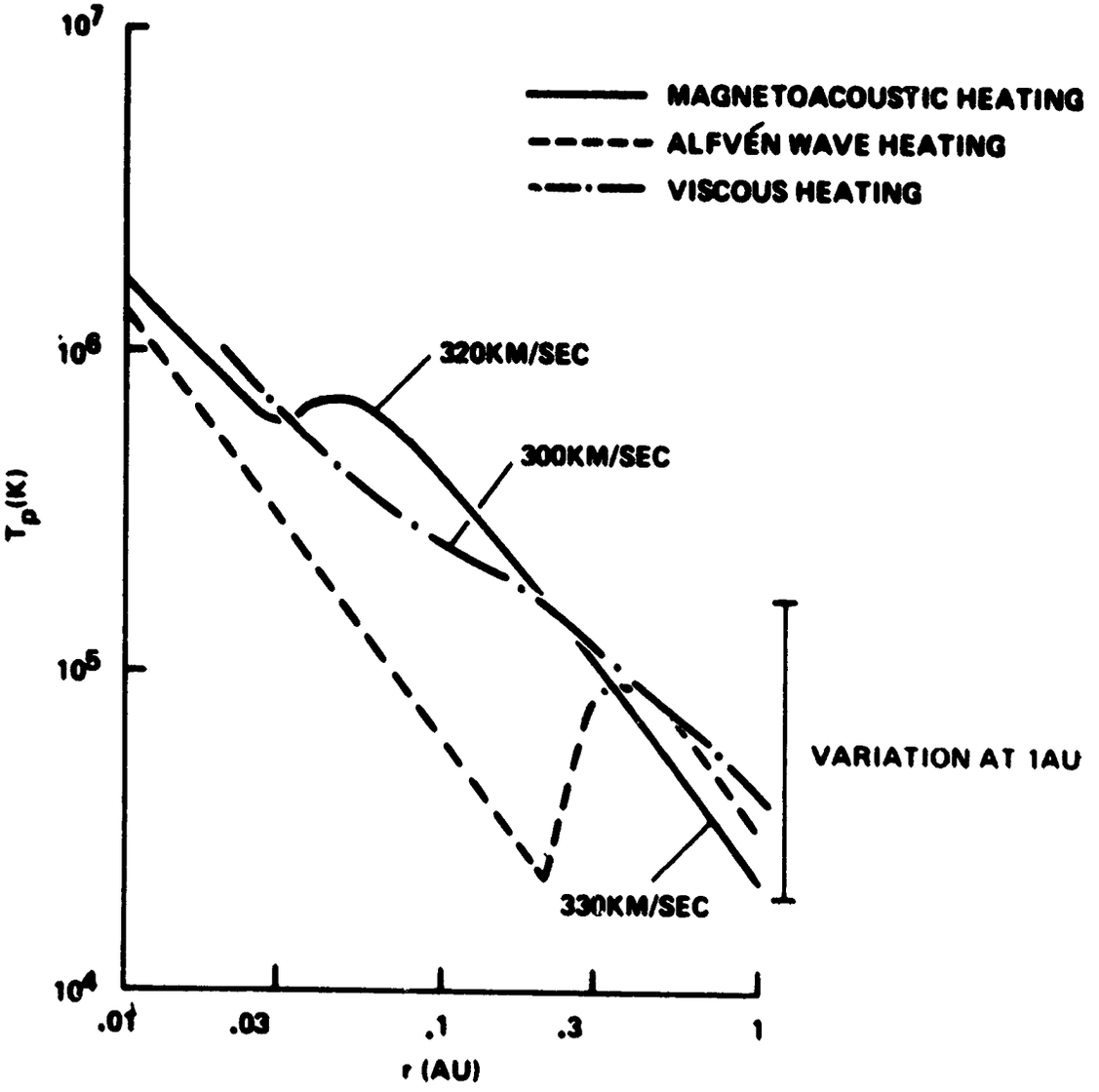


Figure 3: Sample proton temperature profiles for three different models that give nearly the same asymptotic flow speed. The solid curve is by Hartle, Barnes, and Bredekamp (unpublished), dashed curve is after Hollweg (1973), dash-dot curve is after Wolff, Brandt and Southwick (1971).

or beyond; and (b) essentially adiabatic expansion beyond a coronal base near the Sun. If one of these extremes obtains, it should be apparent in solar probe temperature data. As in the case of velocity measurements, a second spacecraft would be of enormous value in reducing ambiguities in the profiles due to turbulent fluctuations.

Thermal conduction is an essential part of solar wind energetics, and probably has its largest effect in the inner core of the wind. Therefore, measurements of electron temperature and heat flux are of key importance in the study of solar wind heating. Since this topic has already been discussed in detail by Feldman and Belcher (see papers, this volume), I will not comment beyond stressing their importance.

Plasma and magnetic measurements, and their correlation with solar observations, will extend our understanding of acceleration and heating of the solar wind. In particular, they will provide an opportunity to confirm or deny some of our ideas about extended heating and acceleration by hydromagnetic turbulence. Fig. 4 gives calculated profiles of the magnetic amplitude of turbulent envelopes based on varying assumptions. Two extreme models are that the turbulence is entirely Alfvenic or entirely magnetoacoustic. Alfven waves of the intensity observed near 1 AU are of low intensity when extrapolated inward by the simple WKB theory. In contrast, if the observed Alfven waves at 1 AU are the residuum of a turbulence field that has been strongly damped on passage from the Sun, Alfvenic turbulence could be quite strong in the inner core of the wind and could account for a major part of its acceleration. Magnetoacoustic turbulence, on the other hand, is strongly dissipative and could exist at significant levels only inside about 20 $R_\odot$. These waves would heat the wind strongly in this region, thereby indirectly accelerating the wind.

It should be fairly easy to distinguish between the extremes in the solar probe mission. The spatial variation of the two kinds of turbulence is

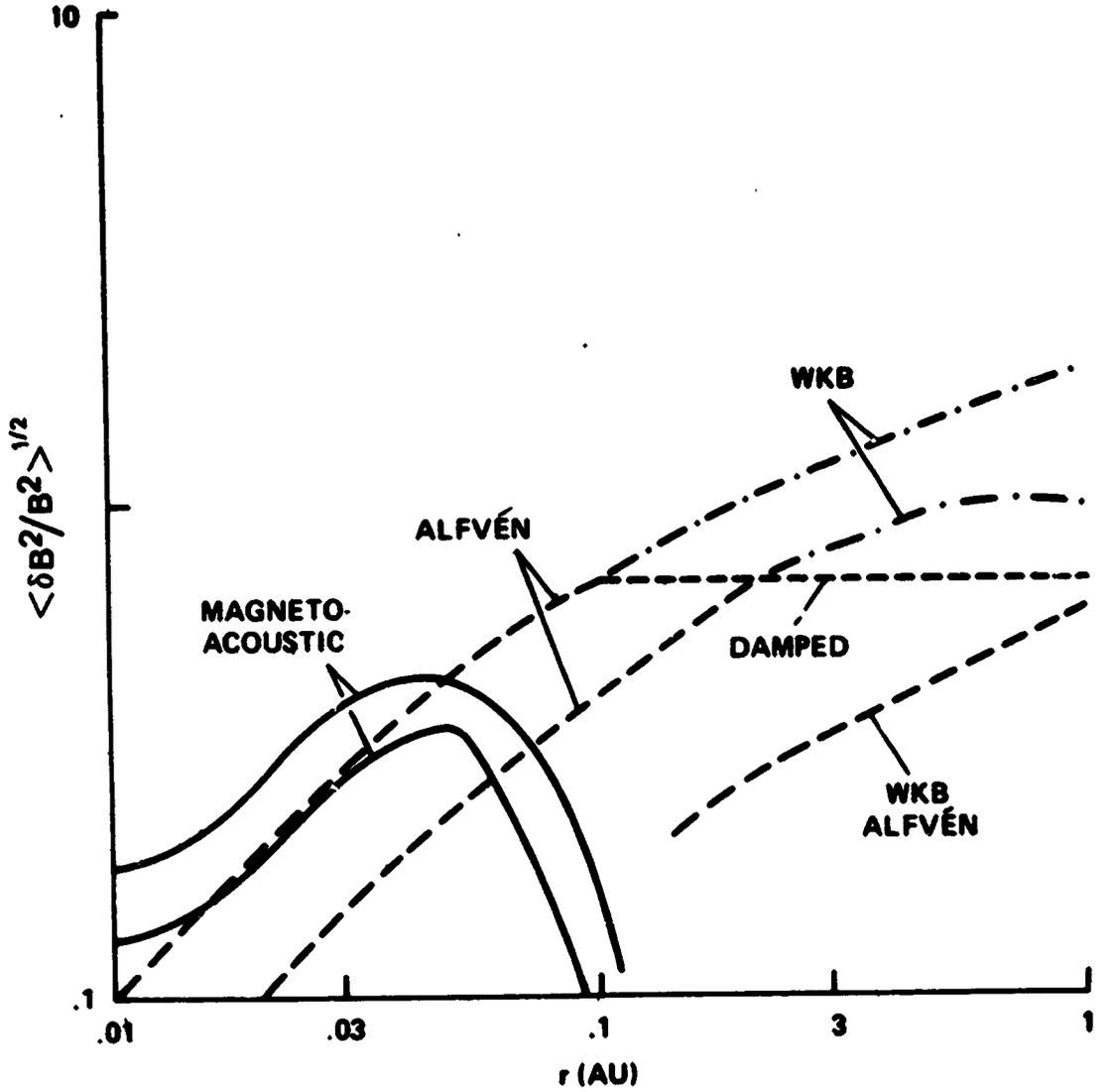


Figure 4: Theoretical profiles of the magnetic amplitude of wave fields according to several models. Solid curves represent magnetoacoustic turbulence as calculated by Hartle, Barnes and Bredekamp (unpublished). Dashed curves represent Alfvenic turbulence calculated in the WKB approximation with and without nonlinear saturation (saturated profile after Hollweg, 1973).

qualititatively different, as is clear from Fig. 4. The local properties of the two kinds of turbulence are also different. Alfven waves appear as fluctuations in the magnetic field direction, with little or no compression in the field strength or plasma density. On the other hand, magnetoacoustic waves are linearly polarized and compressive in both magnetic field strength and density. Plasma and magnetic data from the probe should tell us whether there is a turbulent envelope, and whether it consists mainly of Alfven or magnetoacoustic waves, a mixture of both, or something else entirely. These observations could definitively establish whether turbulent heating and acceleration are an important part of solar wind dynamics.

These comments on the ways in which the solar probe could contribute to our understanding of heating and acceleration are far from exhaustive. A number of other processes have been discussed in the literature, but the pressure of time has forced me to focus on a few of my favorites. In any case, perhaps the most exciting thing about the proposed mission is that it will put us in a position to look back at all the theorizing that has been done. At that point, all that we have done will seem "to have been either wrong, or irrelevant or obvious", as Herbig (1977) said of another field. While we may still not understand all the details of solar wind acceleration after the first probe, we will know an order of magnitude more than we do now.

Note on bibliography: In such a brief paper it is not possible to acknowledge the contributions of many scientists to the development of the field. Extensive references may be found in the review articles of Feldman (1978), Hollweg (1975, 1978), Holzer (1978) and Barnes (1975, 1978), including references to previous reviews.

References

Barnes, A., Plasma processes in the expansion of the solar wind and in the interplanetary medium, Rev. Geophys. Space Phys., 13 (1975), 1049.

Barnes, A., Hydromagnetic waves and turbulence in the solar wind, in Solar System Plasma Physics (C.F. Kennel, L.J. Lanzerotti and E.N. Parker, eds.), North-Holland, 1978.

Ekers, R.D., and L.F. Little, The motion of the solar wind close to the Sun, Astron. Astrophys., 10 (1971), 310.

Feldman, W.C., Kinetic processes in the solar wind, in Solar System Plasma Physics (C.F. Kennel, L.J. Lanzerotti and E.N. Parker, eds.), North-Holland, 1978

Herbig, G.H., Summary of the conference: observations, in Star Formation (T. de Jong and A. Maeder, eds.), IAU Symp. 75, Reidel, 1977.

Hollweg, J.V., Alfven waves in a two-fluid model of the solar wind, Astrophys. J., 181 (1973), 547.

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

Hollweg, J.V., Physical processes in the solar wind, submitted for publication, 1978.

Holzer, T.E., The solar wind and related astrophysical phenomena, in Solar System Plasma Physics (C.F. Kennel, L.J. Lanzerotti, and E.N. Parker, eds.), North-Holland, 1978.

Kopp, R.A., and T.E. Holzer, Dynamics of coronal hole regions, Solar Phys., 49 (1976), 43.

Wolff, C.L., J.C. Brandt, and R.G. Southwick, A two-component model of the quiet solar wind with viscosity, magnetic field, and reduced heat conduction, Astrophys. J., 165 (1971), 181.

D28

A SOLAR WIND EXPERIMENT FOR THE SOLAR PROBE MISSION

K. W. Ogilvie
NASA Goddard Space Flight Center
Greenbelt, Maryland, 20771

ABSTRACT

Electron and ion plasma measurements on Solar Probe are important, especially to elucidate the role of coronal holes as sources of the solar wind and to resolve the long debated question of the importance of wave heating within 20 $R_\odot$ to the acceleration of the solar wind. Adaptable instruments are required to cover the large dynamic range of parameters between 1 AU and 4 $R_\odot$. The baseline mission would be improved, in some respects, by the addition of a second spacecraft, following the first by a distance of several solar radii. A spinning spacecraft would be satisfactory, but not essential, for the present objectives. Attention should be paid to having the heat shield electrically conducting. An instrument for ion and electron (but no composition) measurements would have a mass of ∿10 kg, require 8 w of power, and generate data at a maximum rate of ∿2 kbits/sec.

A SOLAR WIND EXPERIMENT FOR THE SOLAR PROBE MISSION

Scientific Aims

The discovery of coronal holes has focussed attention on the fact that very little is known about the source and acceleration of the solar wind. This will naturally form one of the principal subjects of research with Solar Probe. An important question which can be solved in principle is the importance of waves and the presence of the extended heating region modeled by Hartle and Barnes (1970), thought to exist at 12 $R_\odot$. These effects will make themselves felt principally in the radial variation of bulk speed and ion temperatures, but the hydromagnetic waves have a large velocity amplitude ($\frac{dV}{V} = \frac{dB}{B}$), and may be detectable directly as a local increase in velocity fluctuation. There may also be an effect upon the electron temperature and temperature anisotropy, although this is not certain at present. It has also been suggested that heating in an extended region might be caused by the transfer of energy from electrons to ions at $\sim$ 12 $R_\odot$. The relative importance of these mechanism could be evaluated if it is found that such heating plays a part in the acceleration of the solar wind to high speeds. The relationship of the solar wind to the underlying solar surface structure can also be studied. For example, how much of the solar wind comes from coronal holes, and what is the nature of the source regions of streams? Flow may be far from radial close to the sun, this direction being presumably along the local magnetic field. The co-rotation of the plasma is also interesting, since measurements at 1 AU of the angular momentum loss of the sun are not in a satisfactory state. All these measurements require observations made inside both the critical radius ($\sim$ 10 $R_\odot$) and the Alfvén radius (20 $R_\odot$), and the Solar Probe is ideal for these studies.

Trajectory Considerations

The suggested choice (closest approach 4 $R_\odot$; inclination 90°, nodal angle 45°) seems satisfactory for the plasma measurements, since it is likely to pass across at least extensions of the polar coronal holes.

In 1989 the solar activity will be rising, so that there may be some coronal hole structures near the equator but the dominant holes will probably be the polar holes. Thus a high inclination will almost certainly pass across parts of the surface occupied by holes, and parts not occupied by holes.

There may be difficulties in making solar wind measurements this close to the sun ($T_e \approx T_i = 10^6$ K, U > 100 km sec^{-1}, $n \sim 5 \times 10^4$ cm^{-3}); see Fig. 1. For example, $\lambda_D \approx 30$ cm, is smaller than the spacecraft, which means that there will be wake effects downstream of the craft. The Alfven speed (Fig. 2) is high enough there that shocks are unlikely to form about the spacecraft. The plasma parameters obtained directly will thus be characteristic of the medium, but the spacecraft will be surrounded by an asymmetric sheath even if it is an equipotential. This is likely to trouble the plasma wave experiments.

It would be advantageous to have two spacecraft either separated by a small distance along the orbit, or at different inclinations. In the former case time dependent effects can be studied, while in the latter case there is a greater chance of overflying an equatorial coranal hole. Perhaps one craft could act as a telemetry relay for the other.

A spinning spacecraft has some advantages in terms of solid angle coverage but these in general do not, in the present author's opinion, outweigh the weight penalty.

Ion Measurements

If the aims of the mission are to be met, there are two classes of ion measurements:

1) Accurate proton distribution function measurements.

2) Composition measurements.

1. The distribution function needs to be measured so that heating and acceleration models can be checked. Since the aberration angle is large, it is possible to observe the ions by looking past the heat shield with a wide angle detector (Faraday Cup). It seems unlikely that a detector which looked over the edge of the spacecraft could survive 2500 times the sun's radiation at 1 AU, and the corresponding photo-electrons would cause problems. An external deflection electrode system would probably impact the design of the spacecraft and also perturb the ambient plasma and make it difficult to make measurements of low energy electrons and plasma waves.

2. The principal advantage of this mission for composition studies is that even the rare ions are present in large numbers. The disadvantage is that the thermal and bulk velocities are comparable, so that a high velocity resolution is required over a large range in velocity. Mass and charge states must

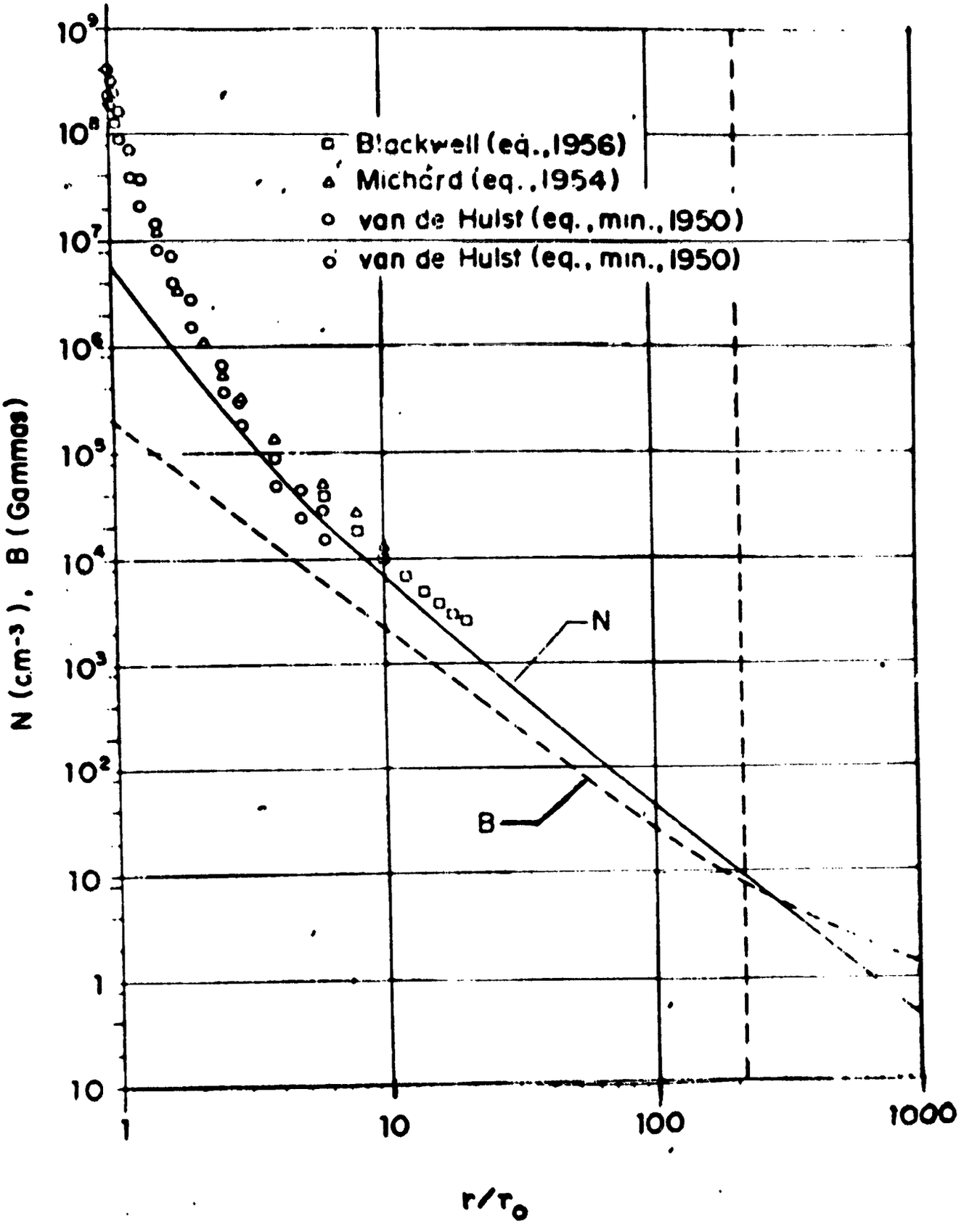


Figure 1. Density and magnetic field as a function of $R/R_{\odot}$.

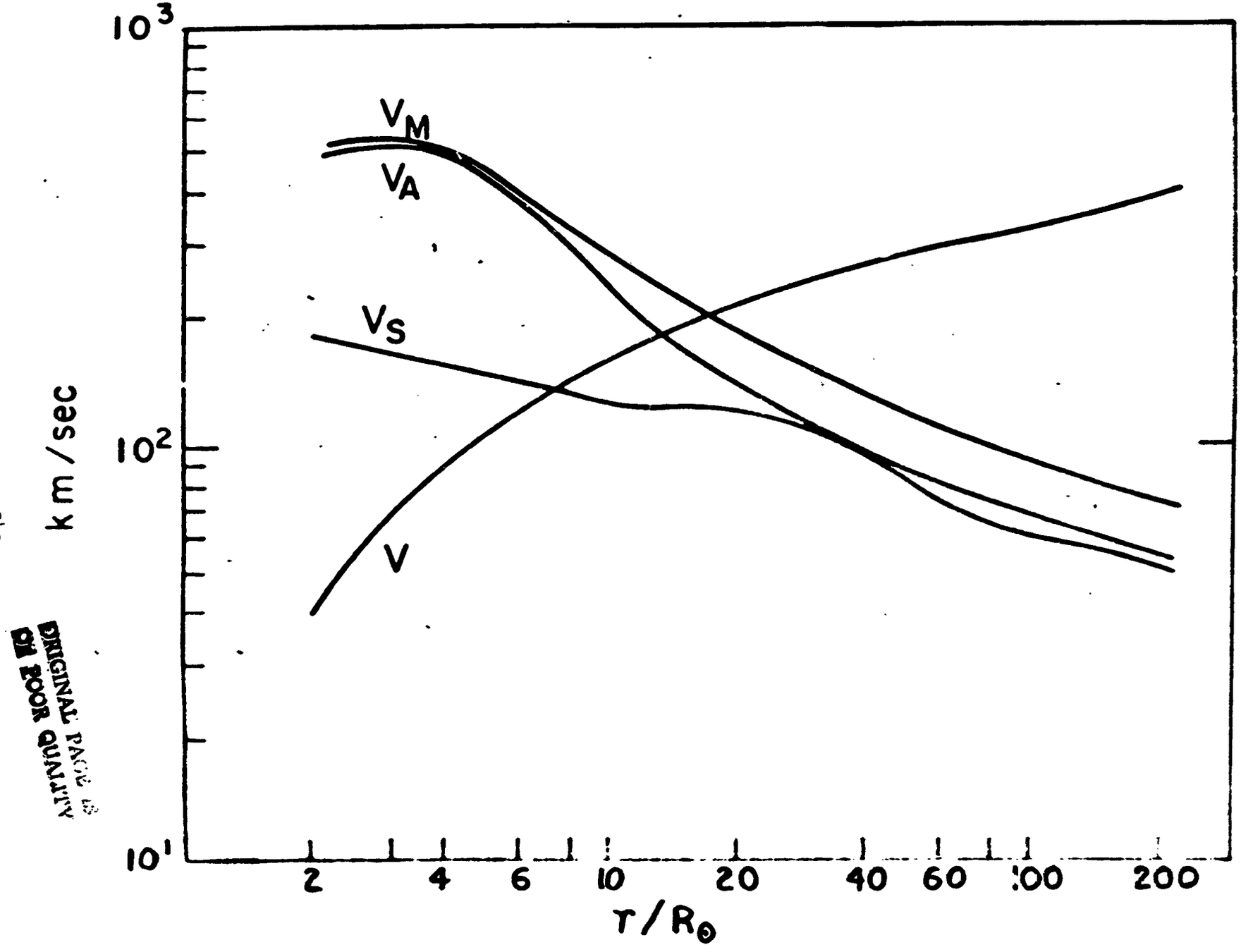


Figure 2. V_M = magnetosonic speed, V_A = Alfvén speed, V_S = sound speed, and V = solar wind speed.

be determined separately, since the most interesting region is highly turbulent.

Apart from direct ion temperature and composition measurements, the scientific aims outlined above can be carried out if we have reliable measurements of

Density
Electron temperature and temperature anisotropy
Bulk velocity
Electron heat flux along the trajectory

There are no substitutes for ion composition and ion temperature measurements, but there are many advantages in obtaining the other plasma parameters by measurements on electrons. The thermal speeds of the electrons are always large compared to the solar wind speed and the vehicle speed, so that electrons can be measured by an instrument which is always behind the heat shield. Thus there will be little trouble in ensuring that the detector survives. The design suggested in Fig. 3 is a six-axis narrow angle spectrometer. This is particularly well suited to measuring drift velocities in any direction since a detector looks up and down each of three mutually perpendicular directions. Each of the six cylindrical electrostatic analyzers has two detectors, so by providing two high voltage supplies and using different apertures and accumulation times, the above quantities may be measured all the way from 5 AU to 4 $R_\odot$. An additional advantage is that even though aberration is large and changes, and the flow may be in any direction when close to the sun, it can be measured by this instrument without any mechanical motion. A similar instrument has been operating on the ISEE-A spacecraft (a spinner) for the last six months, and has proven to be able to measure the solar wind speed to an accuracy of a few percent and its direction to an accuracy of one or two degrees. Figure 4 shows data obtained during a pass through the magnetosheath of the earth. Since the trajectory parameters must be known to high accuracy for the relativity experiment, the bulk speed can be determined to a similar degree of accuracy on this mission. If the spacecraft were spinning, the higher moments of the electron distribution function could be determined more frequently, but the experiment is compatible with a non-spinning spacecraft provided that the magnetic field direction is available. Data can be taken in

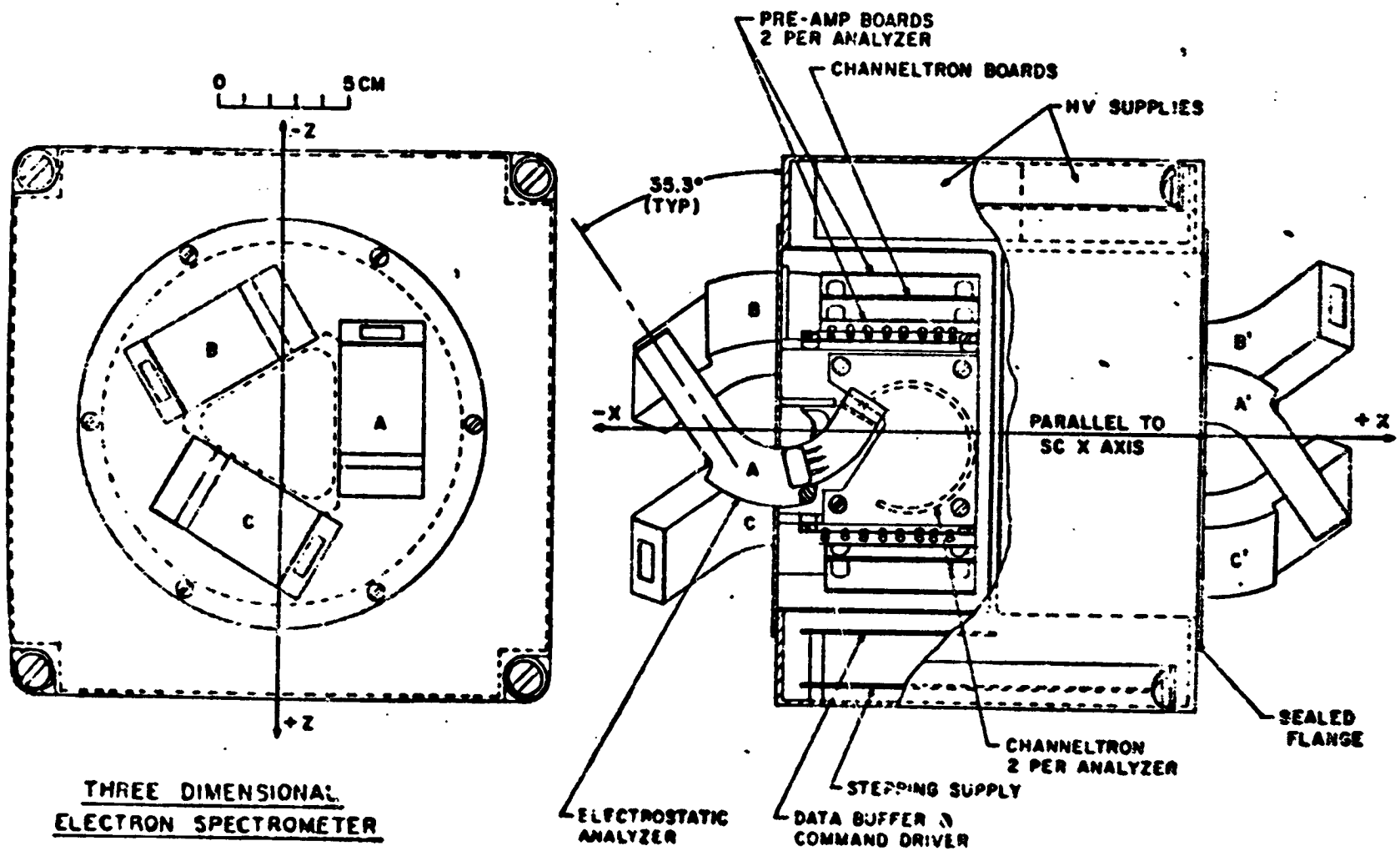


Figure 3. Layout of a possible six-axis electron spectrometer.

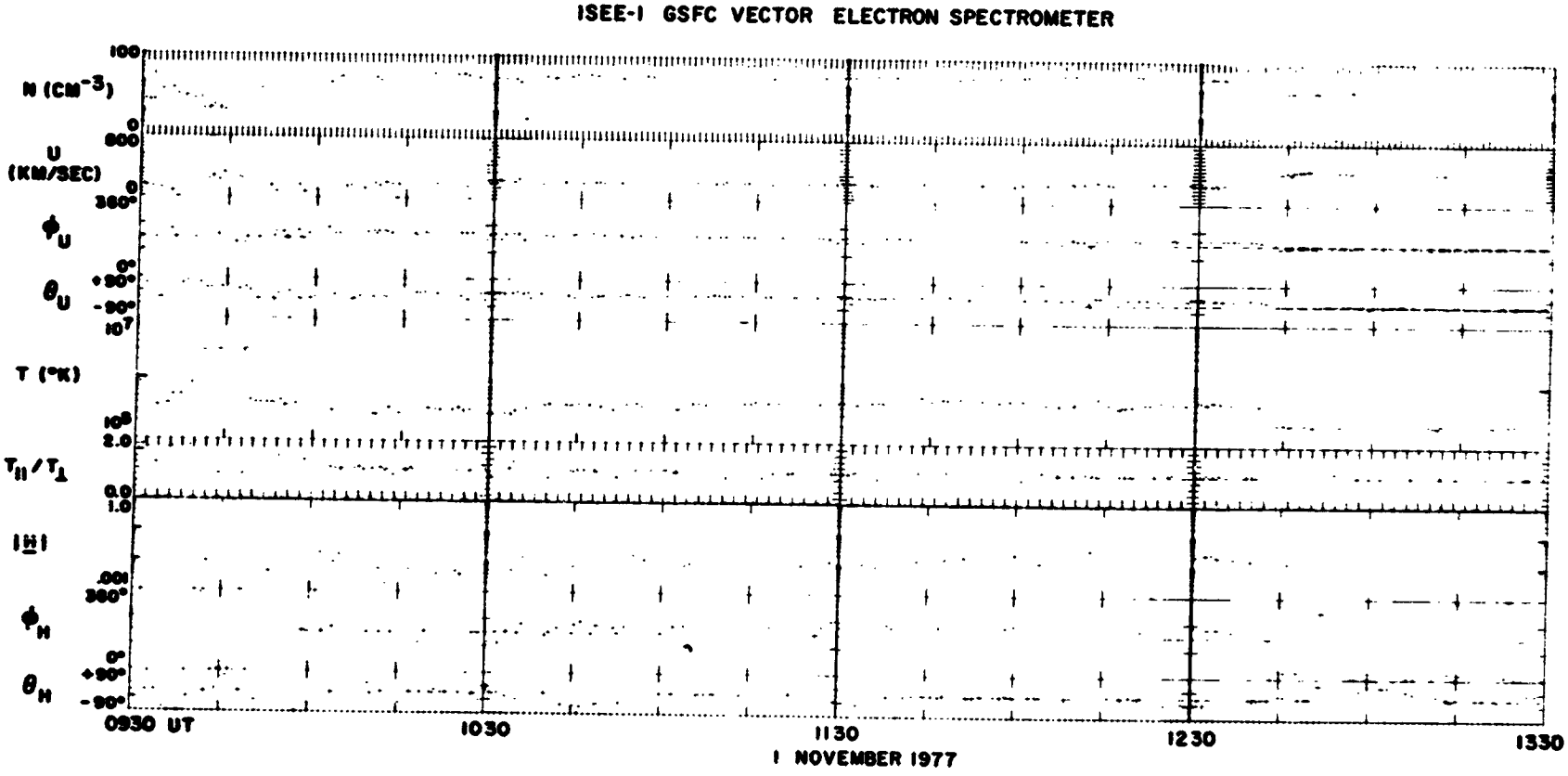


Figure 4. Data obtained (during an outbound passage through the earth's magnetosheath; $n \sim 30$, $T \sim 5 \times 10^5$ K, $U \sim 250$ km sec^{-1}) by the six-axis electron spectrometer on ISEE-1. The quantities plotted are, reading downwards; the density; the bulk speed; azimuth and latitude angles for the speed; the temperature; the temperature anisotropy; the heat flux magnitude and its azimuth and latitude angles. The bow shock is at 1243 and the magnetopause at 0948. Bulk speeds inside the magnetosphere are subject to interpretation on account of the combination of low density ($\sim$1) and high temperature ($\sim 2 \times 10^6$).

a "snapshot" mode and read out comparatively slowly to prevent aliasing; data storage is required to achieve full effectiveness. In this way we arrive at the following requirements:

Mass	4 kg	Electrons only.
Power	3 w	
Telemetry	100-200 bps	

The instrument should be mounted on the anti-sunward end of the spacecraft looking out at various angles to the axis. These angles will be decided after an extensive simulation program designed to ensure that the parameters can be measured with the required accuracy when close to the sun. In view of the small Debye length compared to the typical spacecraft dimensions, special care should be taken to make the spacecraft conducting.

D29

N78-32993

SPECTROSCOPIC MAPPING OF SOLAR WIND VELOCITIES

John L. Kohl
Harvard-Smithsonian Center for Astrophysics

ABSTRACT

During the total solar eclipse of 1970, measurements of resonantly scattered Lyman-alpha radiation from the solar corona revealed a means to determine temperatures and densities in the solar corona beyond 1.5 $R_\odot$. A natural extension of this work is to use the Solar Probe to measure the spectral line profile of Lyman-alpha radiation backscattered toward the Sun from coronal regions between 4 $R_\odot$ and 10 $R_\odot$. The backscattered profile would provide unique and quantitative determinations of the outflow velocities of coronal material into the solar wind. Such information is of critical importance for understanding solar-wind formation and solar-terrestrial effects on the Earth's climate. There is no known way to obtain this information without a Solar Probe-type mission.

I. INTRODUCTION

Measurements of spectral line profiles of coronal hydrogen Lyman-alpha radiation can provide a means for determining the solar wind velocities and temperatures of coronal regions within about 10 solar radii from sun center ($R_\odot$). Outflow velocities of coronal material into the solar wind are needed to understand the supersonic coronal expansion. The kinetic temperature of coronal protons is a key parameter for determining the amount and mechanisms of heating in the various structures of the solar wind acceleration region (presently believed to be between 2 $R_\odot$ and 10 $R_\odot$). Theories of solar-wind acceleration are closely related to a collage of natural phenomena which include solar-terrestrial effects on the Earth's climate, accretion, the escape of matter from stellar atmospheres, and stellar formation.

The present state of knowledge of solar-wind dynamics is severely limited by a complete lack of experimentally determined velocities and temperatures in the acceleration region. Solar-wind velocities have been observed at 1 AU, with some measurements as close as 0.3 AU, and Doppler shifts of coronal lines have provided some evidence of outflow velocities in coronal holes at about 1.1 $R_\odot$. Coronal temperature measurements have been limited to regions within 1.5 $R_\odot$.

The Solar Probe provides a unique opportunity to determine solar-wind velocities in the acceleration region from direct measurements of backscattered Lyman-alpha radiation. Only outwardly directed observations of Lyman-alpha radiation from within the corona are sufficiently sensitive to determine solar-wind velocities. The Harvard-Smithsonian Center for Astrophysics has built rocket instrumentation to observe the Lyman-alpha corona in the absence of a natural solar eclipse (Kohl, Reeves, Kirkham,1978). The first measurements are planned for the fall of 1978. Measurements from near the Earth of coronal Lyman-alpha line profiles are expected to provide coronal hydrogen temperatures but provide very little information about velocities. Observations with a visible light coronagraph are useful for determining electron densities. The combination of outwardly directed observations of Lyman-alpha line profiles from Solar Probe and Earth-based synoptic observations near the time of solar encounter with a Lyman-alpha coronagraph and a visible-light coronagraph would provide velocity, temperature, and density maps for a large segment of the solar corona. Such information can be conveniently displayed as artificial solar eclipse pictures in false color.

II. THE LYMAN-ALPHA CORONA

The Lyman-alpha corona beyond 1.2 $R_\odot$ (see Fig. 1) was observed for the first and only time by an international group of experimenters from Imperial College, London; York University, Canada; Culham Laboratory, U.K.; and the Harvard College Observatory who used rocket borne instrumentation during the total solar eclipse of 1970 (Gabriel et al., 1971). The relatively large intensity and slow decrease in intensity with radial distance compared to other spectral lines was explained by Gabriel (1971) as resonant scattering of chromospheric Lyman-alpha radiation by a proportionately small amount of neutral hydrogen in the high temperature coronal plasma.

Rough estimates of the expected intensity of coronal Lyman-alpha radiation as a function of radial distance from sun center are given in the lower part of Figure 2. The curves between 5 $R_\odot$ and 10 $R_\odot$ are only constrained by the end points and the values at 10 $R_\odot$ may very well be too large. At any point in the corona, the intensity is expected to scale directly with electron density and as the minus 3/2 power of the electron temperature and also to depend on the geometry of the scattering process. The intensities of Figure 2 are based on the 1970 eclipse observations out to 1.8 $R_\odot$ with extrapolations to 10 $R_\odot$ using observations of the K corona (cf. Allen, 1973) which also scale directly with electron density. For simplicity, a constant coronal temperature was assumed. The usual $(r/R_\odot)^{-2/7}$ temperature dependence would result in about a 45% increase in intensity between 4 $R_\odot$ and 10 $R_\odot$. A more rapid fall off in intensity can result from solar-wind velocities in excess of 100-150 km s^{-1}.

Resonant scattering of Lyman-alpha radiation has been described by Beckers and Chipman (1974). The geometry of the scattering process can be described in terms of the angle θ between the radial direction through the coronal point of interest and the direction of observation (see Fig. 2). Relevant spectral line profiles are shown in Figure 3. Considering resonant scattering as an absorption and reemission process, the chromospheric profile can be thought of as the incident radiation and the 1.5×10^6 K gaussian profile as the absorption profile due to the velocity distribution of ground state hydrogen in the direction of the incident radiation. Coronal hydrogen atoms only can absorb photons with wavelengths within their natural line width ($\sim 3 \times 10^{-6}$ nm). The result of the absorption of the chromospheric radiation is a distinctly non-gaussian velocity distribution of excited atoms in the direction of the incident radiation. In perpendicular directions, the velocity distribution of excited atoms would be the same as for ground state atoms.

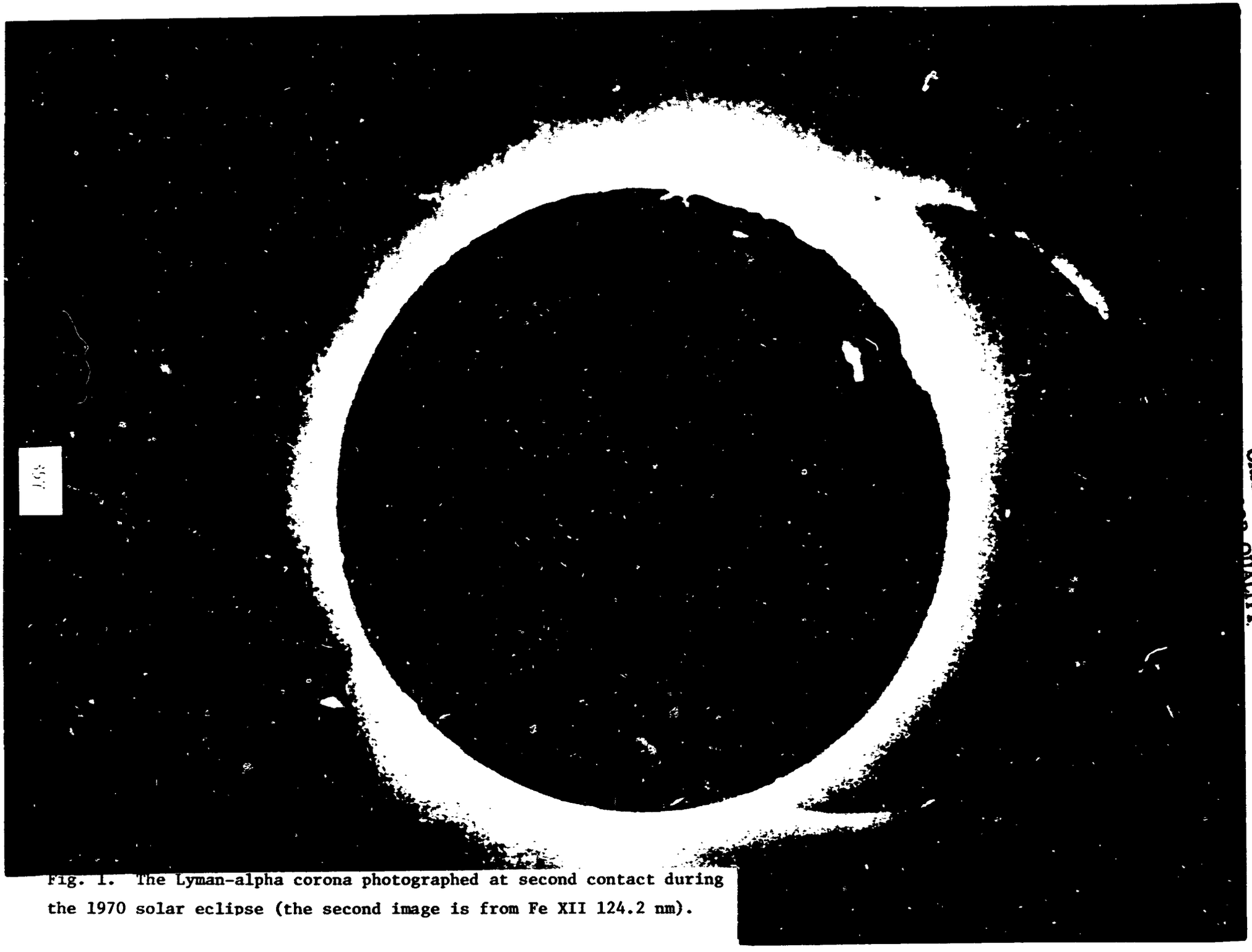


Fig. 1. The Lyman-alpha corona photographed at second contact during the 1970 solar eclipse (the second image is from Fe XII 124.2 nm).

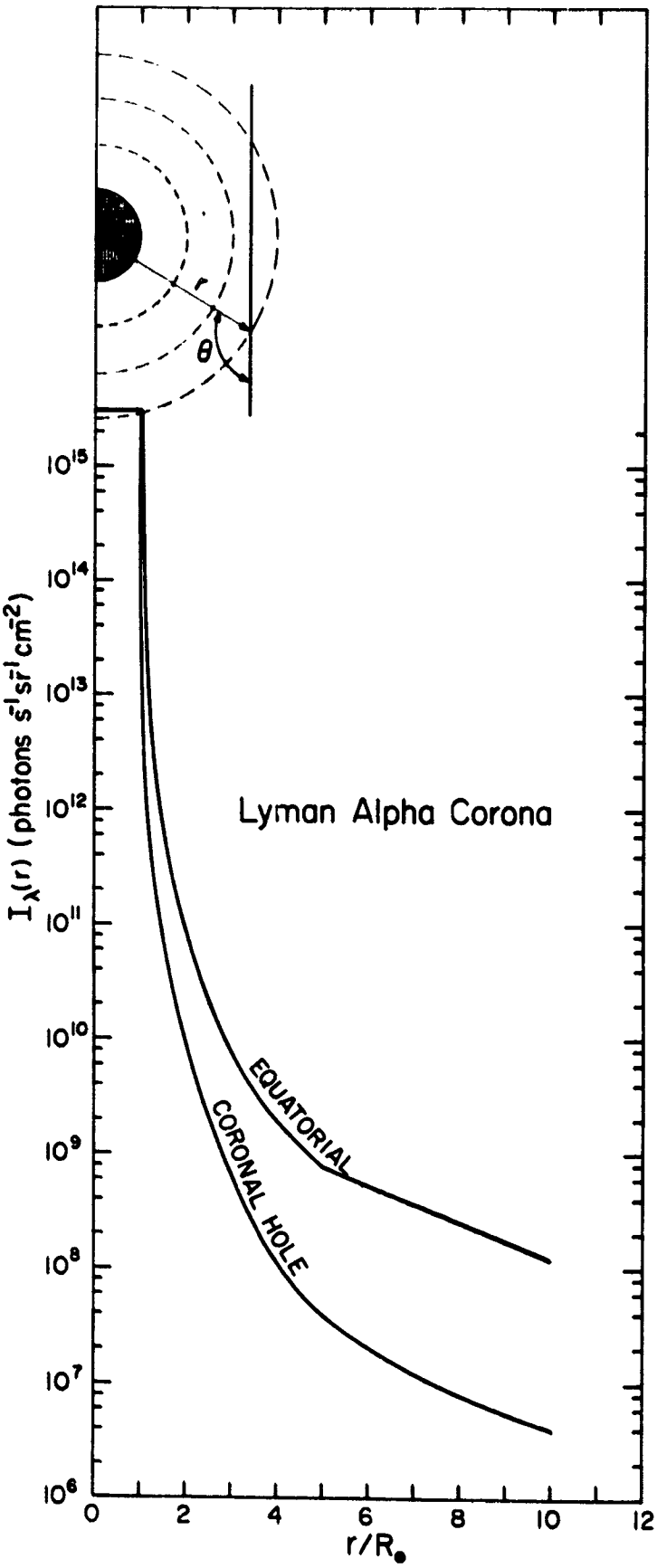


Fig. 2 Estimated intensity of the Lyman-alpha corona versus the distance from Sun center (in solar radii $R_\odot$) for typical equatorial regions and coronal holes. The upper part of the figure represents observations of the Lyman-alpha corona for a line-of-sight inclined at angle θ with respect to a radial line through the point of observation of interest.

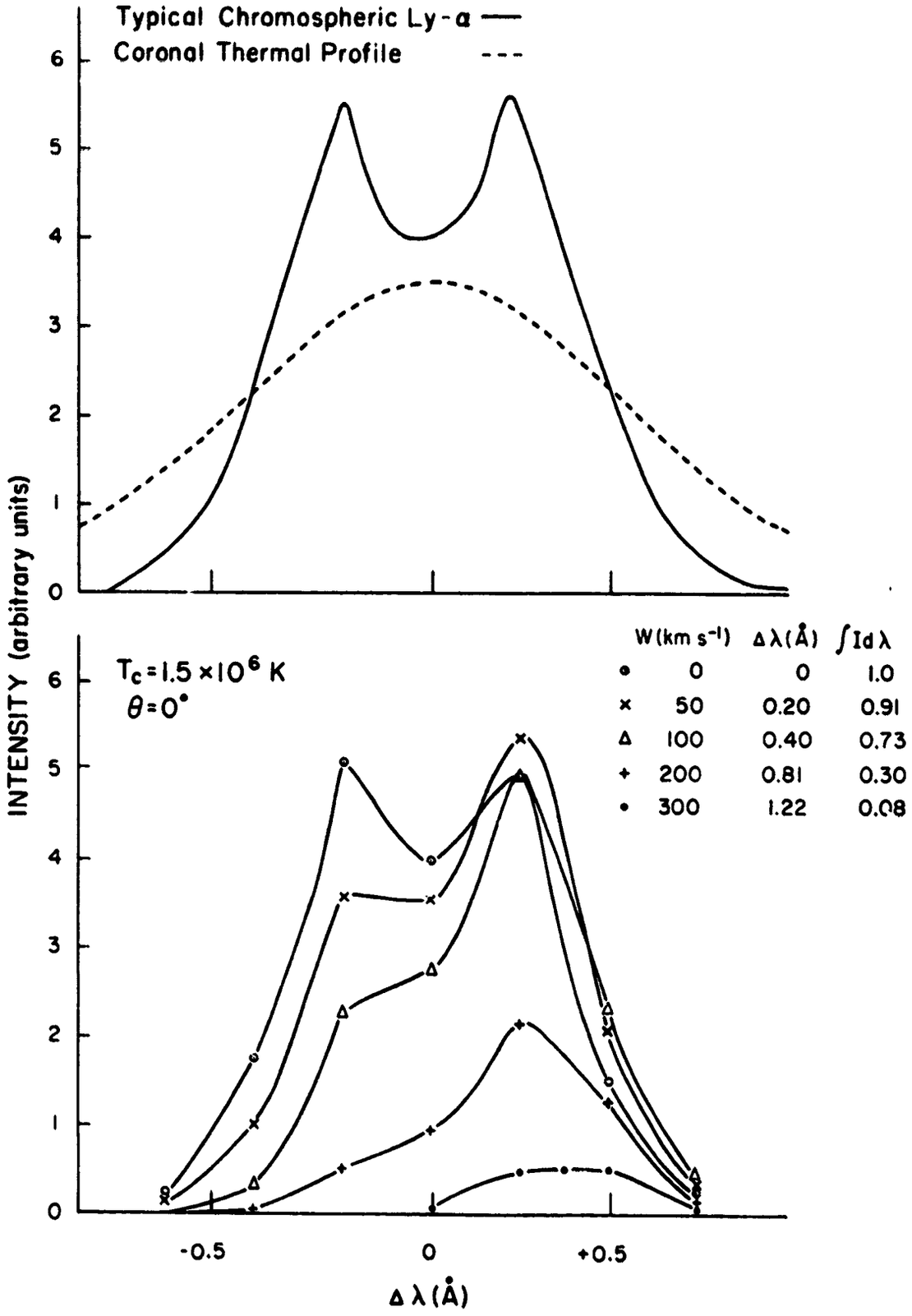


Fig. 3 Line profiles associated with resonant scattering of Lyman-alpha radiation. The upper part of the figure shows a typical emergent Lyman-alpha profile from the disk (solid line) and a 1.5 x 10^6 K thermal gaussian profile (dashed line). The lower part of the figure shows the expected coronal profiles for resonant backscattering toward Sun-center for several different solar-wind velocities.

Because the observed line profile is determined by the velocity distribution of excited atoms in the line-of-sight direction, the profile of scattered radiation at $\theta = 90^{o}$ would be similar to the gaussian curve of Figure 3. However, for backscattered radiation ($\theta = 0^{o}$), the velocity distribution of excited atoms and hence the reemission profile is proportional to the product at each wavelength of the incident chromospheric profile and the velocity profile of ground state coronal hydrogen. A net outflow velocity has the effect of shifting the absorption profile toward longer wavelengths with respect to the chromospheric profile. This dramatically affects the velocity distribution of excited atoms but only in the direction of the incident radiation. The resulting Lyman-alpha profiles for backscattered radiation are quite sensitive to the 25-100 km s^{-1} velocities that are expected over much of the acceleration region. A set of such profiles is shown in the lower part of Figure 3. The solar-wind velocities can be derived from the shapes of the profiles. Neither absolute wavelength positions nor absolute intensities are needed although they both contain useful information.

The backscattered profiles depend on the velocity profile of the absorbers which should be determined independently. A major advantage of hydrogen as opposed to heavier elements is that the thermal velocities are relatively large compared to expected turbulent velocities. Hence, it is probably not necessary to measure the velocity distribution in precisely the direction of the incident radiation.

The overlap of the chromospheric profile and Doppler-shifted coronal profile decreases with increasing outflow velocity. The resulting change in integrated intensity of the scattered radiation affects all θ angles equally. This has been termed "Doppler dimming" and has been described by Hyder and Lites (1970). In principle, velocities can be determined by looking for differences in the intensity gradient between resonantly scattered Lyman-alpha radiation and Thomson scattered visible radiation. However, Doppler dimming of Lyman-alpha is only a 27% effect for outflow velocities of 100 km s^{-1} (cf. Fig. 3). Because of uncertainties in the predicted intensity due to the hydrogen ionization balance, differences in electron temperature versus proton temperature, and line-of-sight uncertainties, Doppler dimming is not expected to be useful for determining velocities less than about 200 km s^{-1}. Even in this case, large uncertainties would be present.

The profiles given in the lower part of Figure 3 represent the contribution from a region of constant temperature and outflow velocity. Actual profiles at $\theta = 0^{o}$ would contain contributions from along the line-of-sight. Since the intensity decreases with radial distance, the contribution functions tend to be peaked toward the smaller radial distances. The decrease in intensity with the expected increase in outflow velocity provides an additional decrease in the contribution from larger radial distances beyond that shown in Figure 2. We have initiated a program to evaluate the expected contribution functions with computer generated observations.

III. LYMAN-ALPHA MEASUREMENTS FROM SOLAR PROBE

The principal objective of a measurement of coronal Lyman-alpha radiation from Solar Probe would be to determine the net outflow velocities of coronal hydrogen into the solar wind for a variety of coronal features between 4 $R_{\odot}$ and 10 $R_{\odot}$. There is no known method to obtain this information without a Solar Probe-type mission. Lyman-alpha observations could also be used to determine thermal and turbulent velocity distributions of coronal hydrogen.

As discussed in Section II, the expected outflow velocities can be determined from measurements of the spectral line profiles of backscattered Lyman alpha radiation ($\theta = 0^{o}$). The required instrument is a simple telescope-spectrometer system placed in the shadow of the heat shield and directed outward from the Sun along the Sun-probe line. A suitable angular resolution of the telescope-spectrometer system would be about 1° which corresponds to 0.017 $R_{\odot}$ at a distance of 1 $R_{\odot}$ from the probe. Such an instrument could be used to obtain the required Lyman-alpha profiles for the $\theta = 0°$ direction. Additional observations from Solar Probe at $\theta = 45°$ and 70° would be useful for determining the velocity distribution of coronal hydrogen which is required for a full analysis of the profiles observed at $\theta = 0°$. Observations of the Lyman-alpha corona from near the Earth could be used to help establish contribution functions from along each line-of-sight (see Sec. II).

There are several positive technical aspects associated with velocity determinations from measurements of backscattered Lyman-alpha radiation. Recall that the expected solar-wind velocities and the thermal velocities of hydrogen in the acceleration region are of similar magnitudes, and the planned speed of the spacecraft through the coronal plasma is also in this general range. This situation would not be troublesome for the Lyman-alpha observations. Because

the regions observed in Lyman-alpha radiation would be well away from Solar Probe, they would not be affected at all by any local disturbance of the environment near the probe. Spacecraft motion would not affect the velocity determination since it would be based on the shapes of the profiles other than wavelength shifts. Measurements of backscattered radiation sample the velocities moving away as well as toward the probe. The Lyman-alpha observations are sensitive to the expected outflow velocities. Unlike electron velocities, the expected net outflow velocity of hydrogen is an appreciable fraction of the hydrogen thermal velocity.

IV. INSTRUMENTATION

We plan to conduct a scientific and engineering study to establish the conceptual design of an instrument for Solar Probe to measure coronal Lyman-alpha radiation. The design will draw on our experiences with a Lyman-alpha coronagraph instrument. Meanwhile, the instrument can be considered to be a short focal length telescope with a baffle system along the telescope axis and a 0.25 meter spectrometer. The instrument would have no moving parts. A multiple anode array detector with about 60 elements would be used. The primary telescope would be directed along the Sun-probe line away from Sun center. Additional telescope mirrors could be directed at $\theta = 45^{\circ}$ and/or 70°. Multiple spectrometer entrance slits would be used for the various telescope images. The instrument would be placed in the shadow of the heat shield. The required spectral resolution is about 0.01 nm and the angular resolution would be about 1°. As a design goal, the weight would be 4 to 6 kg. The power requirements for the high voltage power supply and data system would be about 10 watts. The data rate would be about 48 bps. The expected temperature excursions will require a careful design in order to maintain spectral focus. The electronics would have the usual sensitivity to radiation damage. The acceptable detector background count rate would put constraints on the radiation environment.

V. MEASUREMENTS FROM NEAR THE EARTH

Several types of solar observations from near the Earth would be useful at the time of the solar encounter to support Lyman-alpha observations from Solar Probe. Synoptic observations with a visible light coronagraph during the several days preceding and following the solar encounter could be

used to determine three dimensional electron density maps. Measurements of the emergent chromospheric Lyman-alpha profile would be useful for the analysis of the Lyman-alpha profiles obtained with the Solar Probe instrument. Synoptic observations of the coronal Lyman-alpha profile and intensity could be used to determine thermal and turbulent velocity distributions and help evaluate the line-of-sight contributions of the Solar-Probe observations.

Rocket instrumentation has been developed at this laboratory for measuring the intensity and spectral line profile of coronal Lyman-alpha radiation in the absence of a natural solar eclipse. A scanning spectrometer with photoelectric detection is combined with a linear occulting system which reduces stray light levels near 121.6 nm by about seven orders of magnitude with respect to solar disk radiation. This provides adequate stray light suppression for observations out to 4.0 $R_\odot$ from sun center. Straightforward modifications of the occulting system would be required for measurements beyond 4 $R_\odot$ and spectral resolution of about 0.003 nm would be required to isolate cleanly the narrow geocoronal Lyman-alpha line within the solar coronal line profile.

Because of the rapid decrease in Lyman-alpha intensity with radial distance, coronagraph and eclipse measurements of the Lyman-alpha corona tend to sample coronal regions with scattering angles θ near 90°. The contributions to the observed profiles would be small for points along the lines-of-sight corresponding to θ angles significantly larger than 135° (see Fig. 2). Hence the resulting profiles are insensitive to outflow velocities.

The rocket version of the new Lyman-alpha coronagraph will be used for the first time in the fall of 1978. Coordinated observations of the visible-light corona will be carried out by the High Altitude Observatory. Figure 4 illustrates a typical observing program of the Lyman-alpha coronagraph for the early rocket flights. The Lyman-alpha profile and intensity would be measured for several 4′ by 0.6′ regions of a coronal streamer between 1.5 $R_\odot$ and 4.0 $R_\odot$ and for a coronal hole out to 2.0 $R_\odot$. Coronal temperatures will be determined from the line shape and neutral hydrogen densities will be estimated from the visible light observations. The data from the two instruments will be combined to investigate the ionization equilibrium of hydrogen, the amounts of mechanical heating in the corona and the location where this heat is deposited. Some indication of outflow velocities might be contained in the relative intensities

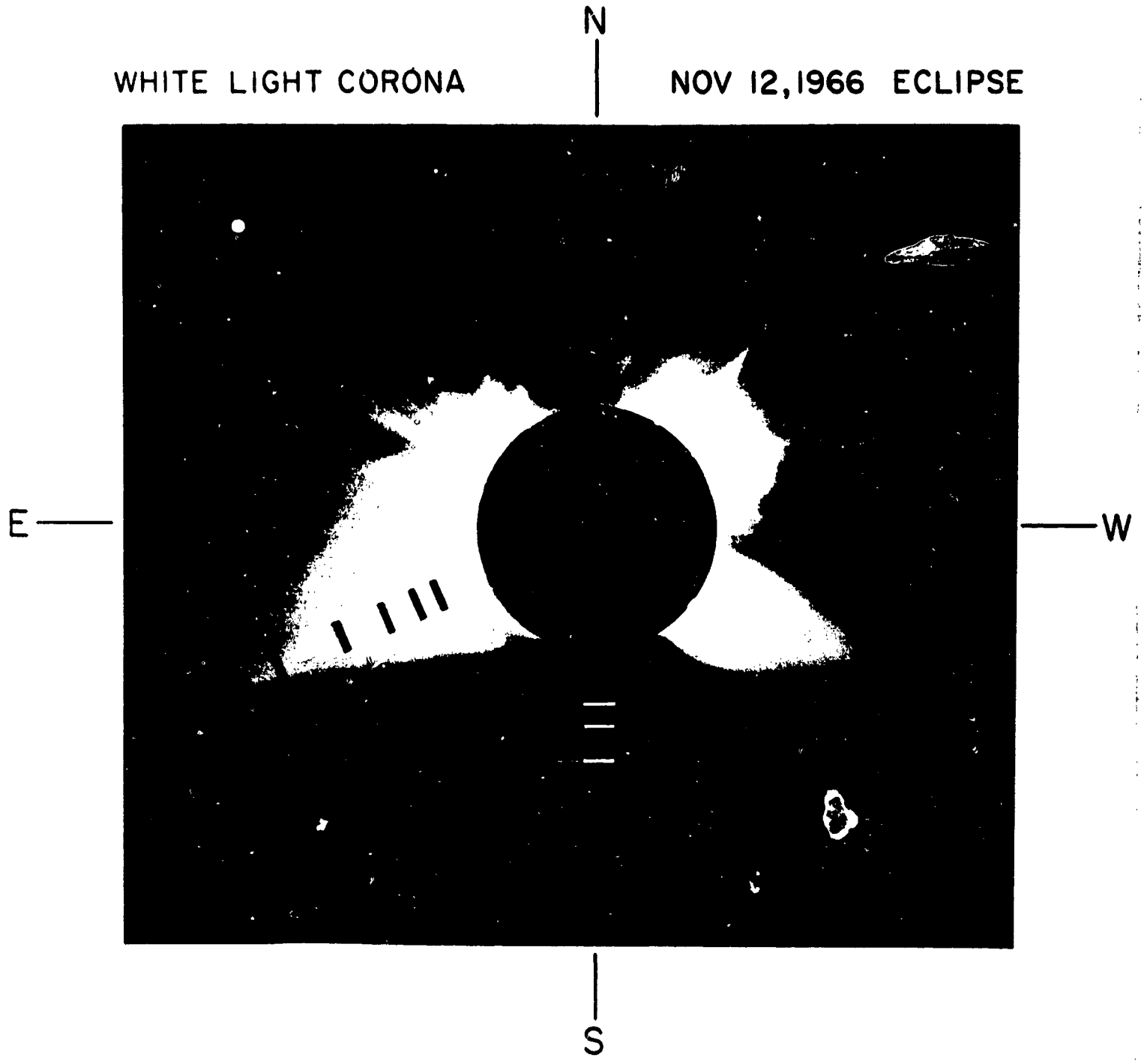


Fig. 4. Typical coronal regions to be observed with a new rocket borne Lyman-alpha coronagraph. The white light corona is from the 1966 solar eclipse (Newkirk et al., 1970).

C-5

of the Lyman-alpha radiation and the visible light.

To obtain extensive observations of the Lyman-alpha corona which include synoptic observations, an upgraded Lyman-alpha coronagraph is being planned for Spacelab.

REFERENCES

Allen, C.W. 1973, Astrophysical Quantities (London: Athlone Press, 3d ed.) chap. 9.

Beckers, J.M. and Chipman, E. 1974, Solar Phys., 34, 151.

Gabriel, A.H. 1971, Solar Phys., 21, 392.

Gabriel, A.H., Garton, W.R.S., Goldberg, L., Jones, T.J.L., Jordan, C., Morgan, F.J., Nicholls, R.W., Parkinson, W.H., Paxton, H.J.B., Reeves, E.M., Shenton, C.B., Speer, R.J., and Wilson, R. 1971, Ap. J., 169, 595.

Hyder, C.L. and Lytes, B.W. 1970, Solar Phys., 14, 147.

Kohl, J.L., Reeves, E.M., and Kirkham, B. 1973, in New Instrumentation for Space Astronomy, ed. by K. van der Hucht and G.S. Vaiana (New York: Pergamon) p. 91.

Newkirk, G., Jr., Dupree, R.G., and Schmahl, E.J. 1970, Solar Phys., 15, 15.

OMIT

A New Method of Modelling Coronal Magnetic Fields

M. Schulz, E. N. Frazier, and D. J. Boucher

The Aerospace Corporation
El Segundo, CA 90245

(Abstract)

We discuss a new method (Schulz et al., 1978) of modelling the coronal magnetic field which uses a non-spherical source surface. This surface is taken to be an isogauss of the underlying potential field generated by those currents which flow below the solar surface. With the Solar Probe, it will be possible to check how good this model is for extrapolation of fields between the photosphere and 1 AU. Of special interest will be the observation of non-radial fields between 4 and 10 $R_\odot$, the fields over active regions (where the source surface exhibits dimples), and across current sheets.

Reference

Schulz, M., Frazier, E. N., and Boucher, D. J., Jr., "Coronal Magnetic-Field Model with Non-Spherical Source Surface," accepted for publication in Solar Physics, 1978.

MAGNETOHYDRODYNAMIC SIMULATION OF CORONAL MASS EJECTIONS INTO THE SOLAR WIND

by

M. Dryer
Space Environment Laboratory
NOAA/ERL
Boulder, Colorado 80302

S. T. Wu and R. S. Steinolfson
School of Science and Engineering
University of Alabama in Huntsville
Huntsville, Alabama 35807

and

E. Tandberg-Hanssen and R. M. Wilson
Space Sciences Laboratory
NASA/Marshall Space Flight Center
Huntsville, Alabama 35812

ABSTRACT

It has long been recognized that the boundary conditions at and close to the Sun must be known for all physical models which purport to describe the flow of both quiescent and transient plasmas into the interplanetary medium. Determination by an *in situ* probe of the background or transient plasma and field characteristics would greatly aid our understanding. For example, assumptions for the background, equilibrium atmosphere for MHD computer simulations must now be considered as *ad hoc*. As a case in point, we note that magnetic field magnitudes and topologies in the low corona cannot be measured directly.

In order to obtain an assessment of the validity of such assumptions, we will review a state-of-art fluid continuum technique to describe the MHD transient response of the corona below 10 solar radii during a well-observed event during the Skylab mission.

More explicitly, a two-dimensional, planar MHD model of solar atmospheric transient events is used to simulate the response of the corona to a flare on 1973 August 21. This event, observed in Hα, He II 304 A, soft X-ray and coronal white light, provided sufficient information to assess the model's ability to simulate major features of an actual solar event. It was found that a thermodynamic input pulse based on data provided by the NASA/Marshall Space Flight Center-Aerospace Corporation X-ray telescope (S-056) on Skylab was sufficient to produce the global geometry, shock and contact surface velocities, excess mass contours and energy budget which were, for the most part, observed by the High Altitude Observatory white-light coronagraph (S-052) on Skylab in the form of a forerunner and coronal transient. Another event, that of 5 September 1973, is also briefly examined.

We conclude that the computer simulation technique, as demonstrated by the encouraging results in the present case, should be subjected to in situ verification of as many of its initial assumptions as possible. Its ability to provide a rational basis for physical understanding of mass ejections suggests its use as one of the tools used in the planning and analysis of such close encounter missions.

A number of non-linear, magnetohydrodynamic (MHD), numerical models are presently under development for the computer simulation of transient events in the solar atmosphere. It is the objective of these simulations to provide insight to the dominant physical processes within such events and, further, to predict their consequences in the interplanetary medium. A continuing confrontation between experiment and theory is necessary, paraphrasing the 1978 Report (Volume 1) of the NAS' Space Science Board, for the health of both as well as to achieve a real understanding of nature.

It is important, then, to review several basic facts concerning numerical simulations so that further confrontations can be assessed in an organized way. First, the output of such simulations is dependent upon a number of parameters. That is, numerical solutions of MHD models are functions of the following: (i) the initially-assumed equilibrium atmosphere; (ii) magnetic topology which is in equilibrium with this atmosphere; and (iii) magnitude and duration of the pulse (i.e., "flare", etc.) which drives the transient. Secondly, the output of such simulations depends upon assumptions involving the presence, magnitude, and type of dissipative processes. However, the solutions are unique for a given set of assumptions and parameters.

We adopt the view that coronal mass ejections into the solar wind may best be described via the fluid continuum assumption wherein the plasma and magnetic field are mutually interactive. Thus, we set aside the study of very low plasma beta examples ($\beta = 16\pi nkT/B^2 << 1.0$) wherein plasma is constrained to act within initially-prescribed magnetic topologies. The panoply of MHD waves must, therefore, be considered, including their formation and propagation as shock waves. The work described here is a summary and slight extension of recent studies by Steinolfson et al. (1978), Dryer et al. (1978), and Dryer and Maxwell (1978).

The next section describes an MHD simulation of the coronal response to flares on 21 August 1973 and 5 September 1973. These events were well-observed by ground-based observatories and by the Skylab ATM instruments.

TWO SIMULATIONS

The Events of 21 August 1973. Spatial and temporal soft X-ray observations of this east limb event were used to model the flare's pulse (Dryer et al., 1978). The initial, one-fluid, atmosphere, in force-free equilibrium with a magnetically, topologically-open portion of a six-lobed hexapole, was in polytropic equilibrium. The polytropic index was taken as 1.08, and the number density and temperature at the base (taken as 1 $R_\odot$ from the sun's center) were taken, based on observations, as $n = 5 \times 10^7$ cm^{-3} and $T = 2 \times 10^6$ $^\circ$K, respectively. Two values (for separate computations) of the magnetic field at the coronal base and on the axis-of-symmetry were taken as $B = 0.834$ G and 2.64 G. The first case will be referred to as the $\beta = 1.0$ case; the second, as the $\beta = 0.1$ case.

The soft X-ray observations were used, via the technique of flux ratios (c.f., Švestka, 1976, or Wilson, 1977), to provide a pressure pulse as a representation of the flare's transfer of energy and momentum to the corona and, ultimately, to the solar wind. Figure 1 shows the observationally-inferred density and temperature. It is seen that the rapid rise was followed by an elevated level of sustained output of thermal energy and momentum transfer. The full set of equations, described in detail by Steinolfson et al. (1978), was solved self-consistently and uniquely for the coronal spatial and temporal response in two dimensions. The asymmetrical shape of the fast, forward MHD shock wave is shown in Figure 2 one hour after the start of the flare (at 1340 UT on 21 August 1973). The larger Lorentz forces somewhat constrain the plasma flow behind the shock for the $\beta = 0.1$ case in comparison with the $\beta = 1.0$ case. The position of the maximum density increase is also shown for each case. Generally, the latter is seen to simulate rather closely the coronal transient observed for this event by Poland and Munro (1976).

The MHD simulation not only provides detailed information which can be compared to quantitative observations, but it also provides information where observations are unavailable. Figures 3 and 4, for example, show the density ratio, plasma velocity vectors, temperature ratio and

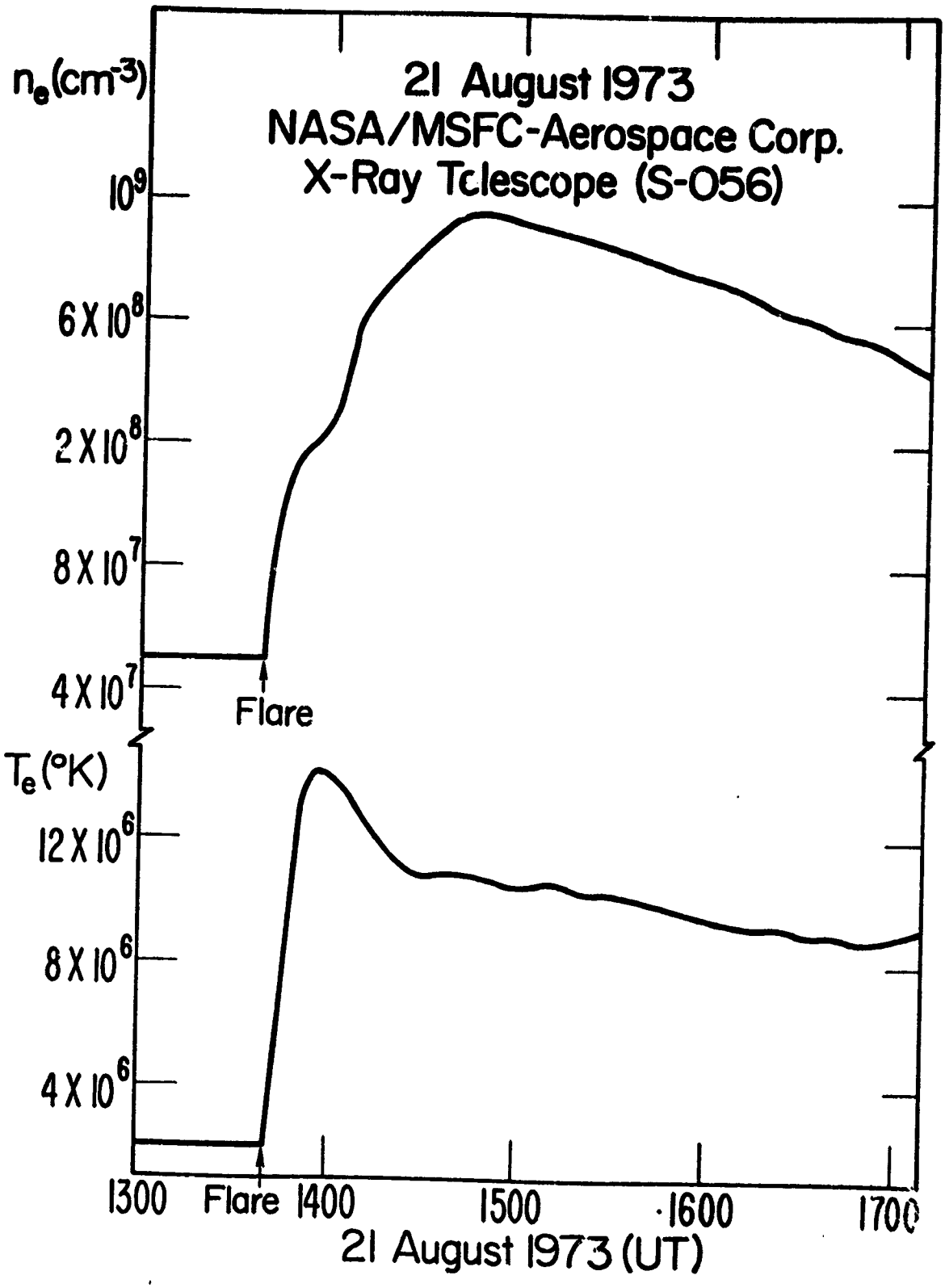


Fig. 1 Time series of electron temperature and density for the flare-initiated transient event on the east limb on 21 August 1973 as derived from the NASA/Marshall-Aerospace Corporation S-056 X-ray experiment on Skylab.

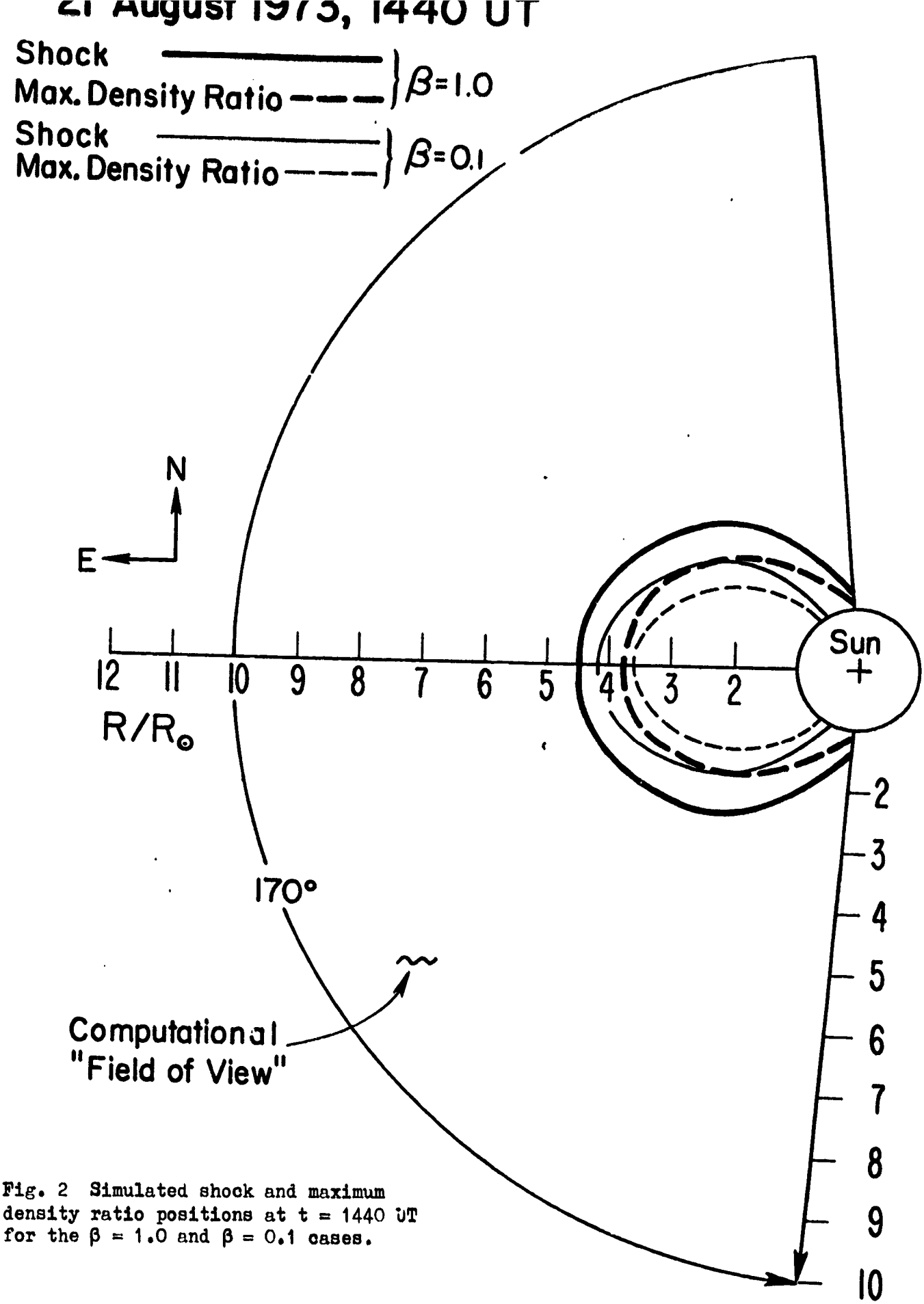


Fig. 2 Simulated shock and maximum density ratio positions at t = 1440 UT for the β = 1.0 and β = 0.1 cases.

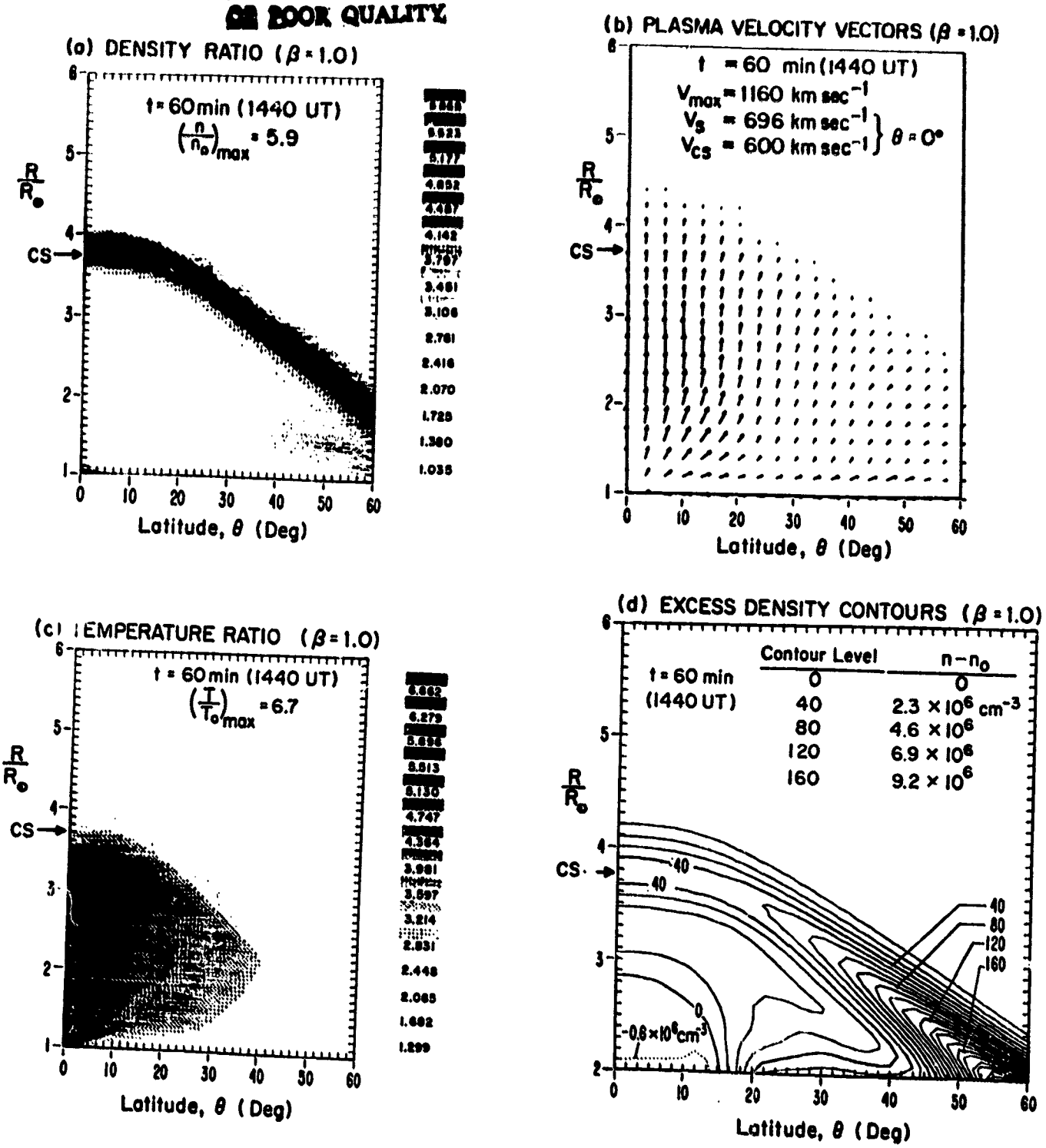


Fig. 3 Cross-sectional views (through one-half of the simulated field-of-view) of various plasma properties for the case: β = 1.0. The numbers assigned as contour levels by the computer graphics routine in Figure 3d are identified in the inset table as the physical change of local density.

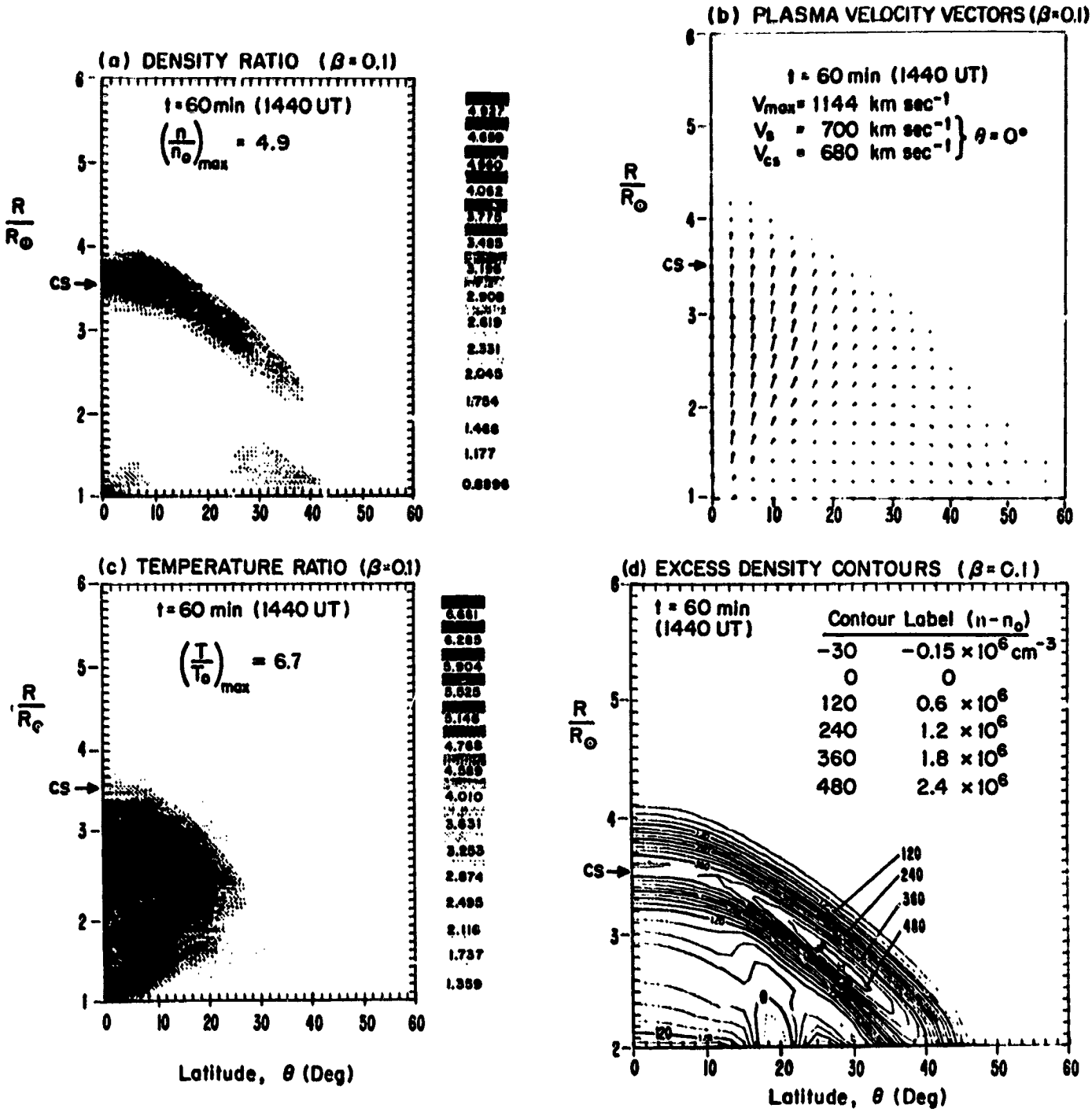


Fig. 4 Same as Figure 3 except for the β = 0.1 case.

excess density contours for the $\beta = 1.0$ and 0.1 cases, respectively. Note that only half of the computational field-of-view needs to be shown in view of symmetry about the solar equator at $\theta = 0^{o}$. The contact surface (denoted by CS) marks the position along the axis-of-symmetry of a plasma parcel which left the coronal base at 1340 UT, the instant of flare initiation.

Parenthetically, we remark that no attempt was made to simulate the "forerunner" as described for this and other events by Jackson and Hildner (1978). We believe that the pre-flare observations of coronal forerunners may be associated (via fast MHD wave density enhancements) with the pre-flare Hα and Fe XIV 5303 A observations reported by Martin and Ramsey (1972) and Bruzek and Demastus (1970). During the post-flare observations in white-light, the forerunners precede the denser coronal transients as discussed by Jackson and Hildner (1978). We further remark that the excess density contours of $n-n_{o}$ (where the subscript refers to the original, undisturbed value of density, etc., at each spatial position) may be directly compared with coronal transient observations. For example, Dulk et al. (1976) show such contours, including both compressions and rarefactions (viz., Figs. 3d and 4d of this paper) for the event of 14-15 September 1973.

The degree of magnetic control of coronal transients has been the topic of considerable discussion (c.f., Dulk et al., 1976; Dulk and McLean, 1978; Dryer et al., 1978). Values of magnetic field magnitude (or, more generally, the plasma β) must either be assumed or they must be inferred on the basis of fragmentary and usually indirect observational data coupled to ad hoc theoretical considerations (c.f., Dulk et al., 1976; Kuijpers, 1975).

With this uncertainty in mind, we show, in Figure 5, both the initial β distribution along the equatorial plane, $\theta = 0^{o}$, and the disturbed β at $t = 75$ min. It is clear that our initial, undisturbed β is unrealistically large at $R \gtrsim 2\ R_{o}$ for both cases. This fact, which stems from the original magnetic topology assumed for our model, is not serious because the non-linear coupled plasma/field interactions differ (for β's larger than one) by essentially only the characteristic speeds. Significant

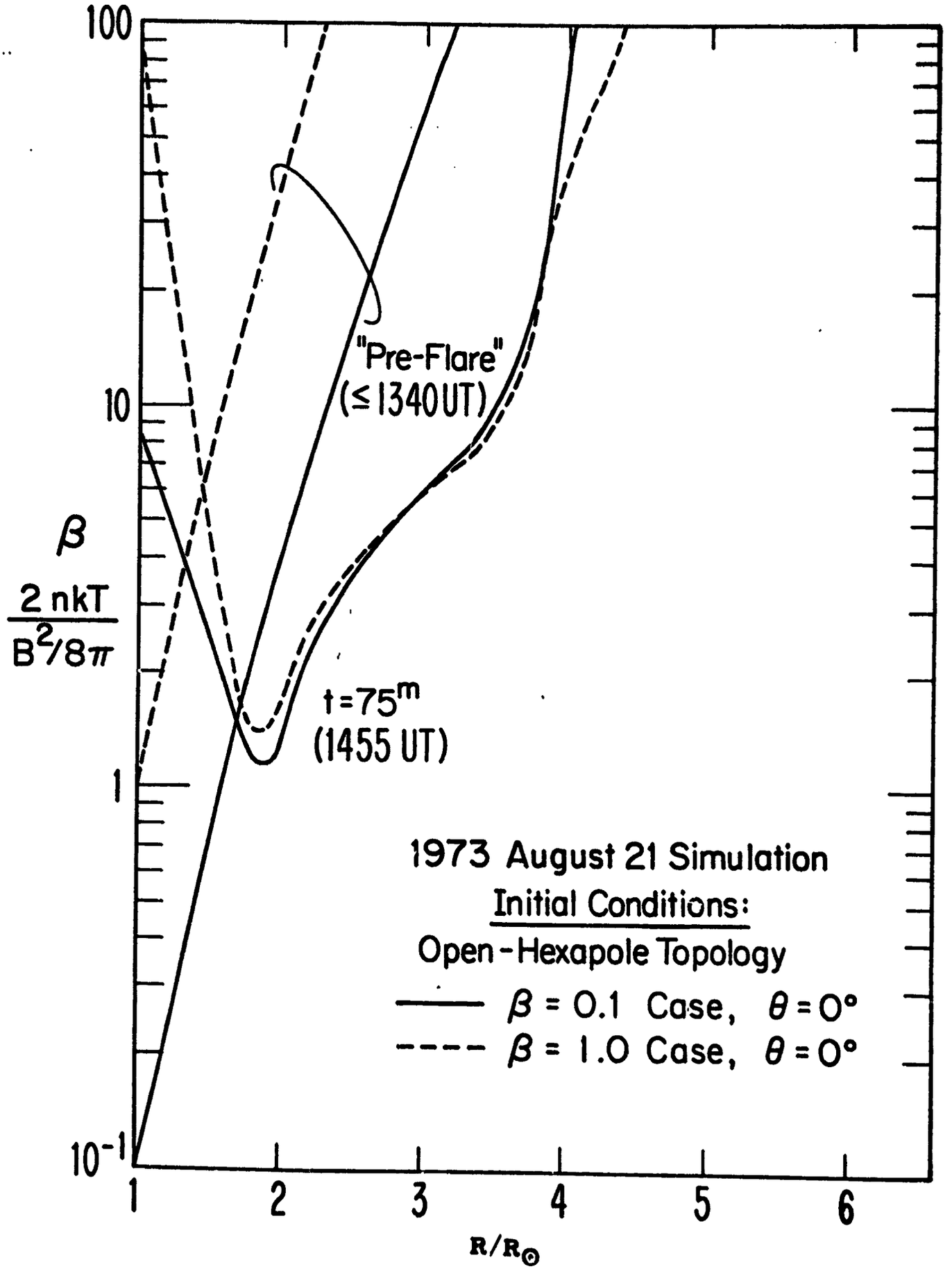


Fig. 5 Variation of the plasma beta in the initial, equilibrium atmosphere as well as within the coronal transient along the equator ($\theta = 0°$).

changes, of course, occur for $\beta << 1.0$. Figure 5 also shows that the temporal behavior of β within the coronal transient is about the same for both the $\beta = 1.0$ and 0.1 cases. It is interesting to note that the region of unrealistically-high values of β is now removed beyond $R \gtrsim 4\ R_o$. A discussion of the deformed magnetic topology is deferred to the summary below of the MHD simulation of the 5 September 1973 event.

It is also instructive to show the thermodynamic variables and plasma velocity in an alternative presentation. Figure 6a shows the thermal pressure ratio in the half-tone format used previously in Figures 3a and 3c, for $\beta = 1.0$, for the density and temperature ratios at 1440 UT. Note that the pressure is continuous across the contact surface (CS), whereas the density and temperature are anti-correlated as a result of the heated, rarefied plasma evaporated from the flare site ($-5^o \leq \theta \leq +5^o$) which is preceded by the shocked, compressed (albeit low temperature) plasma ahead of the contact surface. The absolute values of the density, temperature and velocity (again, corresponding to Fig. 3) are shown in Figures 6b, c and d. It is particularly interesting to note that a fast, reverse MHD shock wave begins to form behind the contact surface along the equator but rapidly weakens in the latitudinal direction. The saw-toothed shape of the velocity profile at $\theta = 0^o$ is reminiscent of forward reverse shock ensembles observed in the solar wind in corotating interaction regions. Whether the double-shock ensemble is produced by a long-lasting flare (as in the present case) or by a corotating coronal hole is immaterial with respect to the physics of the structure itself.

We conclude the discussion of the simulation of the 21 August 1973 event with an estimate of the mass and energy ejected. Since mass (and magnetic) flux is allowed to cross the coronal base in order to satisfy flux conservation, the excess mass within the computational field-of-view may be computed as a function of time. This mass build-up is shown in Figure 7 for the $\beta = 1.0$ case. The cut-off time is taken as 1910 UT when the shock passes the 6 R_o position. At approximately this time, the shock moves beyond the latitudinal boundaries (see Fig. 2) as some excess mass is, accordingly, not included in the computation. As

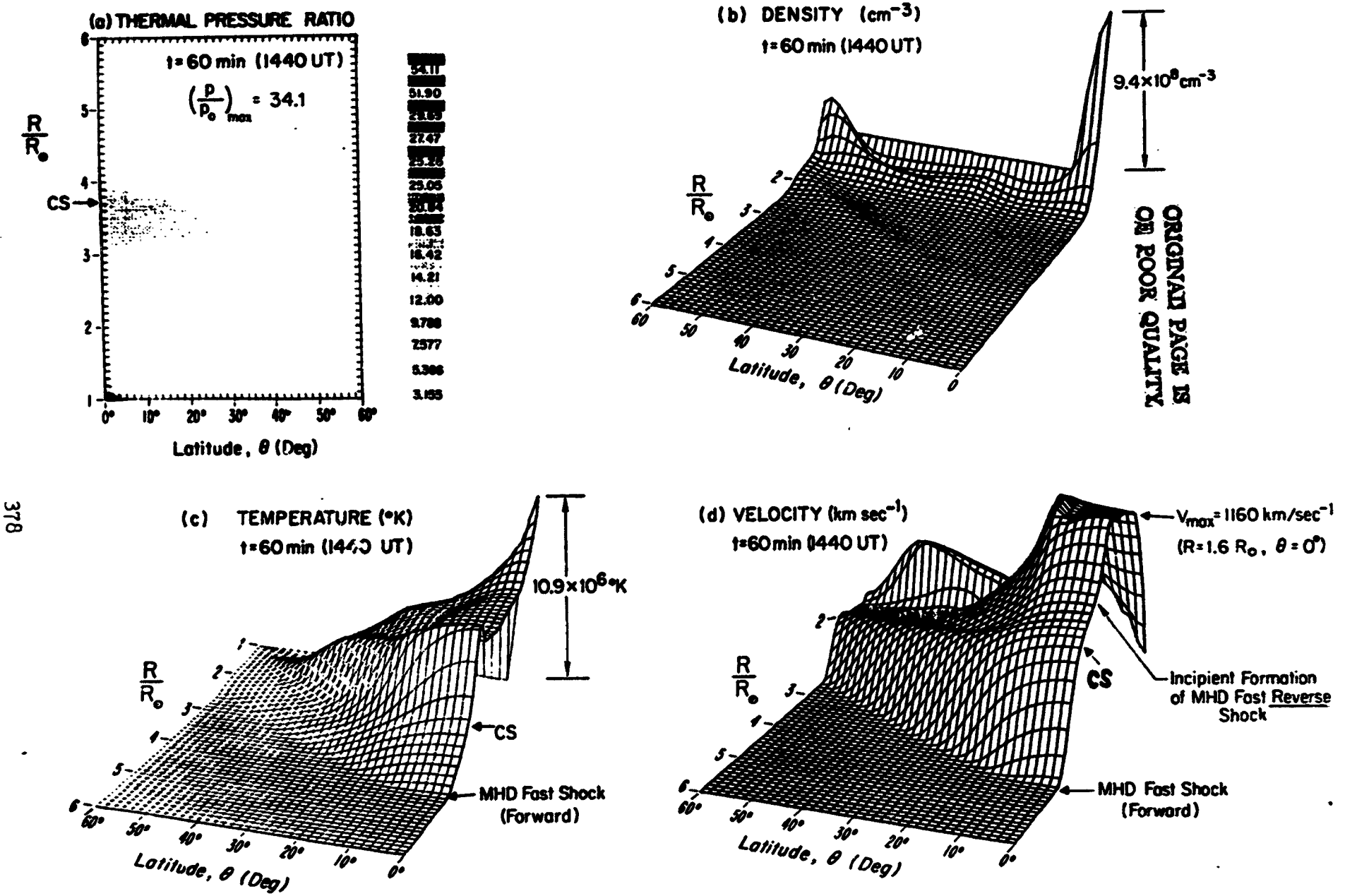


Fig. 6 Pressure ratio, density, temperature and velocity within the coronal transient at 1440 UT as simulated by the MHD model.

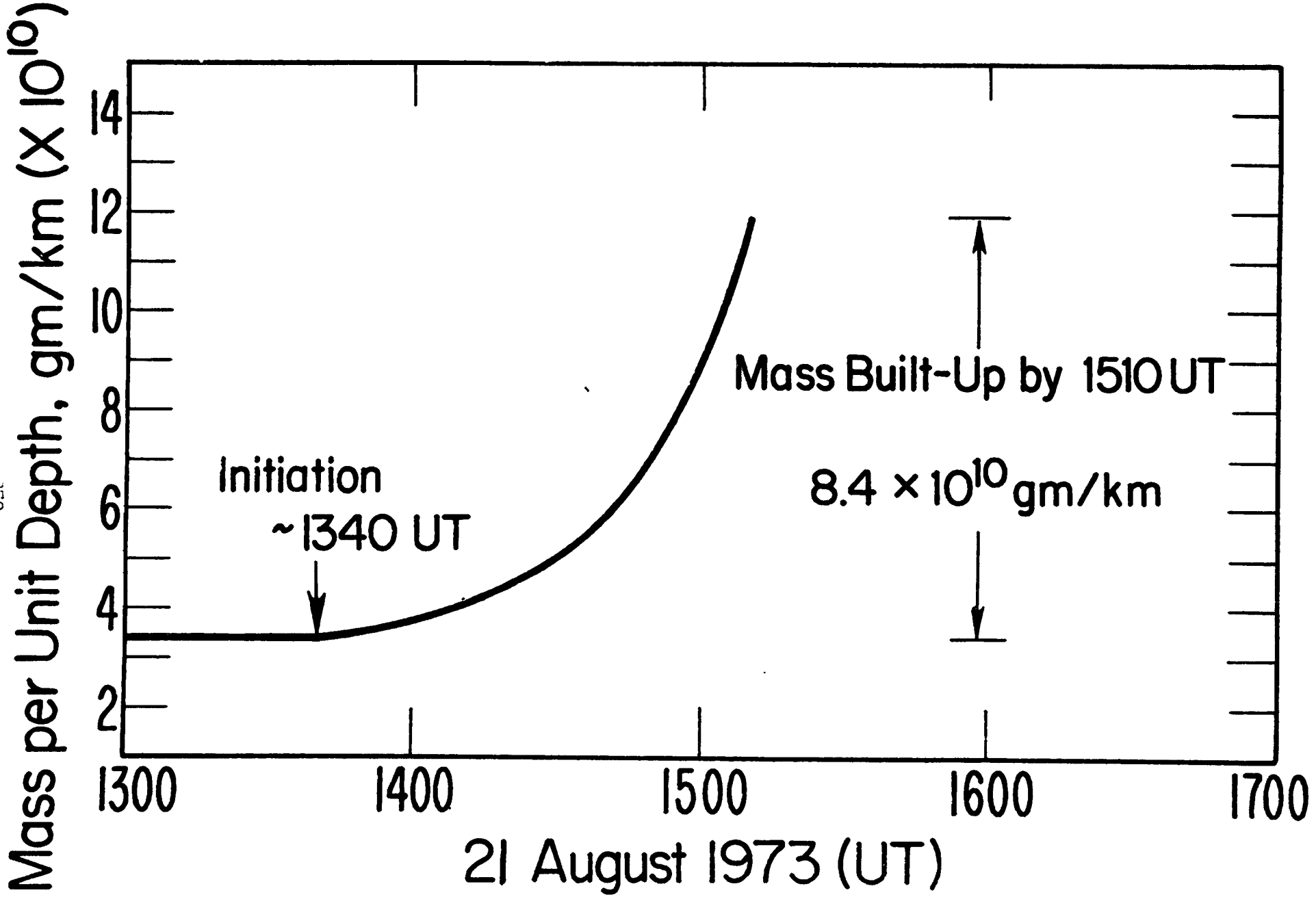


Fig. 7 Simulated mass ejected, per unit depth into the plane of the sky, as a function of time for the $\beta = 1.0$ case.

shown in Figure 7, the two-dimensional simulation provides an estimate of 8.4×10^{10} g km^{-1}. By invoking symmetry and an assumed depth of 0.2 $R_\odot$ (140,000 km), as suggested by Poland and Munro (1976), Dryer et al. (1978) estimated the mass ejected to the solar wind to be 2.3×10^{16} g. The observational estimate, based on the white-light coronal transient, is from $\sim 8 \times 10^{15}$ g (Poland and Munro, 1976; Hildner, 1977; and Jackson and Hildner, 1978) for the radial interval from 2 to 6 $R_\odot$. The fact that the simulated estimate exceeds the observed one, then, by less than an order of magnitude does not detract from what is considered good agreement because substantial ejected mass was undoubtedly present below the observational occulting disc (at 2 $R_\odot$) in the coronagraph.

The excess energy can similarly be computed. Figure 8 shows the build-up of thermal, kinetic, gravitational, and magnetic energies. We should point out that the fact that thermal energy dominates (with the kinetic energy rapidly overtaking the thermal portion) should not be construed as necessarily a general result. For example, as noted above, the role of the magnetic field is still not clear and requires, hopefully, in situ plasma and magnetic measurements. In the present case, then, the net energy addition by 1510 UT is 1.4×10^{27} erg km^{-1}, or 3.9×10^{32} erg after multiplication by an assumed 140,000 km thickness and a factor of 2 to account for symmetry about the equator. The portion of this total estimate for the kinetic and potential forms only, may be compared with the observed estimate (Hildner, 1977), again within the radial ranges noted above. These values are 1.6×10^{32} erg (simulated) as compared with 1.43×10^{31} erg (observed). This order of magnitude difference must be decreased substantially because of the unobserved energy excess below the occulting disc.

The Events of 5 September 1973. An analogous MHD simulation has been performed for the flare of 5 September 1973 which began at 1826 UT (Dryer and Maxwell, 1978). The observations are described in a series of papers prepared under the aegis of the NASA Skylab Solar Flare Workshop. The simulation utilizes the model described above, but several parameters were changed. For example, the coronal base density

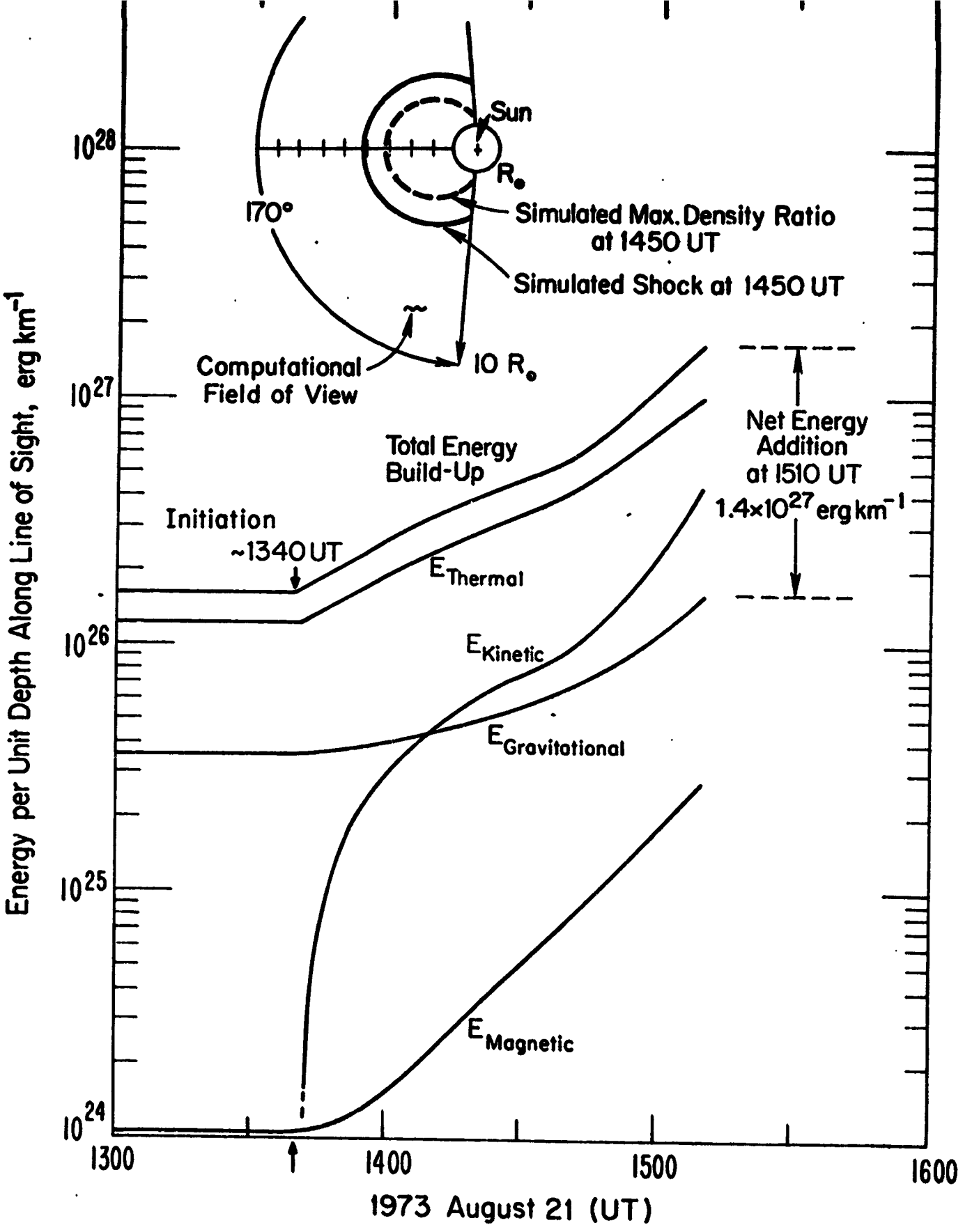


Figure 8. Simulated temporal energy addition to the corona and solar wind. As the total energy is transmitted to the solar wind outside the field of view, the value would decrease to the final equilibrium state which need not be the same as that existing prior to the "flare" at 1340 UT.

was increased to $n = 3 \times 10^6$ cm^{-3}, and the field magnitude was changed accordingly to B = 2.02 G in order to maintain $\beta = 1.0$ at the axis-of-symmetry and coronal base for the initial equilibrium atmosphere. Since soft X-ray observations of density and temperature were not available, Dryer and Maxwell (1978) utilized the short, impulsive radio burst as a suggestion that a 10 min thermal pulse (in a square wave fashion) would be appropriate. The temperature was increased by factors of 10, 20 and 40 in three separate calculations. The discussion here is limited to the case $\delta T = 40$, which produced a shock wave velocity (>1000 km s^{-1}) which compared more closely with that computed from the type II radio burst and a "10 x Saito" coronal density model. The temporal development of the loop structure, as indicated by the density ratio, is shown in Figure 9 (upper row); the temperature ratio, with anomalously-high local maxima near the contact surface, is shown in the lower row. Note that the radial and latitudinal scales are changed at t = 15 min.

Figure 10 shows the thermal pressure development for the same times as those shown (t = 2, 4, 6, 15 min) in Figure 9. The maximum pressure is located at the flare site during the pulse and changes to the vicinity of the contact surface after the flare ends. It is of particular interest to direct attention to the secondary pressure maximum at t = 2 min. This maximum is $(p/p_o) \approx 13$ at $1.11 < R < 1.13\ R_\odot$. The pressure increase which corresponds to this ratio is ~ 2.4 dy cm^{-2}. This simulated result compares favorably with an observed pressure increase of ~ 5.5 dy cm^{-2} inferred by Webb (1977) from intensity measurements of an X-ray filament above the flare site at $\sim$ 1827 UT.

The temporal development of the deformed magnetic field is shown in Figure 11. Conservation of magnetic flux requires that new, emerging flux move upward from below the coronal base as indicated by the dashed lines. Lines of constant vector potential are indicated by the computer-assigned constants which appear in the figure but which have no other physical significance. It is of interest to note: (i) a null is formed at the axis-of-symmetry and moves upward with

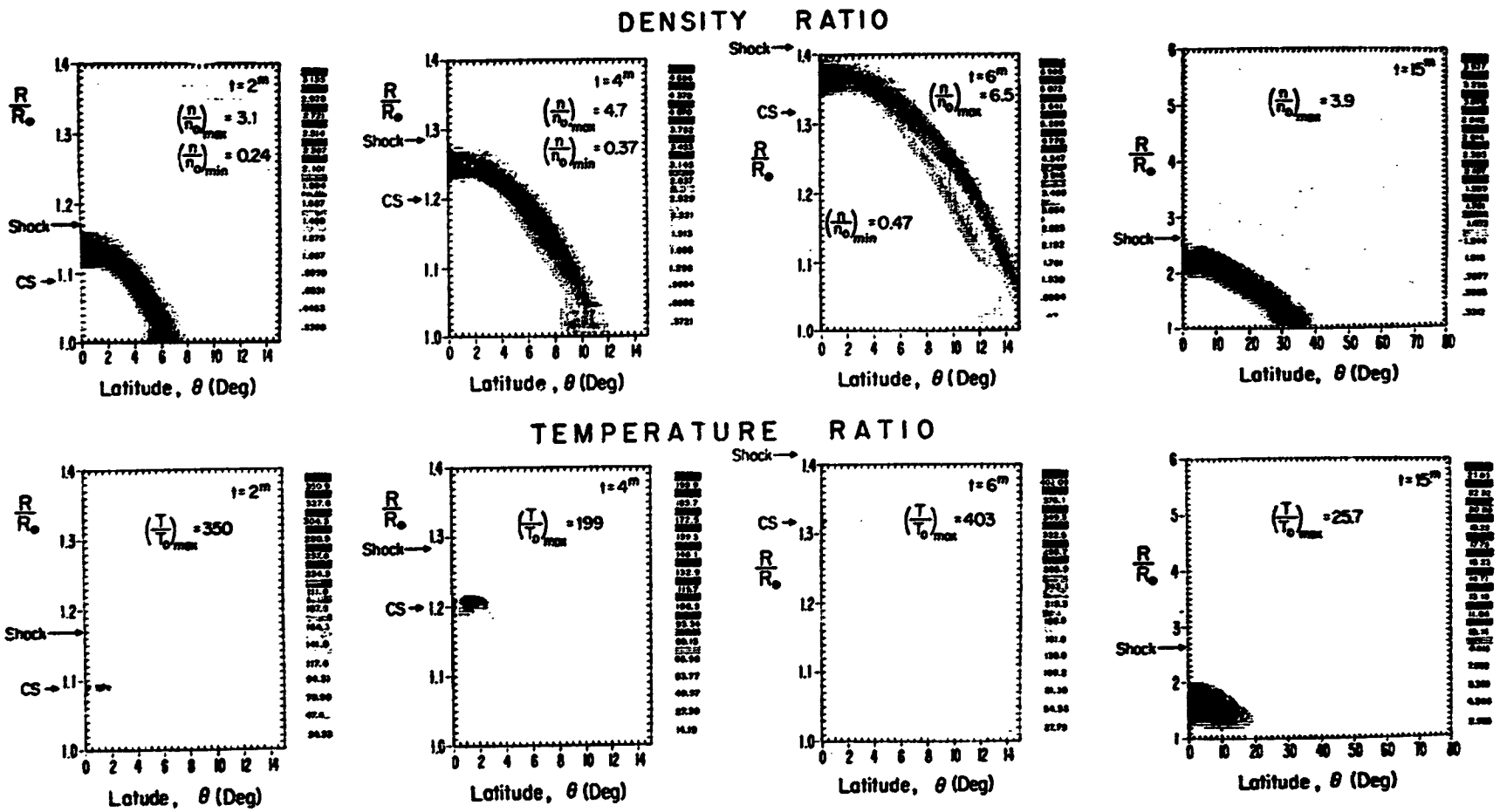


Fig. 9 Temporal development of the density and temperature ratios for the MHD simulation of the 5 September 1973 event.

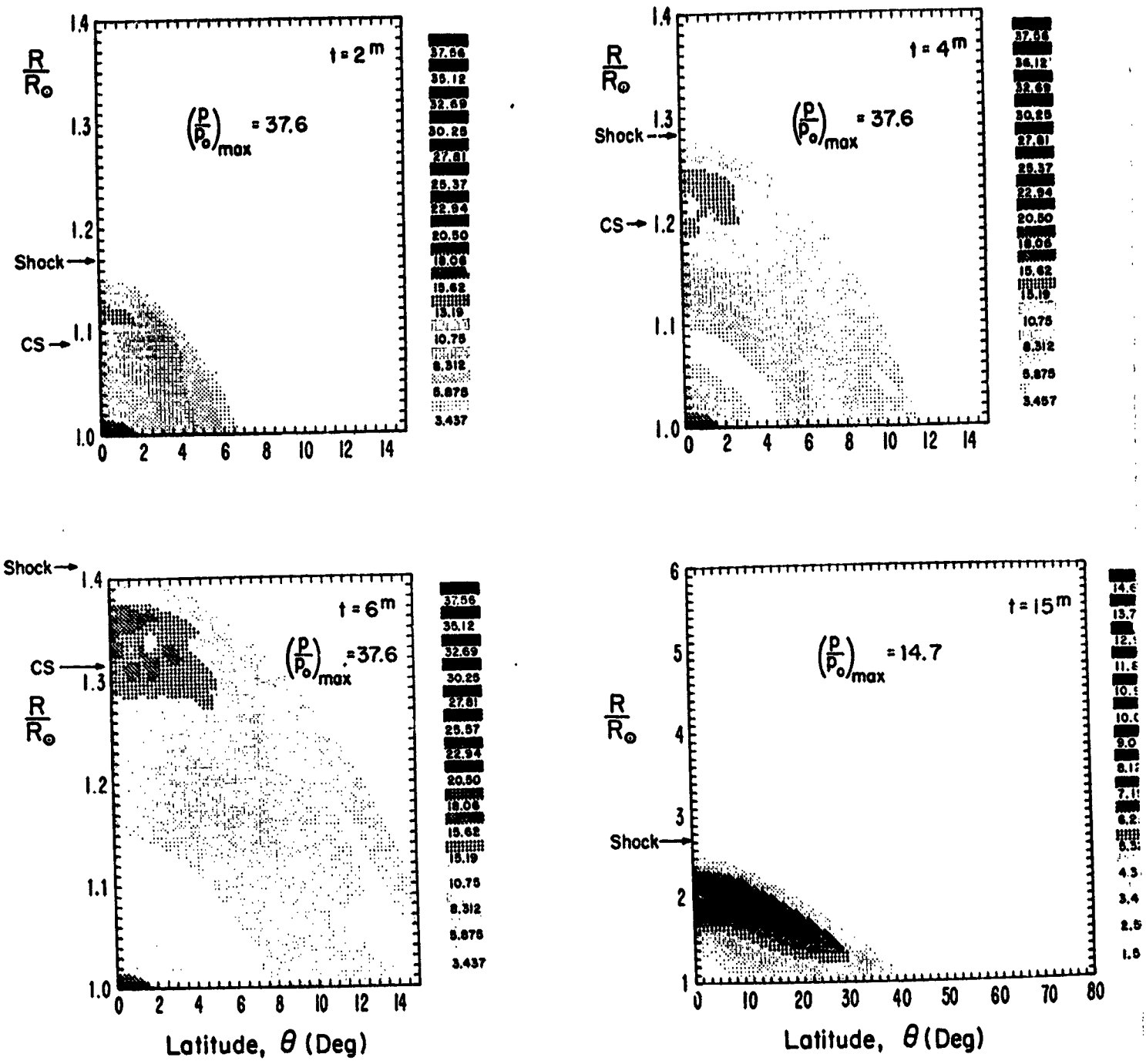


Fig. 10 Temporal development of the pressure ratio for the MHD simulation of the 5 September 1973 event.

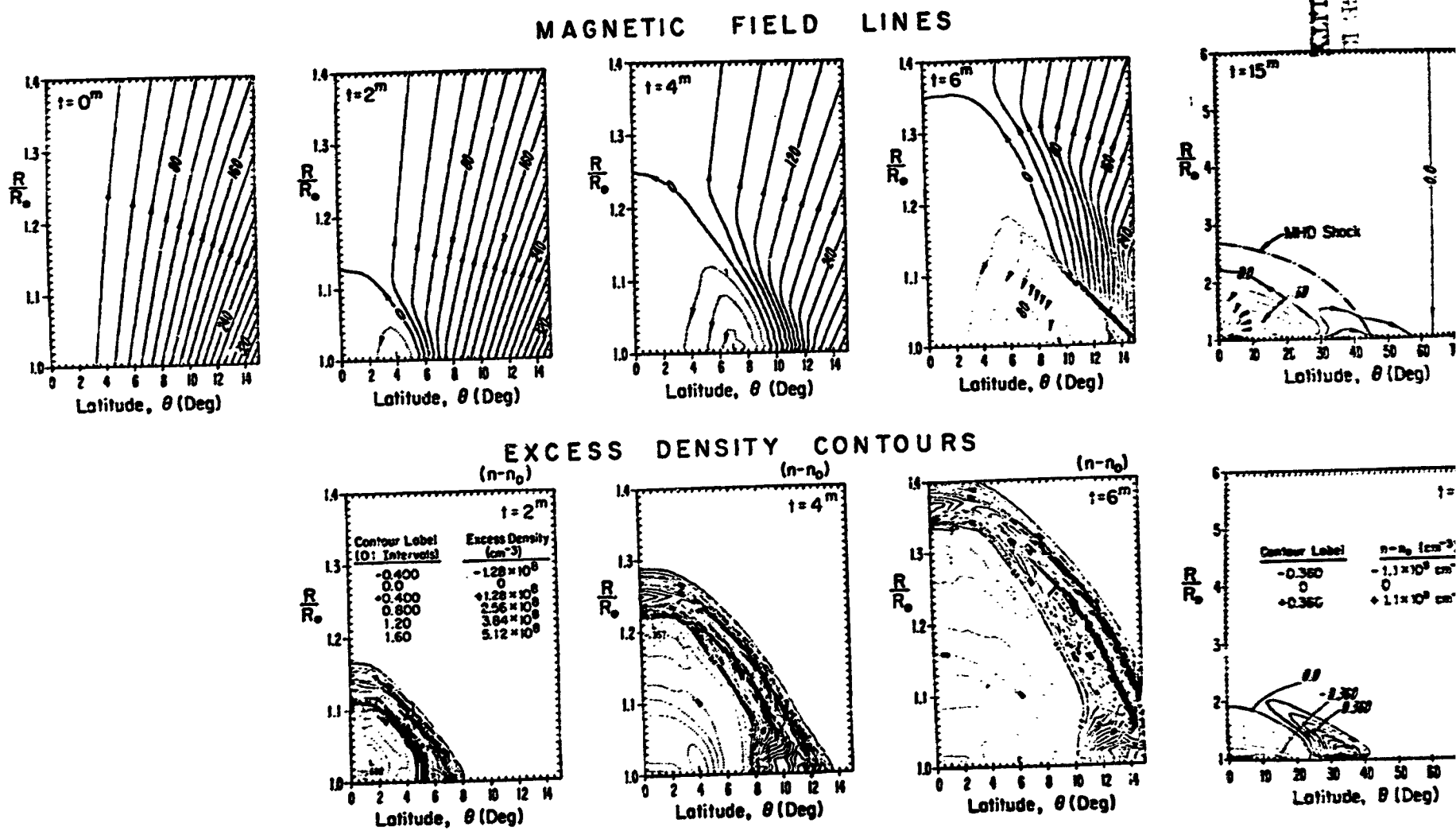


Fig. 11 Temporal development of the deformed magnetic topology and excess density contours for the MHD simulation of the 5 September 1973 event.

time; (ii) a region of strong magnetic field gradient is temporarily formed at t $\sim$ 6 min, indicating strong current formation; and (iii) the fast, forward MHD shock is properly formed as indicated by the appropriate increase of the transverse component across the quasi-parallel shock. At t = 15 min, the shock is observed to be moving laterally toward adjacent poles of the global hexapole.

Finally, Dryer and Maxwell obtained estimates for ejected mass and energy of 6.4×10^{16} g and 2×10^{32} erg, respectively.

CONCLUDING REMARKS

In order to establish a rational framework against which theory and experiment may confront each other, a planar, two-dimensional, time-dependent MHD model has been summarized within the context of two flare events. Although some limited comparisons are encouraging, much remains to be done for future studies of this kind. We have indicated the importance of the assumed, yet flexible, initial conditions and the type of pulse assumed to simulate actual physical events such as flares. It is impossible to overstate the necessity that computer simulation techniques be subjected to <u>in situ</u> verification of as many of the initial assumptions as possible. However, the MHD model's ability to provide a rational basis for improved physical understanding of mass ejections suggests its use as a planning and analysis tool for close encounter missions.

ACKNOWLEDGMENTS

We wish to express our gratitude to Drs. E. Hildner, B. V. Jackson, S. Martin, A. Maxwell, N. Metzler, S. McKenna-Lawlor, Y. Nakagawa, A. I. Poland, and D. Webb for many useful and stimulating discussions. We are grateful to the observers within this group for sharing their data with us in advance of publication. Our appreciation is also extended to the National Center for Atmospheric Research for the use of their CDC 7600 computer and graphics facilities. We also benefited from participation in the NASA Skylab Solar Flare Workshop, sponsored by the National Aeronautics and Space

Administration and the National Science Foundation and managed by the National Center for Atmospheric Research (High Altitude) Observatory). The work performed by STW and RSS was supported by NASA Marshall Space Flight Center under Contract No. NAS8-28097.

REFERENCES

Bruzek, A. and Demastus, H. L., 1970, Solar Phys., 12, 447

Dryer, M., Wu, S. T., Steinolfson, R. S and Wilson, R. M., 1978, submitted to Astrophys. J.

Dryer, M. and Maxwell, A., 1978, in preparation.

Dulk, G. A., Smerd, S. F., MacQueen, R. M. Gosling, J. T., Magun, A., Stewart, R.T., Sheridan, K. V., Robinson, R. D. and Jacques, S., 1976, Solar Phys., 49, 369.

Dulk, G. A. and McLean, D. J., 1978, submitted to Solar Phys.

Hildner, E., 1977, in Study of Travelling Interplanetary Phenomena/1977, (M. A. Shea, D. F. Smart and S. T. Wu, Eds.), pp. 3-21, D. Reidel Publ. Co., Dordrecht-Holland.

Jackson, B. V. and Hildner, E., 1978, submitted to Solar Phys.

Kiujpers, J., 1975, Solar Phys., 44, 173.

Martin, S. F. and Ramsey, H. E., 1972, in Solar Acitvity Observations and Predictions, (P. S. McIntosh and M. Dryer, Eds.), pp. 371-388, MIT press, Cambridge, Mass.

Poland, A. I. and Munro, R. H., 1976, Astrophys. J., 209, 927.

Steinolfson, R. S., Wu, S. T., Dryer, M. and Tandberg-Hanssen, E., 1978, Astrophys. J., in press.

Švestka, Z., 1976, Solar Flares, Vol. 8 of Geophys. and Astrophys. Monographs, (B. M. McCormac, Ed.), D. Reidel Publ. Co., Dordrecht-Holland, p. 128.

Webb, D., 1977, Report to Team 6 (Mass Motions) of the NASA Skylab Solar Flare Workshop, Boulder, Colo.

Wilson, R. M., 1977, NASA Technical Memorandum, X-73367, Marshall Space Flight Center, Alabama.

N78-32993

SCINTILLATION OBSERVATIONS NEAR THE SUN

W. A. Coles[+], B. J. Rickett[+], S. L. Scott*

[+]Applied Physics and Information Science Department,
University of California, San Diego, La Jolla, California

*Owens Valley Radio Observatory,
Big Pine, California 93513

Interplanetary scintillations have been used to probe regions of the interplanetary medium that have not been directly accessible to spacecraft. We present results on the electron density spectrum, the random velocity and the mean velocity of the solar wind in the region from 5 to 100 $R_\odot$. Our results are based on intensity scintillations of incoherent radio sources at different locations and different radio frequencies. The shape of the electron density irregularity spectrum is shown to be well modeled by a power law in wavenumber with a slope that abruptly steepens at higher wavenumbers. This two slope power law model is shown to have a "break" (defined as the wavenumber of the change of slope) that increases with decreasing distance from the Sun. The fractional random velocity, $\sigma_v/\bar{v}$, is shown to be insignificant at distances of greater than 40 $R_\odot$, but shows a steady increase with decreasing solar distance inside of 40 $R_\odot$.

The well-defined break in the spectrum of electron density fluctuations indicates plasma processes in the wavenumber range 0.01 to 0.4 km^{-1}. Plasma detectors on a space probe's close encounter with the Sun should be capable of spectral measurements to frequencies of at least 15 Hz to detect these phenomena.

Interplanetary scintillations have been used to probe regions of the interplanetary medium that have previously been inaccessible to spacecraft. We present results on the electron density fluctuation spectrum, the random velocity, and the mean velocity of the solar wind in the region from 5 to 100 solar radii. Our results are based on intensity scintillations of incoherent radio sources at different locations and different radio frequencies.

Our first result is the variation of the shape of the electron density fluctuation spectrum with heliographic distance and is shown in Figure 1. The different curves are the estimates of the 3-dimensional spatial spectrum of electron density fluctuations parameterized by distance from the Sun. From the shape of these spectra, it is clear that a power law of the form $\kappa^{-\alpha}$ or a power law with two distinct spectral indices is the inferred model for the spectrum. The double lines delineate the range of measured power law slopes while a dotted line indicates an upper or lower limit. Details of data analysis and model fitting for curves a-d may be found in Scott (1978). Curve e is taken from earlier work by Coles and Harmon (1978). The solar distances, spectral slopes, and the wave numbers of the breakpoints in the spectra are given in Table 1.

TABLE 1. SUMMARY OF DATA IN FIGURE 1

Curve	Radial distance solar radii	α Low wave number	α High wave number	Breakpoint km^{-1}
a	5 - 15	2.8 ± 0.2		
b	15 - 40	3.0 ± 0.4	≥ 5.0	0.1
c	40 - 70		5.0 ± 0.5	
d	60 - 100	≤ 2.5	3.7 ± 0.2	0.016
e	130 - 220	2.8	3.2	

Although a break is not observed in curve a or c, the two slopes are very different from each other, but are similar to the two slopes of curve b. It seems likely that we have not observed a break in the a and c curves

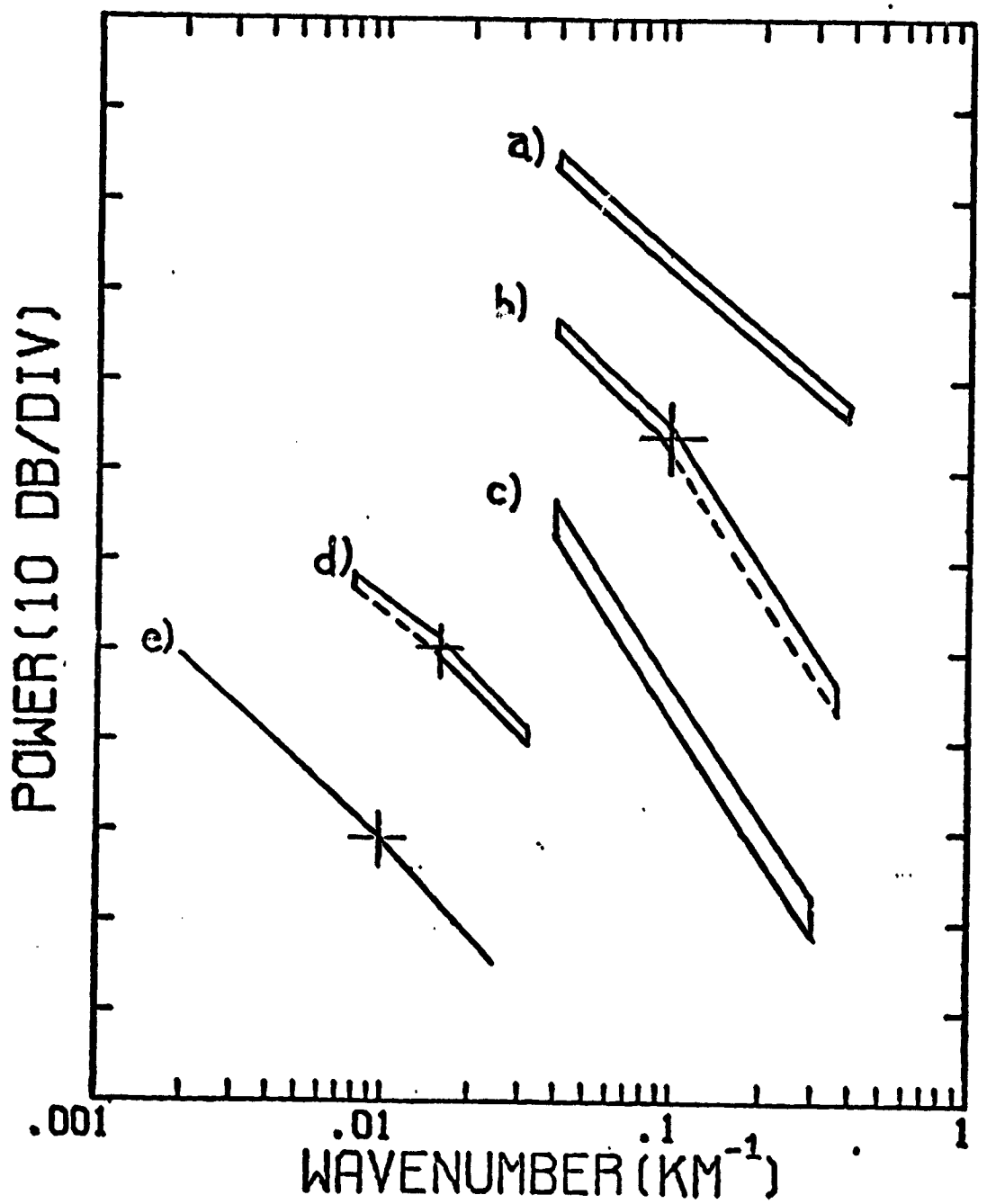


Figure 1. The shape of the density spectrum at different elongations (arbitrary vertical normalization). a) 5-15 solar radii; b) 15-40 solar radii; c) 40-70 solar radii; d) 60-100 solar radii; e) 130-220 solar radii.

because the break exists outside of the observed wavenumber range. Table 1 indicates that there may be an increase in the high-wavenumber α as the Sun is approached, but from our limited data this is not clearly proven.

Figure 2 shows the radial dependence of the wavenumber of the break in the spectral index. The arrows indicate upper or lower limits on the wavenumber of the break. If the wavenumber of the break is assumed to have a simple power law dependence on radial distance, the fit to the data gives a break wavenumber proportional to $r^{-1.2}$.

Estimates of the structure function at different solar distances made by Coles (1978) are shown in Figure 3. The apparent transverse structure function is related to the three dimensional density spectrum, $M3_{n_e}$, by

$$D(\vec{s}) \propto \int (1 - \cos(\vec{\kappa}\cdot\vec{s}))\, M3_{n_e}(\kappa_x, \kappa_y, \kappa_z = 0)\, d\kappa_x\, d\kappa_y$$

Comparison of our spectral estimate at 30 solar radii with the structure function estimate in Figure 3 are in excellent agreement. The radial dependence of the break inferred from the structure function yields a spectral break with a radial dependence of approximately $r^{-0.5}$. Although there is some difference between this value and the radial dependence of $r^{-1.2}$ inferred from Figure 2, the general concept of a break in the irregularity spectrum that moves to lower wavenumber with increasing heliographic distance is now well established. This well defined break indicates plasma processes in the wavenumber range 0.01 to 0.4 km^{-1}.

Our second result concerns the radial dependence of the fractional random velocity component. The diffraction pattern observed by interplanetary scintillations not only drifts across the Earth, but rearranges itself over time. This temporal evolution of the pattern manifests itself in a skewness of the observed temporal cross correlation of intensity at two spaced receivers. We have interpreted this observed skewness as being the result of the pattern being convected with a probability distribution of velocities. This distribution is represented by an rms or "random" velocity, and a mean or drift velocity. Figure 4 shows the radial dependence of the fractional random velocity ($\sigma_v/\bar{v}$) obtained from weak scattering observations. The data have been limited to weak

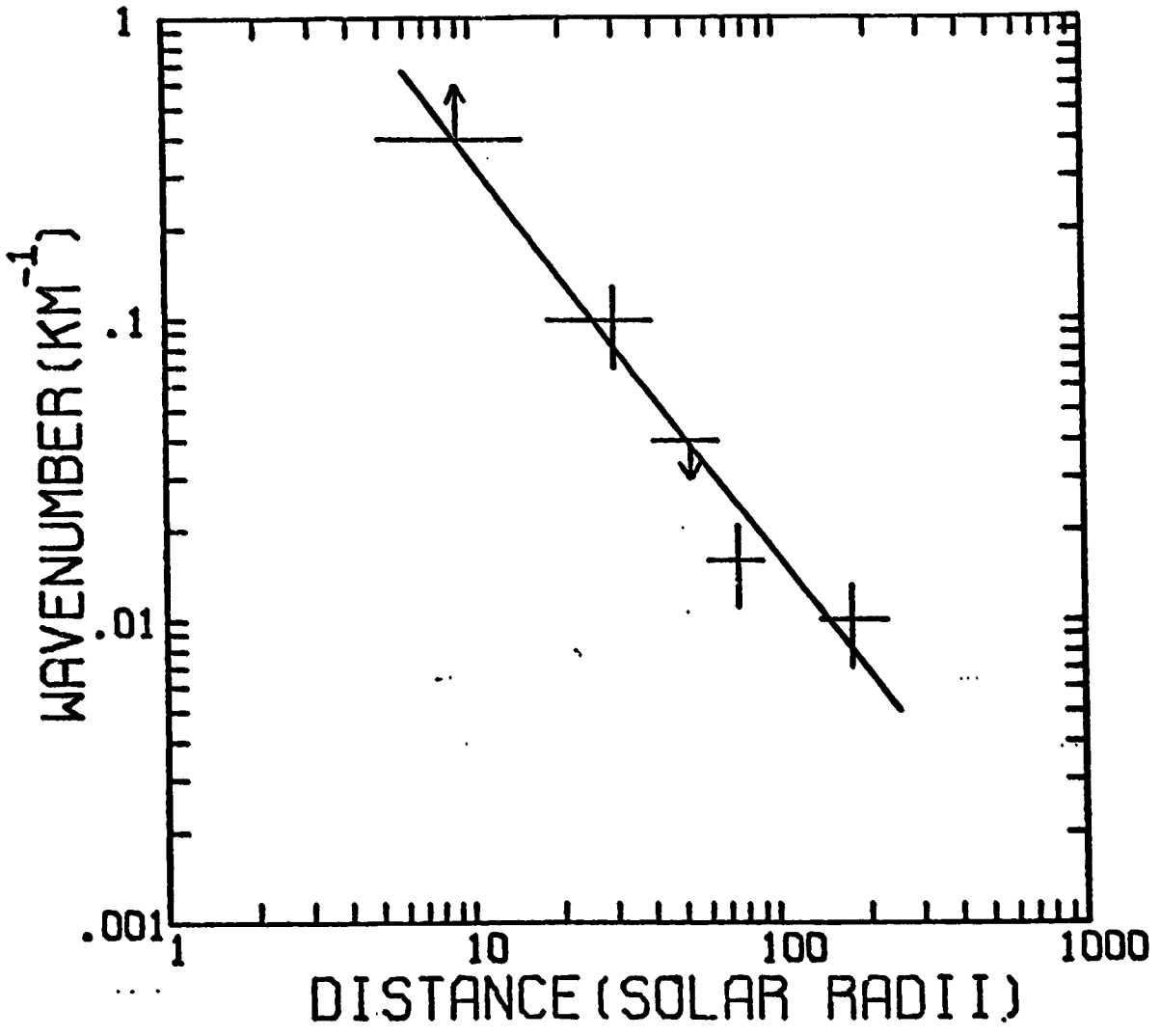


Figure 2. Radial dependence of the spectral break wavenumber.

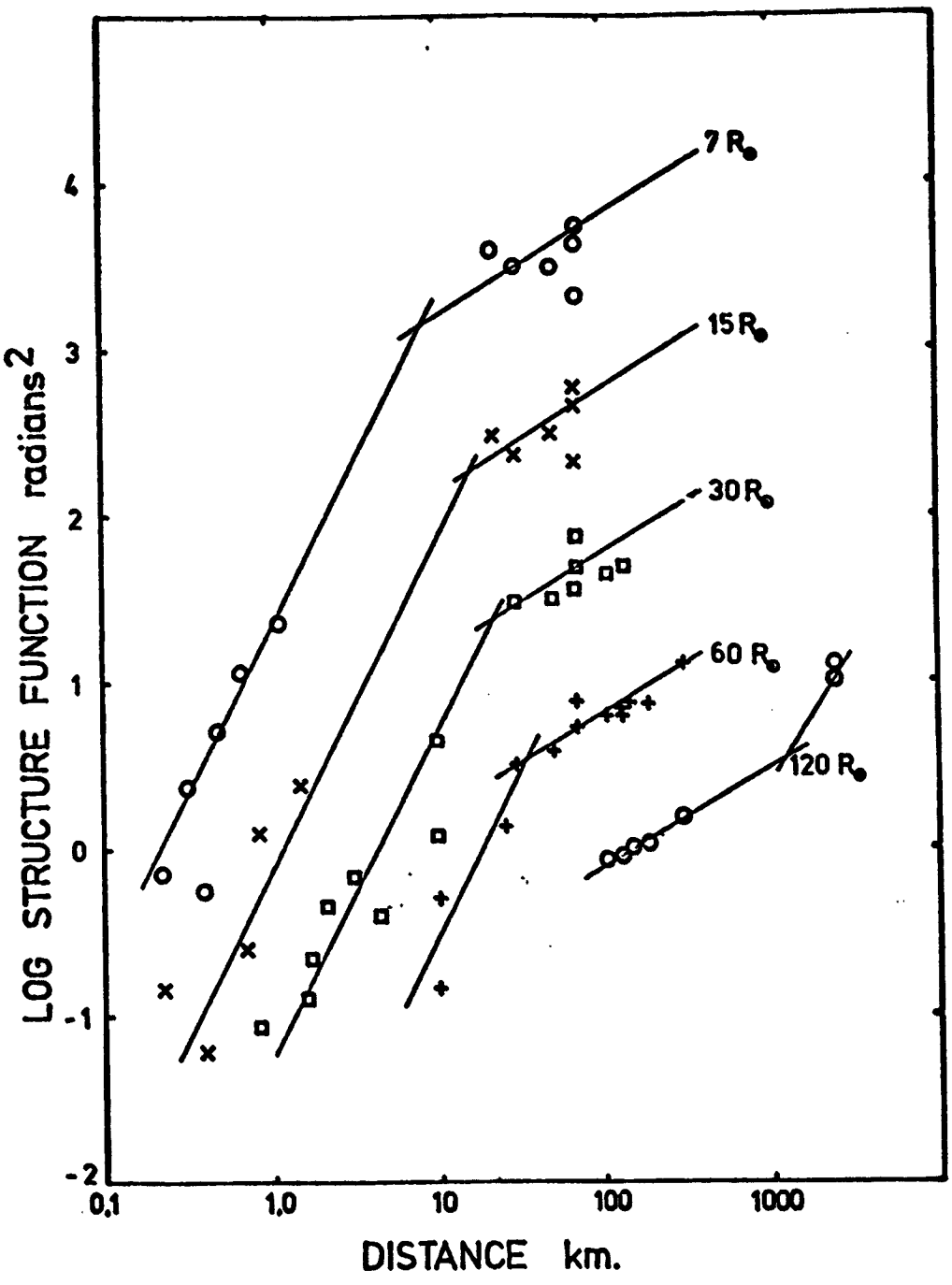


Figure 3. Structure function parameterized by elongation (from Coles, 1978).

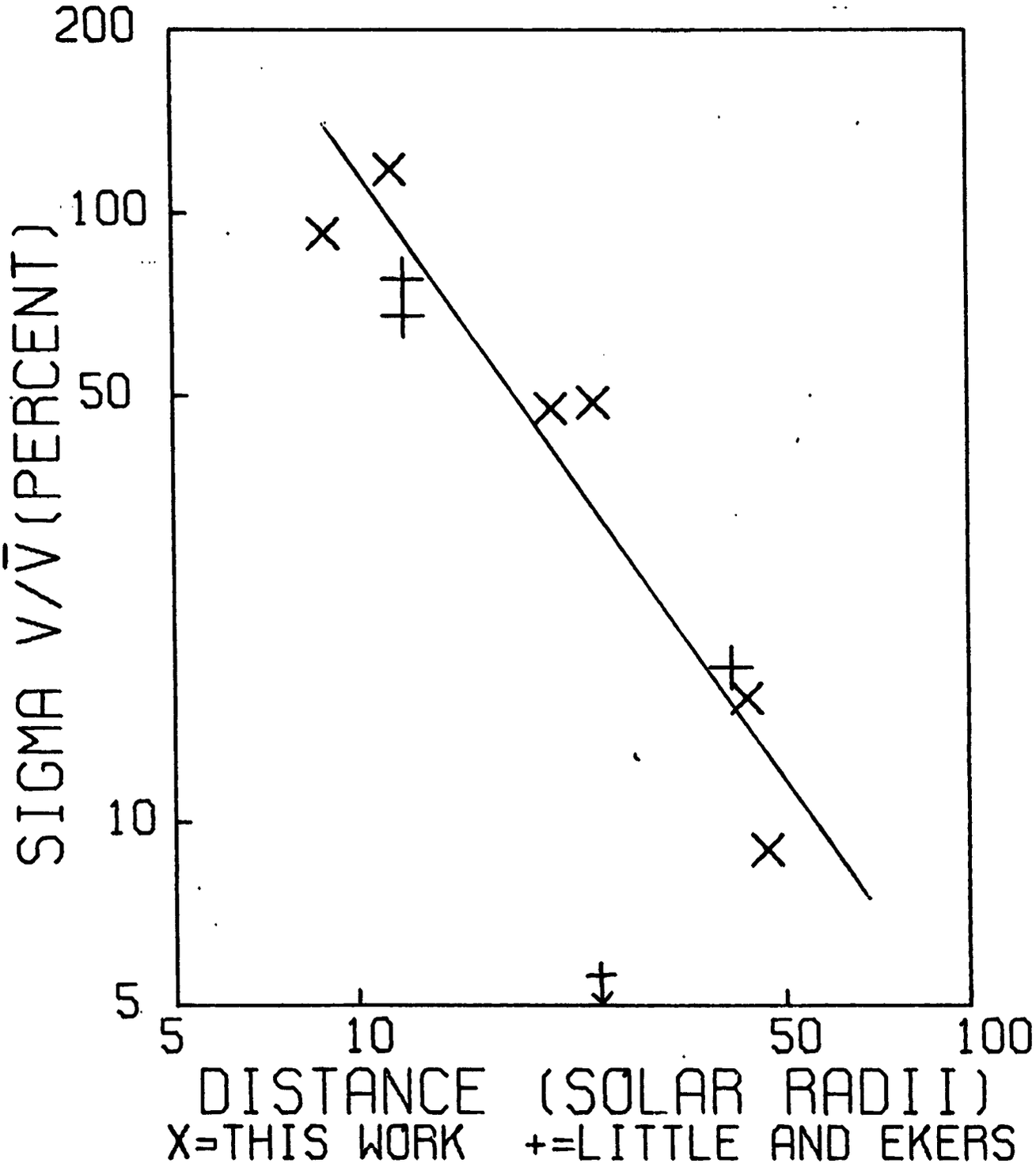


Figure 4. Radial dependence of fractional random velocity.

scattering because of the difficulty of relating the random velocity of the pattern to the random velocity of the medium in strong scattering. In this figure our observations have been combined with those of Ekers and Little (1971). The data point with the arrow represents an observation when no random velocity was observed. The best fit to the data neglecting this point gives a radial dependence of the fractional random velocity proportional to $r^{-1.4}$.

Measurements of the mean velocity with elongation are presented in Figure 5. The data point with the box indicates a measurement at high heliographic latitude ($\sim$57° N). Its high velocity may be the result of increasing velocity at high latitudes as previously reported (Coles and Rickett, 1976). The trend toward lower velocity closer to the Sun is evident. These data are in general agreement with other workers (Woo, 1978; Ekers and Little, 1971) and can be used as additional points in obtaining the acceleration curve of the solar wind.

A space probe mission en route to a close encounter with the Sun can potentially obtain further information on both the evolution of the density fluctuation spectrum with heliographic distance and the nature and causes of the large apparent random velocity inside 40 solar radii. Detection of the break in the spectrum that we have described here requires plasma detectors capable of spectral measurements to frequencies of at least 15 Hz.

References

Coles, W. A., 1978, Space Sci. Rev., 21, 411.

Coles, W. A., Harmon, J. K., 1978, J. Geophys. Res., in press.

Coles, W. A., Rickett, B. J., 1976, J. Geophys. Res., 81,

Ekers, R. D., Little, L. T., 1971, Astron. & Astrophys., 10, 310.

Scott, S. L., Ph.D. Thesis, University of California, San Diego, 1978.

Woo, R., 1978, Astrophys. J., 219, 727.

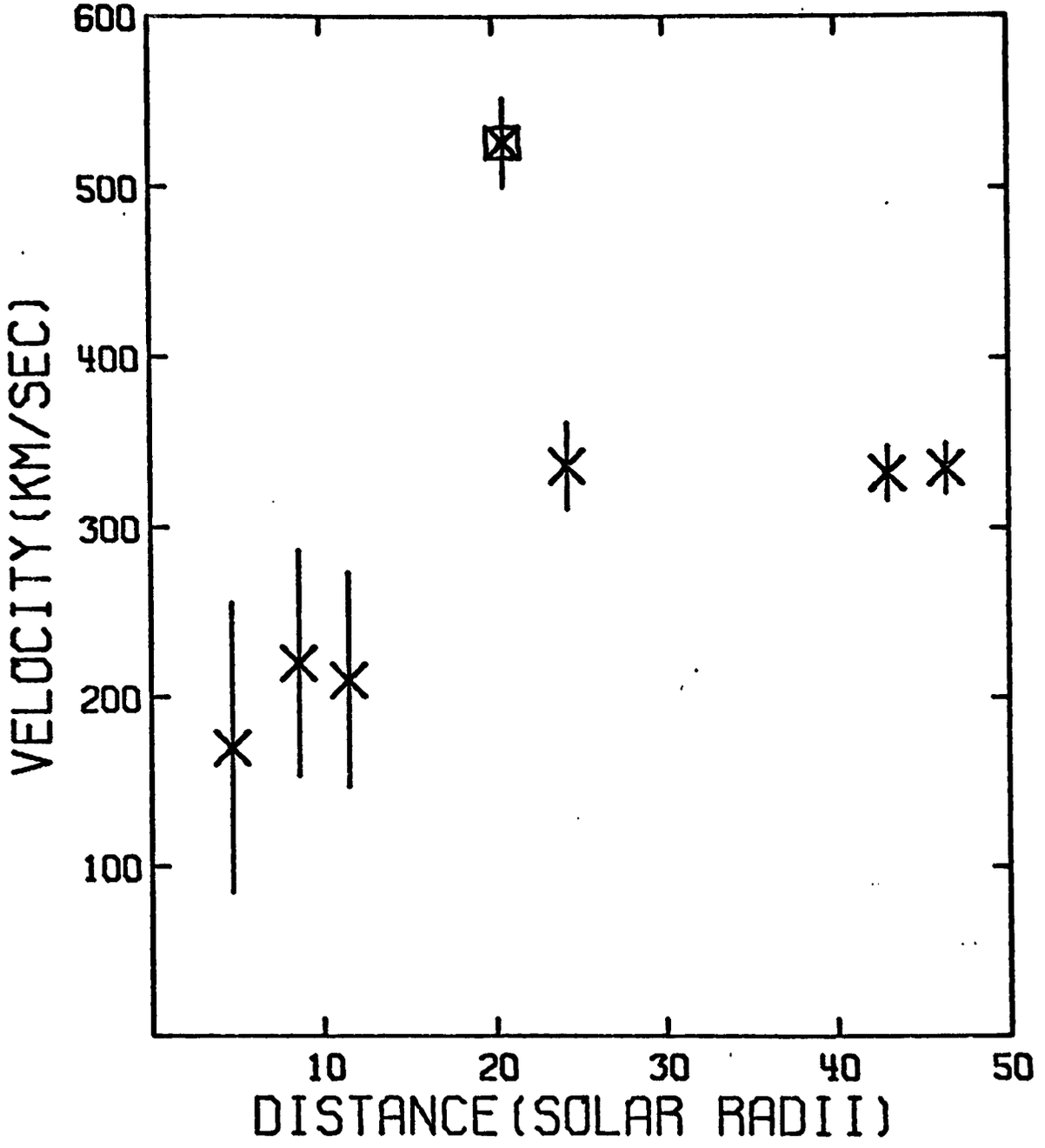


Figure 5. Mean velocity versus elongation.

PLASMA DENSITY MAPPING IN THE SOLAR WIND THROUGH USE OF VHF RADIO TO MEASURE ELECTRON CONTENT

Thomas A. Croft

SRI International
Menlo Park, CA 94025

ABSTRACT

We describe how the electron content between the Solar Probe and the Earth can be observed with a minimum of equipment and give a quantitative rationale for the use of a signal near 400 MHz to supplement the telecommunications signal. The emphasis is on the method of making content observations and on their value. Other investigators are expected to place instruments on the probe for in-situ plasma and magnetic field measurements.

While far from the Sun, during the cruise phase which forms the bulk of the mission duration, the electron content is so low that the S-X dual-frequency system is insufficiently sensitive and a UHF system is optimum. As the probe approaches the Sun, the UHF may be disrupted by scintillation and the variations of the telecommunications signal must be used for the content measurement. By operating the suggested system in different modes as the solar distance changes, we will be able to operate during the entire mission.

INTRODUCTION

The plasma densities of the solar wind and corona have been measured near the Sun by coronagraphs and far from the Sun by in-situ spacecraft, but the intermediate distances (which will be traversed by the Solar Probe) are comparatively sparsely measured. Yet these intermediate distances are a transition region in the macroscopic flow where control by the magnetic field gives way to control by dynamic forces. It will be productive to observe the flow patterns through on-board measurement of plasma density and velocity (as a minimum) coupled with measurement of probe-to-Earth electron content by radio methods. The correlative analysis of such a body of data will reveal the flow parameters both at the spacecraft and in the vicinity where the radio path is nearest the Sun.

When a radio signal travels through a plasma, its speed is reduced by a small amount that is proportional to the number density of the free electrons. It is feasible to measure the total added time required for a signal to transit from a spacecraft to Earth and thus to determine the total number of electrons along the path. This is often visualized as the total electron content (TEC) of a unit area cylinder extending from the transmitter to the receiver.

During the Solar Probe mission there are numerous opportunities to use TEC measurements to determine the density distribution of the solar wind and corona. From experience with other spacecraft under similar circumstances, it is clear that TEC measurements are useful primarily as one element in a correlative data set that includes the on-board measurement of local plasma density and velocity at the probe.

We therefore suggest that the Solar Probe should carry both instruments for measuring the plasma density and its velocity and the radio equipment necessary for determining electron content. Since the tentative plans for this probe already include an S-X system, it might seem that the suggested TEC instrumentation is included. However, a much lower frequency is needed for the Solar Probe mission because an S-X system is relatively insensitive to plasma.

THEORY OF ELECTRON CONTENT MEASUREMENT

The equations that relate radio delay to electron content along a radio path are very simple because we may neglect magnetic and collision terms in the Appleton-Hartree formula. The symbols used in the equations are as follows:

T = time
S = distance
s = distance increment
V = velocity
μ = refractive index
C = speed of light
N = electron concentration
I = electron content = $\int N ds$
F = radio frequency

Subscripts are as follows:

G = group
p = phase
H = the higher of two frequencies
L = the lower of two frequencies

In a plasma, the refractive index is μ, where

$$\mu^2 = 1 - \frac{80.6N}{F^2} \text{ (SI units)},$$

but at 50 MHz and higher in the magnetosphere or the solar wind,

$$\mu \approx 1 - \frac{40.3N}{F^2} \approx \frac{1}{1 + \frac{40.3N}{F^2}}$$

The group and phase delays are given by

$$T_G = \int \frac{ds}{V_G} = \frac{1}{C}\int \frac{ds}{\mu} = \frac{S}{C} + \frac{40.3I}{CF^2}$$

$$T_p = \int \frac{ds}{V_p} = \frac{1}{C}\int \mu ds = \frac{S}{C} - \frac{40.3I}{CF^2}$$

The difference in time delay for two frequencies F_L and F_H is

$$\Delta T_G = T_{GL} - T_{GH} = \frac{40.3I}{C}\left(\frac{1}{F_L^2} - \frac{1}{F_H^2}\right) = -T_{PL} + T_{PH} = -\Delta T_p$$

We may multiply by C to get meters; or we may multiply by F to get wavelengths. For F_L = 400 MHz, F_H = 2300 MHz, and an electron content of $10^{16}/m^2$, these parameters are 8.15 ns, 2.44 m, and 3.26 wavelengths at F_L. For the S-X system these numbers are much smaller: 0.23 ns, 0.07 m, and 0.54 wavelengths.

BRIEF BACKGROUND

Past Dual-Frequency Experiments (Pioneers 6-9 and Mariner 5)

Beginning with the launch of Pioneer 6 in December, 1965, a fairly continuous program of electron content measurement was undertaken using 50- and 423-MHz signals sent uplink to a series of five spacecraft. Pioneers 6, 8 and 9 still work well, but the ground transmitter is silenced. On all these spacecraft, the hardware, software, data processing, and analysis were in the hands of Stanford University and the Stanford Research Institute (now SRI International) forming a cooperative "Center for Radar Astronomy" (Eshleman et al., 1966). The receiver system placed aboard the spacecraft was designed and built by the Stanford Research Institute. The receiving and the on-board data processing systems survived well beyond the design lifetime. The flight hardware that might be used on the Solar Probe could be a direct outgrowth of that dual-frequency receiver system. The major change would be the elimination of one of the two receivers and an increase in the modulation frequency. The timing function performed by the higher of the two frequencies can instead be provided by precision oscillators maintained in a known time reference relative to DSN masers. In this manner, the uplink command signal could replace the reference signal that SRI used to send to the Pioneer spacecraft.

Present Dual-Frequency Experiment (S-X)

When the Deep Space Net (DSN) increased its radio frequency for tracking and telemetry functions, the higher frequency was introduced during an interval when both the old and new frequencies operated on the same spacecraft. During this transition period, it seemed logical to

capitalize on the availability of the DSN dual-frequency capability to perform the scientific measurements, without any need for the scientists to supply a signal. In this manner, the present S-X dual-frequency experiment has come into being. During the 1970s, radio-science measurements aboard spacecraft have ceased to be treated as "experiments" in the usual sense because no hardware has been supplied by the scientist for installation on the spacecraft.

Strengths and Weaknesses of the Past and Present Systems

The older system was very sensitive to plasma so that small percentage variations could be measured in the solar wind. In fact the main limitation on the accuracy of derived results was the imperfection of our knowledge of the electron content of the Earth's ionosphere, a quantity that must be subtracted from the TEC before the interplanetary content can be derived. The greatest weakness in the system, however, was the existence of only a single ground transmitter, resulting in intermittent data during each day. This weakness will probably not be overcome so long as it is necessary for one transmitter, supplied by the scientific team, to send two signals simultaneously. We have therefore devised a method which uses a precision oscillator instead of a second signal, so that a single-frequency transmitter could serve the spacecraft. In this manner, the intermittent character of the data can be avoided because it is feasible to use other existing transmitters at other longitudes.

The present system suffers mainly from the fact that there is so little time delay difference due to the plasma along the path--only 0.23 ns for a TEC of $10^{16}/m^2$. The experience with operation of this system on Viking and Voyager indicates that system uncertainties in determining the group delay difference between the S- and the X-band frequencies are not likely to be much less than 10 ns, from which we obtain a TEC error of some $40 \times 10^{16}/m^2$. For the Viking solar corona experiment, the standard error in the measurements was found to be about $70 \times 10^{16}/m^2$ (Tyler et al., 1977). To demonstrate why this is a serious problem for the Solar Probe, we examine the magnitude of the electron content that will be encountered during the mission.

TYPICAL VALUES OF TEC (AVERAGE SOLAR WIND DENSITY)

We assume that electron density varies with the inverse square of heliocentric distance and that it has a value of 7.6 electrons/cm^3 at the orbit of Earth. These two assumptions establish the density of the solar wind everywhere near the ecliptic plane. Using this model of the solar wind, we have determined the contours of constant TEC. The result is plotted in Figure 1, on which we overlaid the orbits of Jupiter and the inner planets. In this drawing, the Sun and Earth are stationary and the other planets move in their orbits. The electron content between any point and Earth is the value indicated by the contour intervals. For example, the content between Earth and the location where the Mars symbol is drawn would be about 300 x $10^{16}/m^2$.

We have translated the proposed trajectory of the Solar Probe into the coordinates of Figure 1. The probe first moves along large expanding spirals to the right, reaching Jupiter at about the point where the Jupiter symbol is drawn on this figure. After that, the spirals continue clockwise, but they become smaller until at last the probe dives toward the Sun. There are a total of three turns along the spiral and, as a result, there are three passages through superior conjunction. The first of these occurs on (approximately) day 300 when the spacecraft is close to the ecliptic plane. The second occurs on day 690 when the spacecraft is somewhat to the north, having recently passed by Jupiter. The third conjunction occurs on day 1060 when the spacecraft will be well to the north of the Sun as seen from Earth.

From this exercise we find that the TEC throughout the 3 years prior to passage by the Sun will be less than 200 x $10^{16}/m^2$ more than a half of the time and less than 500 x $10^{16}/m^2$ about 90 percent of the time. It is only during the first superior conjunction that the electron content is expected to rise to much more than 1000 x $10^{16}/m^2$.

With an error of 40 (and perhaps 70) x $10^{16}/m^2$, the measurement of electron content history with the S-X dual-frequency system would do little more than verify that the solar wind exists; it would not be sufficiently accurate to detect the small variations that provide the scientifically significant data.

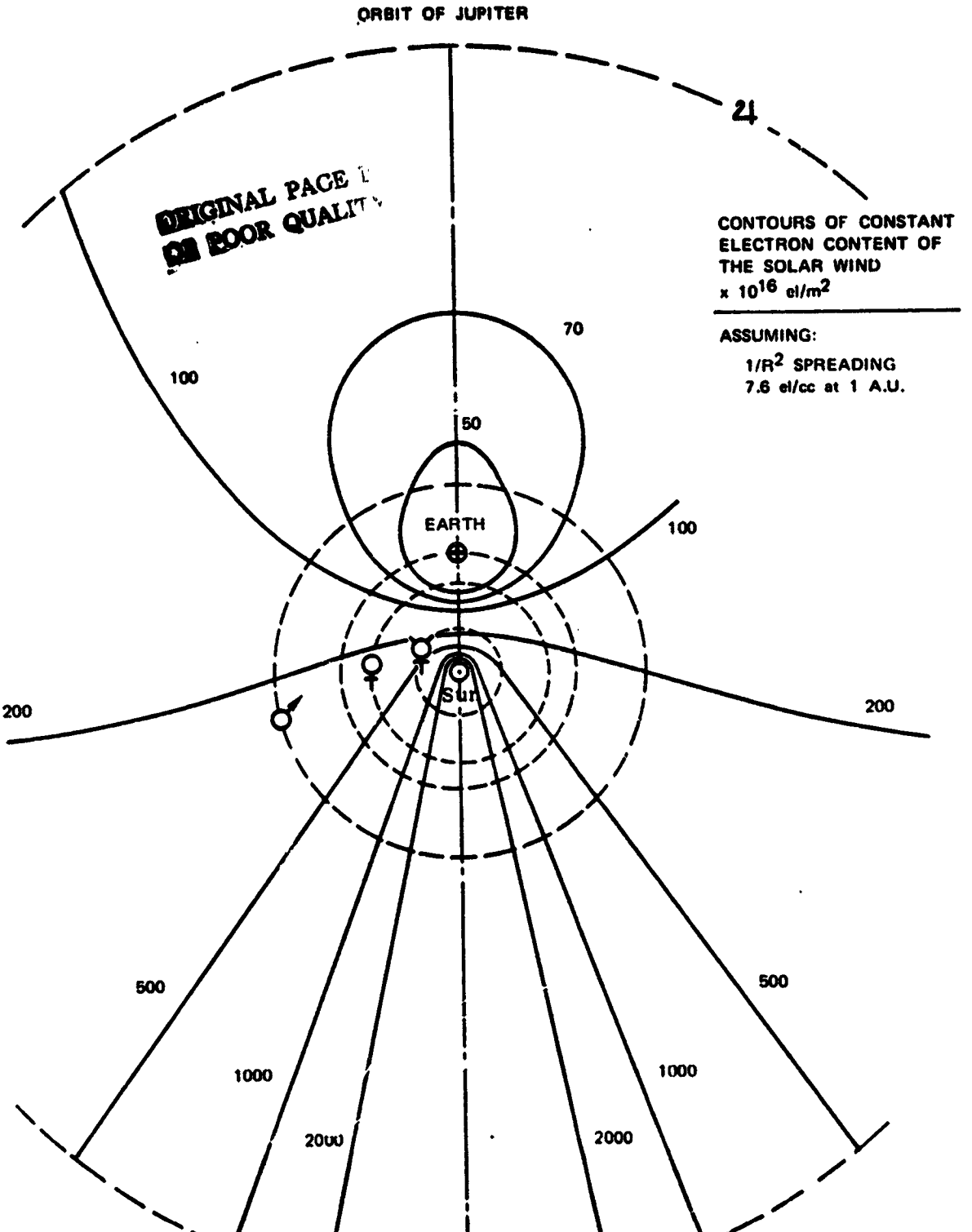


FIGURE 1 CHART OF THE SOLAR SYSTEM SHOWING THE AVERAGE ELECTRON CONTENT OF THE SOLAR WIND BETWEEN EARTH AND ANY POINT NEAR THE ECLIPTIC. For example, if content is measured from earth to a spacecraft located along the line labeled "200," that content would fluctuate about an average value of approximately 200 x 10^{16} el/m^2.

Th.. USE OF UHF OR VHF

The obvious way to improve sensitivity is to reduce the frequency, F_L, as indicated by the sensitivity curves shown in Figure 2.

The extreme sensitivity of the older dual-frequency experiments was due to the fact that F_L was only 50 MHz and the system was readily capable of measuring an electron content $1 \times 10^{16}/m^2$. For the Solar Probe these low frequencies would be disrupted while the probe is in the vicinity of the Sun during the prime portion of the mission. The optimum frequency for a Solar Probe would be near 400 MHz, with the exact frequency to be determined from a careful optimization study.

IMPACT ON SPACECRAFT DESIGN

There are several ways in which a VHF or UHF link could be implemented on a Solar Probe. We have performed a preliminary design study of the appraoch previously mentioned wherein a precision oscillator serves in place of the second signal. In this mode of operation, the radio science measurement aboard a spacecraft resumes its role as a typical "experiment" in the sense that the scientists supply a module for installation aboard the spacecraft. The main parameters of one module design are summarized in Table 1 for a fairly typical instrument consisting of a receiver and its associated data processing system with modest telemetry requirements. These weight, volume, power, and telemetry estimates are likely to be valid for any similar instrument which operates in an upling mode.

THE ION DRIVE OPTION

The implementation of ion drive might change the picture considerably. During a preceding study, we evaluated the potential for a downlink UHF signal from the spacecraft which uses the large power-generation capability needed for ion drive. During those times when the ion drive is not on, a powerful downlink signal might be generated with a reasonable amount of hardware, because the power could be used more wastefully than is typically the case. Although I do not wish to pursue this subject further here, I only state that if ion drive does become a serious contender for this mission, a downlink UHF signal should be considered because it offers a number of somewhat unexpected and valuable benefits.

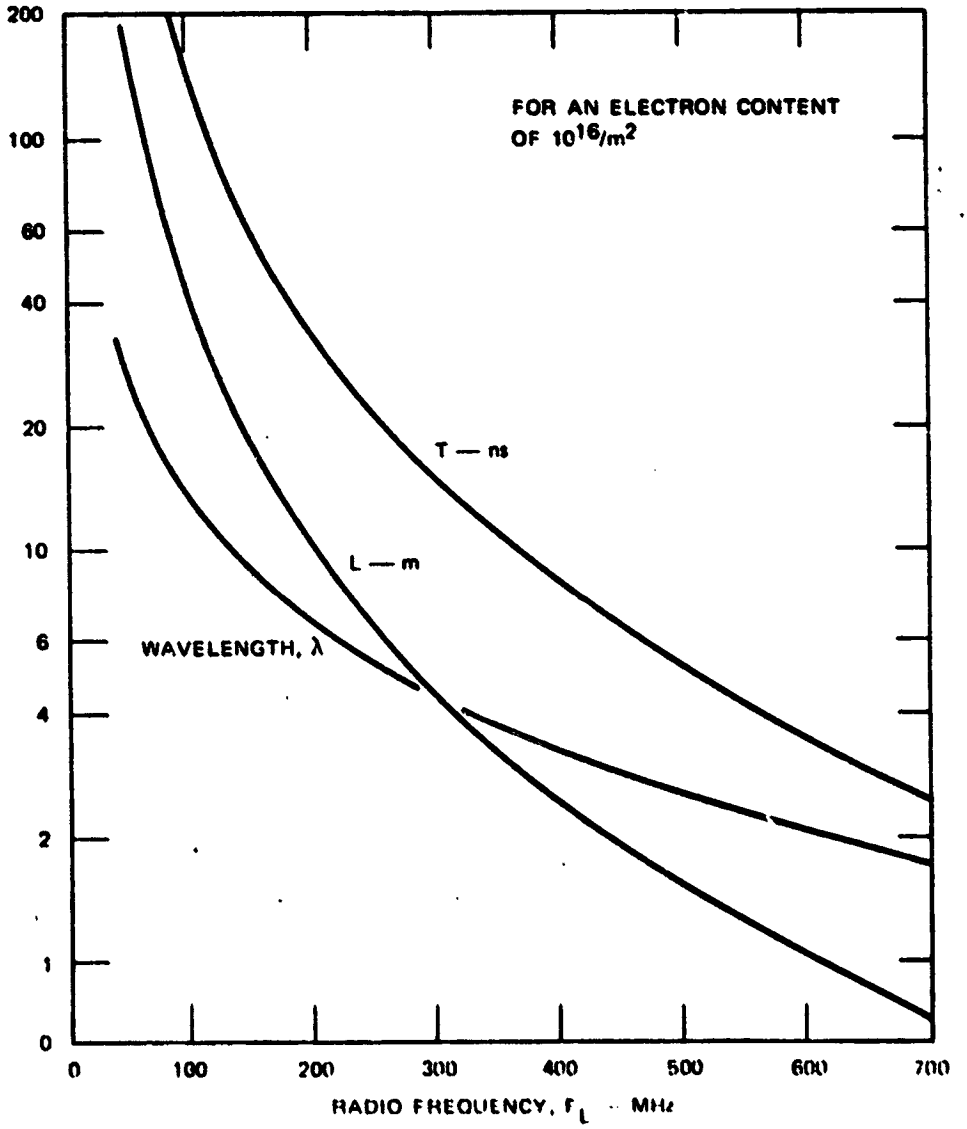


FIGURE 2 THE DELAY DIFFERENCE FOR TWO SIGNALS SENT THROUGH A REGION IN WHICH THE TEC IS $10^{16}/m^2$, EXPRESSED THREE WAYS AS A TIME DELAY, T; AS A SEPARATION DISTANCE, L, AND AS A SEPARATION DISTANCE MEASURED IN WAVELENGTHS AT THE LOWER FREQUENCY (useful in Doppler calculation)

Table 1

ONE REPRESENTATIVE DESIGN OF A RECEIVER SYSTEM CAPABLE OF MEASURING TEC

Pa Parameter	Value
Weight	<4 lb (1.8 kg)
Volume	≤216 inch^3
Dimensions	≤6 by 6 by 6 inches
Power required	<1.5 watts continuous, when on
Data rate	Four 8-bit data words each minute Three 8-bit housekeeping words each minute
Input signals	UHF radio signal; VCO voltage and modulation signal from S-band receiver; digital data control pulses; calibration signals
Output signals	Differential Doppler; modulation phase, UHF signal strength; S-band signal strength, three housekeeping signals
Command	On-off
EMI Constraints	UHF interference <<-145 dBm

SCIENTIFIC VALUE OF THE TEC MEASUREMENT

The value of total electron content measurements arises in several different regimes:

(1) During passage through the magnetosphere of Jupiter, TEC measurements offer an avenue for studying the plasma distribution, not just the density at the spacecraft.

(2) During the passage by the Sun, the electron content can be derived from the Doppler variations of the telecommunication signal interpolated between the content as measured before and after this event using the dual-frequency system. It is possible that the UHF uplink system would continue to work during the solar passage because of the fortunate choice of a trajectory that is on the near side of the Sun.

(3) During the bulk of the cruise of the Solar Probe, this instrument (coupled with plasma density and velocity measurements made on-board) will provide excellent diagnostic information concerning the solar wind flow regimes.

(4) During the solar conjunctions, and particularly during the second and third conjunction when the spacecraft is at northerly latitudes, the electron content can be interpreted to infer the plasma density profiles over the north pole of the Sun. Each new measurement of these profiles will be of unique value because the profiles change continuously from day to day and through the solar cycle.

REFERENCES

Eshleman, V. R., O. K. Garriott, R. L. Leadabrand, and A. M. Peterson, "The Interplanetary Electron Number Density from Preliminary Analysis of the Pioneer 6 Radio Propagation Experiment," J. Geophys. Res., Vol. 71, No. 13, pp. 3324-3327 (July 1, 1966).

Tyler, G. L., J. P. Brenkle, T. A. Komarek, and A. I. Zygielbaum, "The Viking Solar Corona Experiment," J. Geophys. Res., Vol. 82, No. 28, pp. 4335-4340 (September 30, 1977).

SECTION VII.

INTERPLANETARY DUST

INTENTIONALLY

D33

WHAT IS THE FATE OF INTERPLANETARY DUST?

Martha S. Hanner
Jet Propulsion Laboratory

The Solar Probe will provide a unique opportunity to observe the evolution of dust particles near the sun.

Space investigations in recent years have led to the following general picture of the interplanetary dust:

Size distribution: The size distribution near 1 AU has been derived from microcraters on lunar rock samples between 10^{-3} and 10^{-16} gm [1,2,3] A bimodal distribution is evident, suggesting that the particles smaller than 10^{-12} gm represent a separate component. The mass concentration has a maximum near 10^{-5} gm [4]. Based on this size distribution, the scattered sunlight observed as zodiacal light originates mainly from particles in the range 10^{-5} to 10^{-7} gm [5,6] .

Spatial distribution: The zodiacal light photometers on Helios 1 and 2 recorded a brightness gradient corresponding to a radial distribution of the dust, $n(r) \propto r^{-\nu}$, $\nu = 1.3 \pm 0.1$, valid from 0.1 to $\gtrsim 1$ AU [7,8] . No evidence for a dust-free zone was found, to a limiting solar distance of 19 solar radii [8]. The direct dust detector carried on Helios generally confirms this radial gradient [9] . The zodiacal light experiment on Pioneers 10 and 11 derived a radial gradient $\nu = 1.5 \pm 0.5$ from 1 to 3 AU [10] . On the other hand, a radial distribution increasing outward from 0.7 to 2.5 AU was derived for radio meteors, mass 10^{-3} to 10^{-4} gm [11] . The half-width of the dust cloud perpendicular to the ecliptic is a few tenths AU. The off-ecliptic dust distribution will be mapped from the Solar Polar mission [12] .

Composition: Interplanetary dust samples have been collected in the earth's stratosphere and analyzed [13]. Their composition indicates primitive solar system material, generally similar in elemental abundances to carbonaceous chondrites, although more porous, fine-grained, and complex in structure than the meteorites. Basic grain size within the particles is 0.1 μm or smaller. The polarization of the zodiacal light can be approximated by fluffy, absorbing particles [14] consistent with the appearance of the collected samples.

Dynamics: Impact detectors on Pioneers 8 and 9, in solar orbit between 0.75 and 1.09 AU, detected two dust populations [15] . Particles with mass $\geq 10^{-12}$ gm were generally in direct Keplerian orbit, with low velocities ~10 km/sec relative to the detectors. Small particles, called β meteoroids, with $m < 10^{-12}$ gm, were observed moving outward from the solar direction with velocities 50 - 100 km/sec. These two components have been confirmed by the Helios dust detectors [9] and support the bimodal flux curve derived from the lunar microcraters.

Origin: Whipple [16] has estimated that an average production rate on the order of 10 tons/sec is needed in order to maintain the interplanetary dust cloud. Dohnanyi [17] has concluded that asteroidal debris is not a major source. Comets represent the most probable origin of the dust [16, 18, 19]. A few orbits of impacting dust particles have been directly associated with comets [20] . One particle detected by Pioneer 9 was identified as interstellar [21] . However, present-day comets are not an adequate source; much of the dust released by comets does not remain gravitationally bound to the solar system [22,23,24] . It seems that either an unusually large, bright comet is captured every $\sim 10^4$ years (perhaps Encke?) or more dust is emitted from long-period comets in the form of large particles in bound orbits than now estimated [16,18,26]. The interplanetary dust cloud may indeed be variable, with the present epoch representing a high density.

The Solar Probe mission provides a unique opportunity to study the evolution of the interplanetary dust and its eventual destruction near the sun. Understanding the solar interaction with the dust has implications beyond the solar system in understanding the evolution of circumstellar dust.

Dust particles in the inner solar system have short lifetimes in cosmic terms. A dust grain of a few microns radius will spiral inward from 1 AU to the sun under the influence of Poynting-Robertson drag in about 10^4 years [25,16] . Near the sun, two destructive processes become important. Collisions may fragment the grains into submicron components

which can be blown out of the solar system. The grains which survive collisions will vaporize near the sun. The relative importance of these two processes is not yet clearly established, nor is the role of electromagnetic forces well understood.

Zook and Berg have computed the rate of collisional breakup of micrometeoroids [27]. The collisional rate will be proportional to the square of the number density and the average impacting velocity,

$$C' \propto n^2 v$$

If $n(r) \propto r^{-1.3}$, as observed from Helios, and $v \propto r^{-0.5}$, then

$$C' \propto r^{-3.1}$$

The number of collisions in a spherical shell of radius r, thickness dr, will be

$$C' \propto r^{-1.1}\, dr$$

There is evidence that $n(r) \propto r^{-1.3}$ extends well within 0.1 AU [8]. Thus, most collisions will take place at small solar distances, inside 0.3 AU, the limit directly observable by the dust detector on Helios.

One must then ask whether the collisions will result in particle breakup. Zook and Berg conclude that if $\log (M/M_{min}) \sim 3$, where M_{min} is the minimum mass of a colliding particle which can fragment a particle of mass M, most particles of mass $\geq 10^{-5}$ gm will be fragmented between 1 AU and 0.1 AU. Even if $\log(M/M_{min})$ is ~ 2, the probability of fragmentation is ~ 0.5 for these larger particles. The submicron particle fragments will be rapidly blown outward by radiation pressure. Zook and Berg conclude that collisional fragmentation is sufficient to account for the flux of beta meteoroids observed by Pioneers 8 and 9. Their conclusion is independently supported by Dohnanyi [28].

Evidence that dust grains are being vaporized close to the sun comes from infrared observations of the corona. Eclipse observations in November 1966 by MacQueen [29] at 2.2 μ and Peterson [30] at 2.2 μ and 3.5 μ show a prominent emission peak near 4 $R_\odot$ and a shallower one at 3.3 - 3.5 $R_\odot$ superimposed on a smooth brightness variation with solar distance. The 4 $R_\odot$ peak at 2.2 μ was subsequently

confirmed from balloon observations and new peaks detected at 8.7 and 9.2 $R_\odot$ [29] . Lena et al [32] identified peaks at 4 $R_\odot$ and 6.4 $R_\odot$ in the 10μ band during the June 1973 eclipse. Peterson [31] derived a grain temperature of 2160 ± 200 K for the peak at 4 $R_\odot$ from observations at 0.84μ, 1.57μ, and 2.23μ during the March 1970 eclipse. This high temperature would indicate a refractory material such as graphite, rather than a silicate material. However, spectral scans by Lena et al. [32] across the 10μ band gave a peak at 9.5μ, corresponding to silicate, and a minimum near 10.8μ, which would indicate a lack of opaque material such as graphite.

To interpret these observations, one must consider not only the grain temperature and sublimation rate as a function of solar distance and grain composition, but also the forces affecting the particle motion, which are a function of particle size and composition [33-38]. It is convenient to define β, the ratio of the radiation pressure force to the gravitational force, since both vary inversely with the square of solar distance

$$\beta = \frac{F_{rad}}{F_{grav}} = \frac{3 R_o^2}{4c G M_\odot} \cdot \frac{1}{\rho a} \int_o^\infty Q_{pr}(m,a,\lambda)\, F_\odot(\lambda)\, d\lambda$$

where R_o = 1 AU, Q_{pr} is the efficiency factor for radiation pressure computed by Mie theory as a function of refractive index m, particle radius a, and wavelength λ, and $F_\odot(\lambda)$ is the monochromatic solar flux. Thus, β varies inversely with particle radius and density ρ, modified by the variation in Q_{pr} with particle size. For particles of micron size and larger, β is negligible. It reaches a maximum near a = 0.2μm, for both absorbing and dielectric materials [34,38]; $\beta_{max} > 1$ for iron and magnetite (i.e. for absorbing materials) and $0.4 < \beta_{max} < 1$ for silicates (olivine, obsidian, andesite). For silicate grains, a < 0.1μ, β decreases rapidly; however for small magnetite grains, β remains larger than 0.5 [38]. In principle, it is not necessary to have $\beta > 1$ for a particle to leave the solar system, but only necessary that the total energy of the particle be positive, or $\beta > 0.5$ [27].

The fate of a grain (expulsion versus complete vaporization) will depend on its rate of sublimation compared to the rate of change of orbital parameters. For example, Lamy predicts that iron particles would only reach 24 $R_\odot$, then would spiral rapidly outward whereas obsidian grains would penetrate to 2-3 $R_\odot$ and vaporize completely [34]. The relative dwell time of the grains at different solar distances will be reflected in the relative concentration, which can be measured from the solar probe. Lamy predicts that grains with $a < 0.2\,\mu m$ do not have a significant concentration in the inner corona, a prediction which can be verified from the solar probe. When one considers that actual grains have a complex mineralogy, compared to the simple homogeneous spheres treated theoretically, it is clear that their evolution near the sun is not simple to predict. For example, the addition of a few per cent free carbon to silicate material (similar to collected dust samples) drastically lowers the visible albedo, increasing the absorption of solar radiation [39].

The fundamental measurements necessary to characterize the fragmentation and vaporization processes are the mass and velocity distributions of the particles as a function of heliocentric distance. The type of impact ionization detectors already flown on HEOS and Helios and planned for the Galileo and Solar Polar missions can be adapted for the solar probe [40]. The instrument measures the plasma generated by an impacting dust particle. From the pulse height and rise time of the signal, particle mass and velocity can be determined. Sensitivity is sufficient to detect sub-micron particles. In the Helios instrument, time-of-flight spectrum of the positive ions was also measured, to determine composition. For adaptation to the solar probe, calibration at higher velocities is required. Improvement in the accuracy of mass and velocity determination would also be desirable. In order to observe the angular distribution of particle velocities, several sensors, mounted at different orientations, are required.

The impact detector will provide statistics on the velocity

distribution primarily for sub-micron and micron-sized particles. The velocity distribution of the larger particles ($\sim 10\mu$ - 100μ range) can be mapped by measuring the Doppler shift of the Fraunhofer lines in the zodiacal light. Such an experiment has not yet been flown on a space probe. A suitable instrument for this purpose has now been developed [41]; similar instruments, with sufficient sensitivity and resolution, are being used in the Atmospheric Explorer series. This experiment would provide new information on the average velocity distribution as a function of heliocentric distance throughout the inner solar system.

More comprehensive infrared observations of the solar corona are also needed; however these can best be carried out from earth orbit. Improved spectral coverage, to derive the temperature of the emitting grains as a function of solar distance and to map the detailed structure across the silicate bands are particularly important. The spatial structure and temporal constancy of the emission peaks also need study. Such observations should be considered as important support for the solar probe mission.

REFERENCES

[1] Hörz, F., Brownlee, D.E., Fechtig, H., Hartung, J.B., Morrison, D.A., Neukum, G., Schneider, E., Vedder, J.F., and Gault, D.E., 1975. "Lunar Microcraters:Implications for the Micrometeoroid Complex", Planet. Space Sci. 23, 151.

[2] Fechtig, H., Hartung, J.B., Nagel, K., Neukum, G., and Storzer, D., 1974. "Lunar Microcrater Studies, derived Meteoroid Fluxes, and Comparisons with Satellite-borne Experiments", 5th Lunar Sci. Conf., Vol. 3, p. 2463.

[3] Fechtig, H., 1976. "In-Situ Records of Interplanetary Dust Particles - Methods and Results", Lecture Notes in Physics, Vol. 48: "Interplanetary Dust and Zodiacal Light", Proc. of IAU-Colloquium No. 31, June 1975, Edited H. Elsässer und H. Fechtig, Springer Verlag, Berlin-Heidelberg-New York, p. 143.

[4] Giese, R.H., and Grün, E., 1976. "The Compatibility of Recent Micrometeoroid Flux Curves with Observations and Models of the Zodiacal Light", Proc. IAU Colloquium No. 31, op. cit., p. 135.

[5] Röser, S., and Staude, H.J., 1978. "The Zodical Light from 1500 Å to 60 Microns: Mie Scattering and Thermal Emission", Astron. and Astrophys., in press.

[6] Millman, P.M., 1976. "Meteors and Interplanetary Dust", Proc. IAU Colloquium No. 31, op. cit., p. 359.

[7] Link, H., Leinert, C., Pitz, E., and Salm, N., 1976. "Preliminary Results of the Helios A Zodiacal Light Experiment", Proc. IAU Colloquium No. 31, op. cit., p. 24.

[8] Leinert, C., Hanner, M., Link, H., and Pitz, E., 1978. "Search for a Dust Free Zone Around the Sun from the Helios 1 Solar Probe", Astron. and Astrophys. 64, 119.

[9] Grün, E., Fechtig, H., Kissel, J., and Gammelin, P., 1977. "Micrometeoroid Data from the First Two Orbits of Helios 1", Jour. Geophys. 42, 717.

[10] Hanner, M.S., Sparrow, J.G., Weinberg, J.L., Beeson, D.E., 1976. "Pioneer 10 Observations of Zodiacal Light Brightness Near the Ecliptic: Changes with Heliocentric Distance", IAU Colloquium No. 31, op. cit., p. 29.

[11] Southworth, R.B., and Sekanina, Z., 1973. "Physical and Dynamical Studies of Meteors", NASA Report CR-2316.

[12] Fechtig, H., Giese, R.H., Hanner, M.S., and Zook, H.A., 1976. "Investigation of Interplanetary Dust from Out-of-Ecliptic Space Probes", The Study of the Sun and Interplanetary Medium in Three Dimensions, GSFC May 1975, ed. L.A. Fisk and W.I. Axford, p. 298.

[13] Brownlee, D.E., Tomandl, D.A., and Olszewski, E., 1977. "Interplanetary Dust: A New Source of Extraterrestrial Material for Laboratory Studies", Proc. 8th Lunar Sci. Conf., p. 149. and

Brownlee, D.E., Rajan, R.S., and Tomandl, D.A., 1977. "A Chemical and Textural Comparison between Carbonaceous Chondrites and Interplanetary Dust", in Comets Asteroids, Meteorites, IAU Colloquium No. 39, August 1976, ed. A.H. Delsemme, p. 137.

[14] Giese, R.H., Weiss, K., Zerull, R.H., and Ono, T., 1978. "Large Fluffy Particles: A Possible Explanation of the Optical Properties of Interplanetary Dust", Astron. and Astrophys. 65, p. 265.

[15] Berg, O.E., and Grün, E., 1973. "Evidence of Hyperbolic Cosmic Dust Particles", Space Research XIII, 1047.

[16] Whipple, F.L., 1967. "On Maintaining the Meteoritic Complex", in the Zodiacal Light and the Interplanetary Medium, ed. J.L. Weinberg, NASA SP-150, p. 409.

[17] Dohnanyi, J.S., 1976. "Sources of Interplanetary Dust: Asteroids" IAU Colloquium No. 31, op. cit. p. 187.

[18] Whipple, F.L., 1955. "A Comet Model III. the Zodiacal Light, Ap. J. 121, 750.

[19] Delsemme, A.H., 1976. "Can Comets be the only Source of Interplanetary Dust?", IAU Colloquium No. 31, op. cit., p. 481.

[20] Grün, E., Kissel, J., and Hoffmann, H.J., 1976. "Dust Emission from Comet Kohoutek (1973f) at Large Distances from the Sun" IAU Colloquium No. 31, op. cit., p. 334.

[21] Wolf, H., Rhee, J., Berg, O.E., 1976. "Orbital Elements of Dust Particles Intercepted by Pioneers 8 and 9", IAU Colloquium No. 31, op. cit., p. 165.

[22] Harwit, M., 1963. "Origins of the Zodiacal Dust Cloud", JGR 68, 2171.

[23] Röser, S., 1976. "Can Short Period Comets Maintain the Zodical Cloud?", IAU Colloquium No. 31, op. cit. p. 319.

[24] Delsemme, A.H., 1976. "The Production Rate of Dust by Comets", IAU Colloquium No. 31, op. cit., p. 314.

[25] Wyatt, S.P., and Whipple, F.L., 1950. "The Poynting-Robertson Effect on Meteor Orbits", Ap. J., 111, 134.

[26] Sekanina, Z. and Schuster, H.E., 1978. "Meteoroids from Periodic Comet D'Arrest and Interplanetary Dust", Astron. and Astrophys. 65, 29.

[27] Zook, H.A., and Berg, O.E., 1975. "A Source for Hyperbolic Cosmic Dust Particles", Plan. Space Sci. 23, 183.

[28] Dohnanyi, J.S., 1976. "Flux of Hyperbolic Meteoroids", IAU Colloquium No. 31, op. cit., p. 170.

[29] MacQueen, R.M., 1968. "Infrared Observations of the Outer Solar Corona", Ap. J., 154, 1059.

[30] Peterson, A.W., 1969. "The Coronal Brightness at 2.3 Microns", Ap. J. 155, 1009.

[31] Peterson, A.W., 1971. "A Determination of the Vaporization Temperature of Circumsolar Dust at 4 $R_\odot$", Bull. AAS, 3, 500.

[32] Léna, P., Viala, Y., Hall, D., and Soufflot, A., 1974. "The Thermal Emission of the Dust Corona during the Eclipse of June 30, 1973 II. Photometric and Spectral Observations", Astron. and Astrophys. 37, 81.

[33] Kaiser, C.E., 1970. "The Thermal Emission of the F-Corona" Ap. J. 159, 77.

[34] Lamy, Ph. L., 1974. "The Dynamics of Circum-Solar Dust Grains", Astron. and Astrophys. 33, 191.

[35] Lamy, Ph. L., 1974. "Interaction of Interplanetary Dust Grains with the Solar Radiation Field", Astron. and Astrophys. 35, 197.

[36] Mukai, T., Yamamoto, T., Hasegawa, H. Fujiwara, A., and Koike, C., 1974. "On the Circumsolar Grain Materials", Publ. Astron. Soc. Japan 26, 445.

[37] Mukai, T., and Mukai, S., 1973. "Temperature and Motion of the Grains in Interplanetary Space", Publ. Astron. Soc. Japan, 25, 481.

[38] Schwehm, G., and Rohde, M., 1977. "Dynamical Effects on Circum-Solar Dust Grains", Jour. Geophys. 42, 727.

[39] Johnson, T.V., and Fanale, F.P., 1973. "Optical Properties of Carbonaceous Chondrites and Their Relationship to Asteroids", Jour. Geophys. Res., 78, 8507.

[40] Dietzel, H., Eichhorn, G., Fechtig, H., Grün, E., Hoffmann, H.J., and Kissel, J., 1973. "The HEOS 2 and Helios Micrometeoriod Experiments", Jour. Phys. E: Scientific Instr. 6, 209.

[41] Torr, D. G., Young, E., Torr, M. R., and Nagy, A. F., 1978. "Zodiacal Light Dynamics Experiment", this Volume.

D34

N78-32998

ZODIACAL LIGHT DYNAMICS EXPERIMENT

(A Wideband Imaging Fabry-Perot Interferometer)

D. G. Torr, E. Young, M. R. Torr, A. F. Nagy

Space Physics Research Laboratory
University of Michigan
Ann Arbor, Michigan 48109

ABSTRACT

The Solar Probe will provide an ideal platform from which to study dynamics of dust particles near the Sun by measuring the detailed character of the Fraunhofer structure of the zodiacal light. The suggested instrument is a wideband imaging Fabry-Perot interferometer with state-of-the-art technology in both the optics and the detector.

The instrument would function as a high-resolution imaging device providing wavelength resolution of 0.03 Å over about a 20 Å range. It would have sensitivity comparable with conventional photometers. The wideband imaging capability would provide sky maps of the zodiacal light on a despun spacecraft without mechanical scanning. The total mass of this instrument is about 4 kg; power comsumption is approximately 4 watts. The sensitivity is 10^{-2} S10 units.

The Solar Probe mission would allow the velocity distribution of the dust to be mapped along most of the trajectory of the spacecraft.

INTRODUCTION

In the previous paper, M. Hanner presented the scientific rationale for the measurement of the velocity distribution of interplanetary dust. It is our belief that such a measurement would make a valuable contribution to knowledge of the physics of the solar system by providing direct experimental evidence of the forces acting on the dust particles. Detailed information on the velocity distribution of dust particles as a function of position with respect to the Sun is contained in the Doppler signature of the Fraunhofer profiles of sunlight scattered from the dust particles. The development of an instrument capable of measuring the detailed character of the Fraunhofer structure lies well within the reach of modern optical and detector technology. We discuss below the requirements that should be met in the design of instrumentation to achieve this end.

MEASUREMENTS

The measurement of the detailed spectral signature of the zodiacal light, which varies from 10^{-2} to 5×10^{3} S10 units over the mission, requires significantly greater sensitivity, dynamic range and spectral resolution than has been achieved with instrumentation flown previously. In the next section we discuss techniques by which this can be achieved on the Solar Probe. The resultant spectra must have low noise, i.e., less than 0.1% of the continuum values. These requirements demand an instrument of high resolution and throughput.

Figure 1 shows a typical example of Fraunhofer absorption lines in the wavelength interval 5000-5020 Å, (Beckers, Bridges and Gilliam, 1976). This spectrum illustrates that the amplitudes of the deepest absorption feature are about 80% of the average continuum value (900 W m^{-2} μ^{-1}) and that the line widths range typically from 0.1 to a few angstroms. Deep lines which might be selected for observation are $\sim$ 0.2 Å broad. We expect Doppler shifts in the order of $\pm$ 0.3 Å as

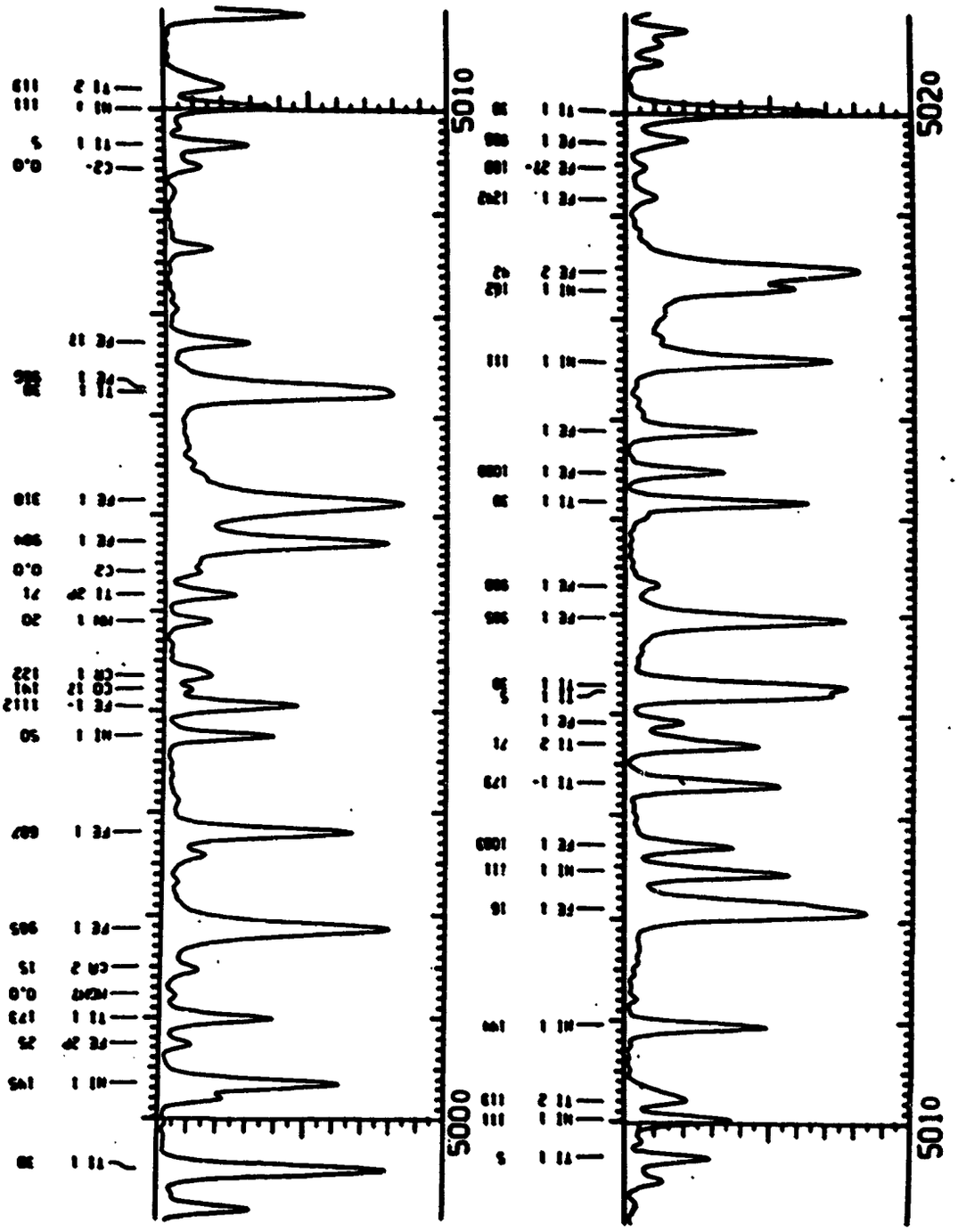


Figure 1. Fraunhofer Absorption Lines.

well as considerable line broadening due to variations in velocity along the line of sight. We therefore select a resolution of $\sim$ 0.03 - 0.05 Å which will facilitate deconvolution of the instrumental and Doppler broadening effects.

INSTRUMENT

A Wideband Imaging Fabry-Perot Interferometer

A Fabry-Perot interferometer will provide adequate resolution and throughput to meet the objectives discussed above. An instrument of this type is currently being built at the University of Michigan by P. B. Hays for flight on the Dynamics Explorer satellites. This instrument was designed primarily to measure the Doppler shift and broadening of airglow emissions originating in the Earth's daytime atmosphere. The design therefore incorporates a large baffle system for protection against scattered sunlight. In the case of the Solar Probe the main measurements of the mission would be made with the heat shield acting as a sunshade. Therefore, provided the instrument is suitably located to minimize stray light scattered off the spacecraft and associated instrumentation, the baffle system could be largely dispensed with, and the field of view of the instrument increased correspondingly. The advantages of enlarging the field of view of the Fabry-Perot are not obvious and require some explanation. The main advantages of a large field of view would be the elimination or reduction of mechanical scanning to acquire spatial coverage and a great enhancement of the data acquisition rate. The sky could be imaged on an area array detector.

In the conventional Fabry-Perot mode of operation, a small field of view is used and the spacing of the etalon plates is adjusted to scan through the desired spectral feature. In an imaging mode of operation the variation in angle of incidence across the field of view provides

the wavelength scan. The technique requires the isolation of a single order of the etalon. Therefore the passband of the filter must be matched to that of the etalon at all angles as illustrated in Figure 2. An interferometer with these characteristics has been built and successfully operated by one of us (Eugene Young) at the University of Washington in Seattle. In this mode of operation the Fabry-Perot provides a wavelength scan in the radial direction across the detector and a spatial scan in azimuth. This means that different spectral features are imaged from different parts of the sky. However, since a Doppler measurement can be carried out on any suitable spectral absorption feature, the convolution of spatial and spectral information does not matter. The solar Fraunhofer absorption spectrum is rich with features in the vicinity of 5000 Å so that at least one line per Å may be expected, thus providing about 20 or more spatial segments for a field of view of about 10°. The spectrographic and spatial imaging capability would also mean that all data would be acquired simultaneously with significant enhancement of the data acquisition rate or effective throughput of the instrument compared to a mechanical scanning device.

Figure 3 depicts a three dimensional viewing mode to illustrate the imaging concept. In this mode, the conical mirror gives a meridional scan while the central-axis hole provides a view perpendicular to the meridian. The instrument could be oriented and baffled so that neither of the two fields of view shown come within 50° of the Sun. Thus operation could be maintained even when the heat shield is oriented away from the Sun.

Figure 4 schematically illustrates the pattern projected on the detector. A suitable multiplier would be an ITT proximity focussed image intensifier fiberoptically coupled to several CCD area array detectors.

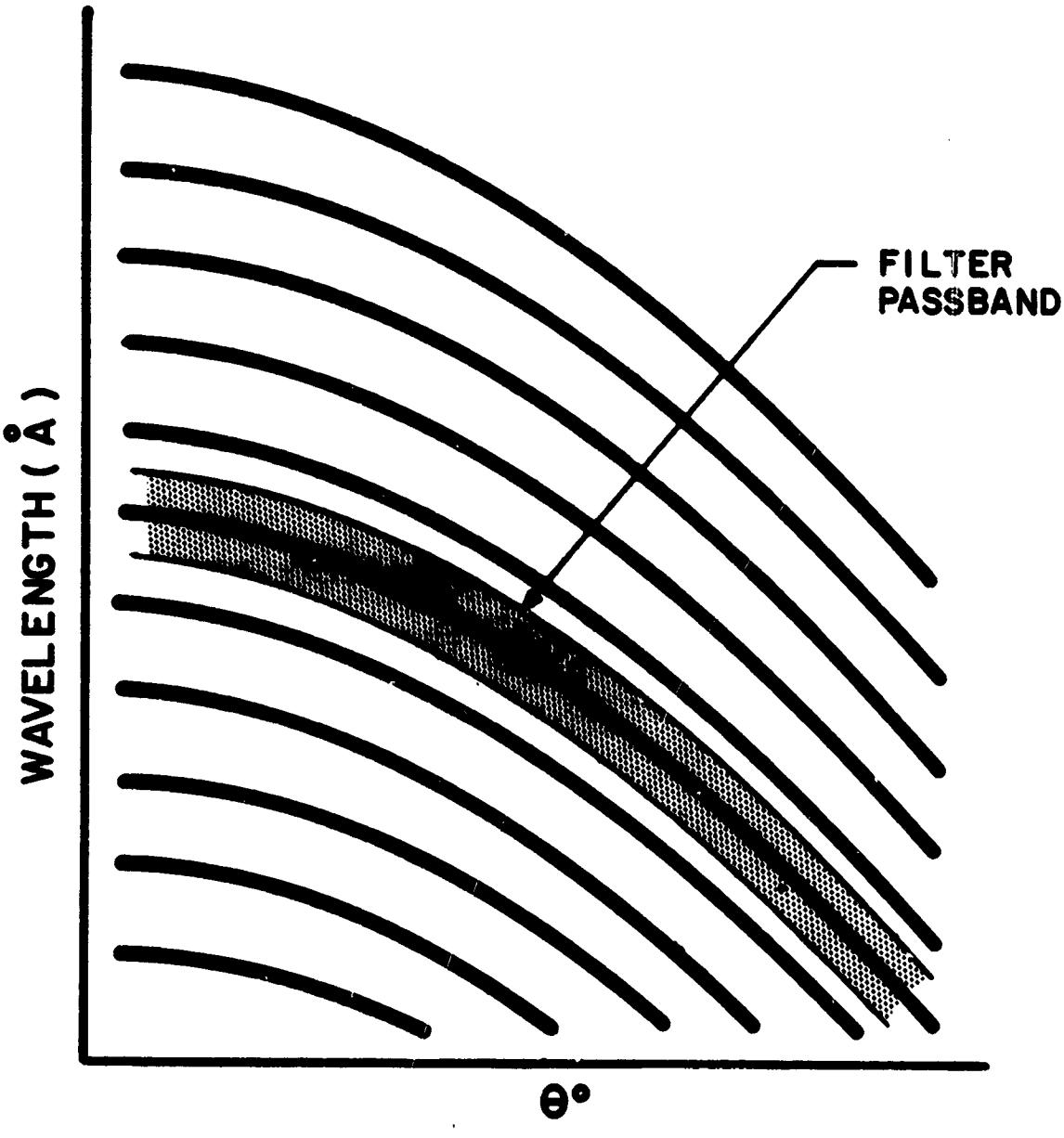


Figure 2. Passband schematic.

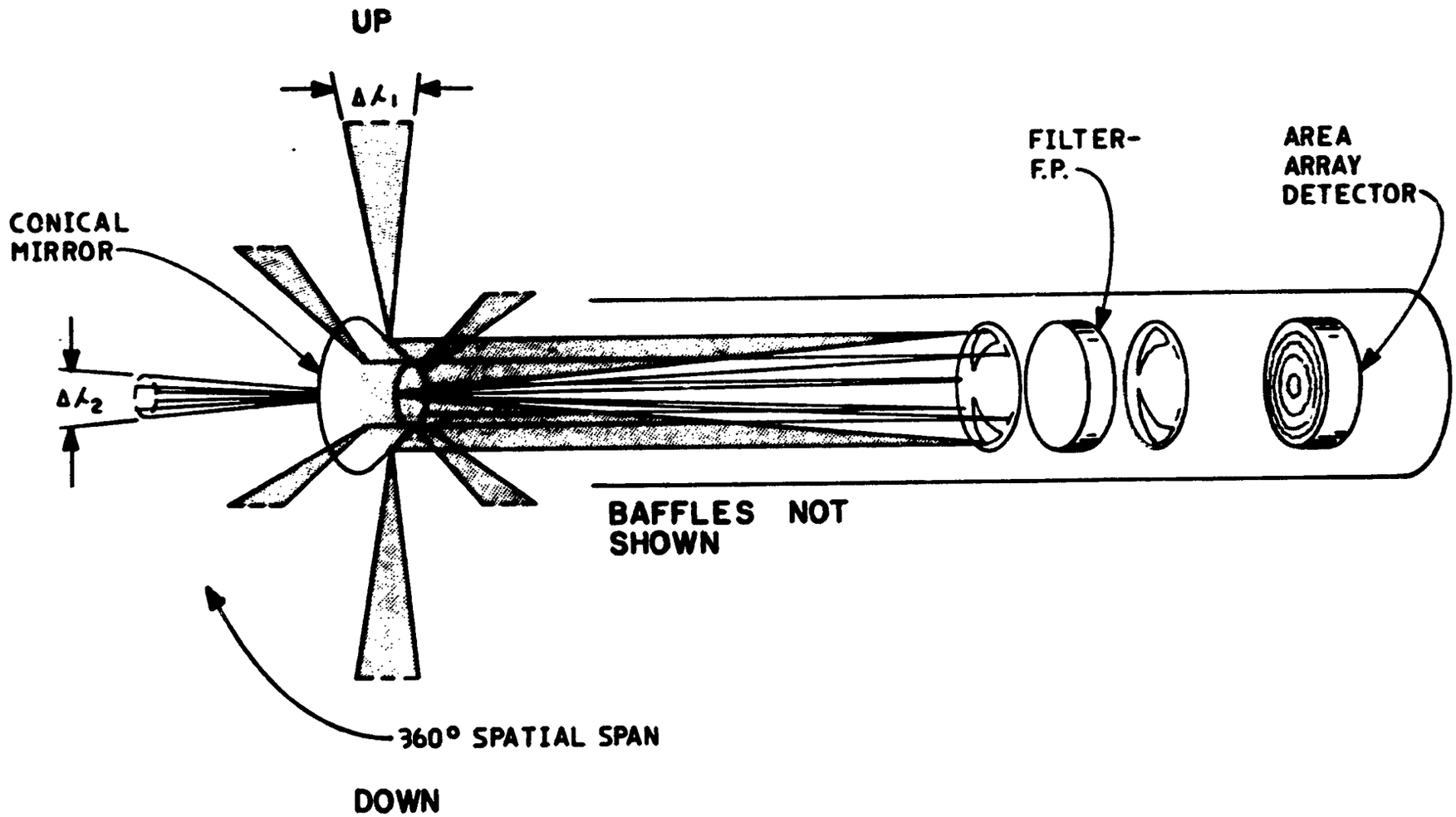


Figure 3. Schematic illustrating possible viewing mode for fixed spacecraft orientation.

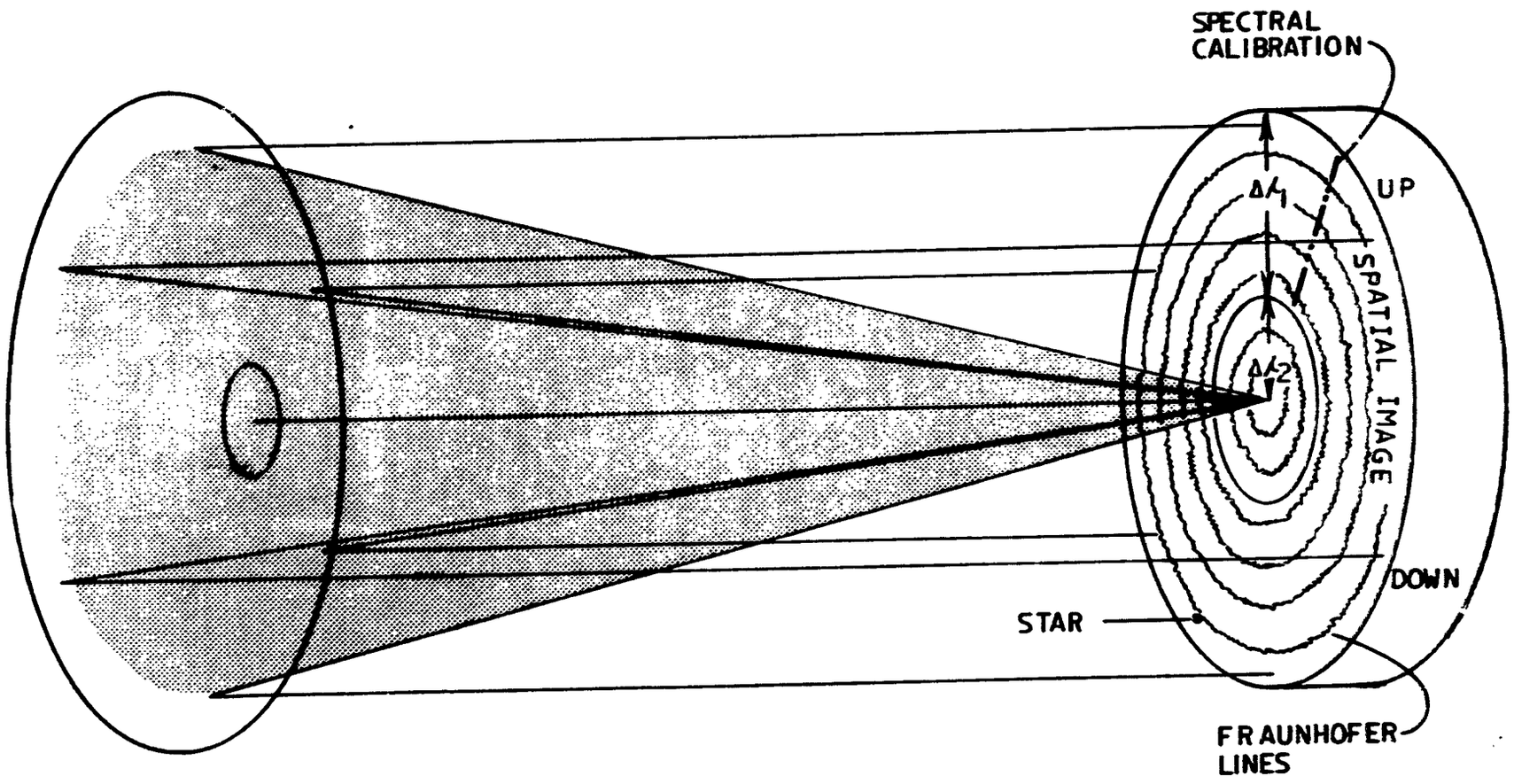


Figure 4. Imaging on the detector.

A resolution of 0.05 Å, for example, would be obtained using two solid etalons in tandem. The use of solid etalons permits a large finesse per etalon (30 to 40), while a finesse of only 26 per simple etalon would yield a spectral range of 20 Å at 0.03 Å resolution. A 4-cm diameter ITT tube coupled to a grid comprising several Fairchild 211 CCD's gives more than 800 detector resolution elements across the 20 Å.

Sensitivity

A field of view of 10° translates into a throughput of about 1 cm^2 ster which, for typical efficiency of the front-end optics amounts to ∿ 30 photoelectron events s^{-1} $S10^{-1}$. For a bandwidth of 0.03 Å, the sensitivity, c, is 1 s^{-1} $S10^{-1}$ at the photocathode. These sensitivities are for the total field of view. If the field of view is divided into 30 spatial sectors, for example, c is reduced to 0.03 photoelectron events s^{-1} $S10^{-1}$ $sector^{-1}$.

Photocathode darkcount is not a limiting factor when area array detection is used for signals with spectral characteristics. CCD darkcount can also be effectively eliminated by radiative cooling.

Dimensions and Power

The shape of the instrument is depicted in Figure 3. The instrument is cylindrical with baffles protruding at the mirror locations. Precise dimensions will depend on the baffling required around the conical mirror. The maximum length will be 45 cm. Optical elements will be about 4 cm in diameter. The volume of the optical housing should not exceed 7000 cm^3 and the electronics housing 2000 cm^3.

Power consumption using 1978 components would be 6 w. Projection to 1983 using manufacturer-supplied trend curves is 3 w.

Other instrument requirements are a temperature range for normal operation of -30 to +60°C, a telemetry rate of 50 bits/s, and 60,000 words of memory storage.

REFERENCES

Beckers, J. M., L. A. Bridges and L. B. Gilliam, A high resolution spectral atlas of the solar irradiance from 380 to 700 nm. II. AFGL-TR-76-0126(11), Environmental Research Papers No. 565, 1976.

SECTION VIII.

GRAVITATIONAL WAVES

D35

GRAVITATIONAL WAVE DETECTION WITH THE SOLAR PROBE: I. MOTIVATION

Kip S. Thorne

N78-32999

California Institute of Technology, Pasadena

What is a gravitational wave? Just as electric and magnetic fields have their wave aspect (electromagnetic waves), so according to general relativity, gravitational fields should have a wave aspect (gravitational waves). Gravitational waves, like electromagnetic, should propagate with the speed of light and should exert forces on matter in a direction transverse to their propagation. Electromagnetic waves push charged particles back and forth transversely relative to uncharged particles. Since gravity couples equally strongly to all objects, gravitational waves should push all forms of matter back and forth transversely with the same acceleration —in other words, they should push inertial frames back and forth. If the amount of pushing were the same everywhere in space, gravitational waves would be undetectable figments of our mathematics. But the amount of pushing is predicted to differ from one region of space to another — a prediction which means that (1) objects and inertial frames in neighboring regions of space should be pushed back and forth transversely relative to each other, and (2) neighboring inertial frames cannot mesh to form a single global inertial frame. In other words, spacetime is curved. In this sense one can say that "gravitational waves are a curvature of spacetime which propagates with the speed of light".[1]

Why is it important to detect gravitational waves? Obviously, the detection of gravitational waves would be a great verification of the predictions of relativistic theories of gravity. Moreover, once waves have been detected, one can test theories of gravity by measuring the propagation speed and polarization properties of the waves.[2] Such tests cut through "the space of all theories" in a direction roughly orthogonal to other experiments: theories that predict identical Shapiro time delay and perihelion shift have very different types of gravitational waves.

Testing theories of gravity is not the most important payoff of gravitational-wave detection. More important is the possible role of gravitational waves as a new window onto the universe — a tool for studying astrophysical phenomena that can be studied poorly or not at all with electromagnetic waves.

What astrophysical information might one learn from gravitational waves? Because matter is electrically neutral on macroscopic scales, cosmic electromagnetic waves are typically an incoherent superposition of emission from huge numbers of individual electrons. By contrast, the fact that all objects have positive mass prevents macroscopic gravitational shielding, and causes the strongest predicted gravitational waves to come from coherent, bulk motions

of huge conglomerates of matter. The waves are especially strong when the velocities of the emitting matter are near the speed of light, and when the density is so high that internal gravity is relativistic. Moreover relativistic gravity alone, in the form of bulk dynamical motions of curved empty space, should generate strong waves. The prime example is a collision between two black holes.

Bulk, high-speed motions and relativistic gravity are central to current speculations about the energy sources of supernovae and about violent activity in galactic nuclei and quasars.[3] The gravitational waves from such phenomena should carry detailed information about the bulk motions and gravity that generated them. Moreover, although the sources of the gravitational waves will typically be buried beneath vast layers of obscuring matter, the waves should emerge through that matter with impunity — no absorption; no scattering; no distortion. This contrasts with electromagnetic waves, which must diffuse out through thick layers of matter, and will lose most of their information as they go. Thus, gravitational waves are a promising probe of violent strong-gravity phenomena in the centers of collapsing stars, in quasars, and in galactic nuclei — phenomena which can be studied only poorly, and perhaps not at all, by electromagnetic waves.

Figure 1 shows an example of the information that gravitational waves might carry. These are the waves predicted to come from a collision between two black holes. Plotted vertically is the dimensionless amplitude of the wave ψ — which is equal to the relative velocity, and hence the relative Doppler shift, that the waves would produce between Earth and a very distant spacecraft. See Estabrook's lecture, below. The amplitude ψ, shown as a function of time, is called the waveform of the gravitational waves. This waveform contains the following information: (1) The fact that the final value of ψ differs from the initial value tells us definitively that the source, initially or finally or both, consisted of two or more bodies moving at high speed relative to each other — in this case two black holes in the initial state. (2) The details of the initial rise and peak, and the difference $\psi_{final} - \psi_{initial}$ carry information about the impact parameter and relative velocity of the collision, and about the orientation of the observer relative to the collision. (3) The damped oscillations, $\psi = e^{-t/\tau}\sin\omega t$, are produced by pulsations of the final coalesced black hole which results from the collision. The two parameters τ (damping time) and ω (frequency) are determined by the mass M and angular momentum J of the

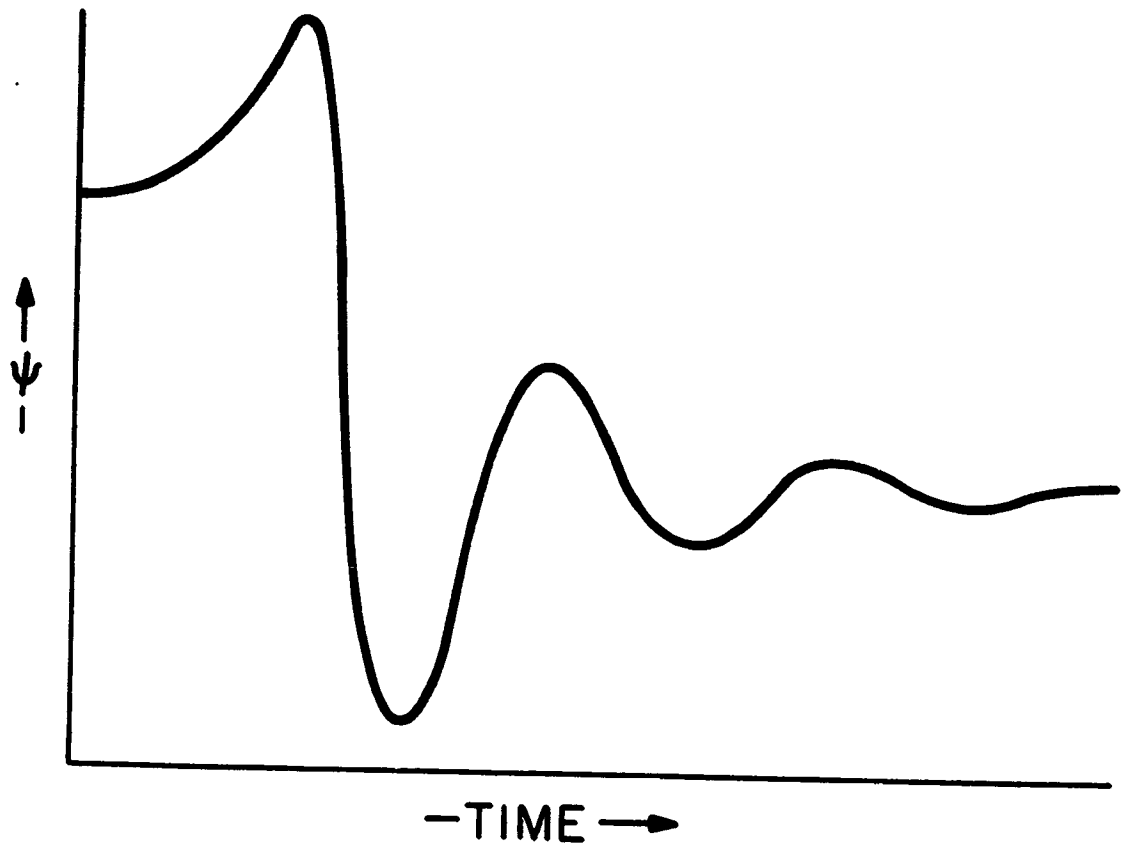


Figure 1. The predicted gravitational waveform from a collision between two black holes — qualitative; cf. refs 4,5.

final hole. From measured values of τ and ω one can hope to deduce M and J — which is valuable information since all properties of the final hole are determined entirely by its mass and angular momentum.

What is the status of attempts to detect gravitational waves? The search for gravitational waves was pioneered by Joseph Weber at the University of Maryland.[6] Weber, and others after him, used one-ton aluminum bars as antennas for kilohertz gravitational waves. Despite early hints of success, it now appears that such waves have not been detected.[7] Nevertheless, prospects for success within the next decade are fairly good. Spurred on by the lure of high payoff, eight different research groups in the US, Canada, Italy, Russia, Japan, and Australia are now building second-generation bar antennas and are designing the third generation.[8]

Most prospective sources should produce waves at frequencies below one kilohertz, where bar antennas are less promising than other detection techniques. For frequencies between one kilohertz and one hertz, ground-based laser-interferometer detectors look best. They are being developed by Weiss at MIT, Drever in Glasgow, and Billing in Munich.[8] Below one hertz, where the waves from quasars and active galactic nuclei should lie, Doppler tracking of spacecraft is the only promising technique in the next decade.[8] Ground-based detectors cannot operate below one hertz because of seismic noise.

Why is the Solar Probe a special mission for gravitational-wave detection? Spacecraft tracking for gravitational-wave detection is in its extreme infancy. No serious efforts have been made on past missions. The Galileo and Solar Polar Missions will be a modest and reasonable beginning — if they have X-band tracking on the uplink as well as down. However, their lack of dual frequency on both links will substantially impede the removal of plasma dispersion noise.

The Solar Probe is the only mission now planned or contemplated which is likely to have most of the features required for an optimal effort at gravitational-wave detection. These features are: (i) dual frequency tracking on all links, (ii) a high-stability onboard clock, which makes possible four Doppler tracking links (one-way up; one-way down; round-trip up-down; and round-trip down-up.), (iii) a drag-free capability, and (iv) two spacecraft, permitting coincidence experiments. Of these requirements, the first three will likely be met; the fourth is less likely but highly desirable. The need for these features will be documented in the lectures that follow.

The advantage of the Solar Probe is that other, non-gravitational-wave experiments will likely require the above features. Without these features, attempts to determine the solar quadrupole moment and various tests of general relacivity will likely be debilitated. Moreover, there is no evident conflict between gravitational-wave detection and other experiments. Gravitational-wave detection must be done during the cruise phase of the mission, when the spacecraft is far from the Earth and far from the Sun. The other experiments will be active near Solar encounter.

If an optimal effort at gravitational-wave detection is not made on the Solar Probe, it is hard to imagine when another equally good opportunity will come along.

How does the Solar Probe's expected sensitivity compare with the strengths of predicted gravitational waves? A reliable estimate of the Solar-Probe sensitivity has not yet been made. Detailed studies of various noise sources are much needed; see the lectures that follow. Rough estimates by Estabrook (below) suggest that gravitational-wave bursts with frequencies $.01 \lesssim \nu \lesssim .0001$ and amplitudes $\psi \gtrsim 10^{-16}$ might be detectable. Integration into the noise might improve this to $\psi \gtrsim 3 \times 10^{-18}$ for continuous-wave sources.

These very rough estimates are compared, in Figure 2, with even cruder estimates of the waves one might hope for from collisions of black holes and black-hole births in quasars and active galactic nuclei at the Hubble distance.[4,9] In the .001 - .0001 Hz band, where the Solar Probe might be sensitive enough for detection, the black-hole sources would have masses $\sim 1 \times 10^{7} - 10^{8}\ M_{\odot}$. This is not unreasonably large in view of current speculations about quasars and galactic nuclei,[3] and in view of the recent identification[10] of a possible $5 \times 10^{9}\ M_{\odot}$ black hole in the nucleus of the giant elliptical galaxy M 87. Estimates of the mean time between gravitational wave bursts at Earth due to such sources range from one week to 300 years or more.[9] Thus, the search for such waves is a high-risk, high-potential-payoff enterprise.

Figure 2 also shows, dotted, the strengths of the continuous waves to be expected from binary stars ($\psi \simeq 10^{-20}$ at the strongest; far too weak for detection by Solar Probe) and also an upper limit on the strength of any continuous-wave background from the big-bang or from discrete sources. This upper limit is the strength of waves whose energy density would close the Universe and would begin to conflict with measurements of the Universal deceleration parameter. Although the Solar-Probe sensitivity is much better than this upper limit, and although

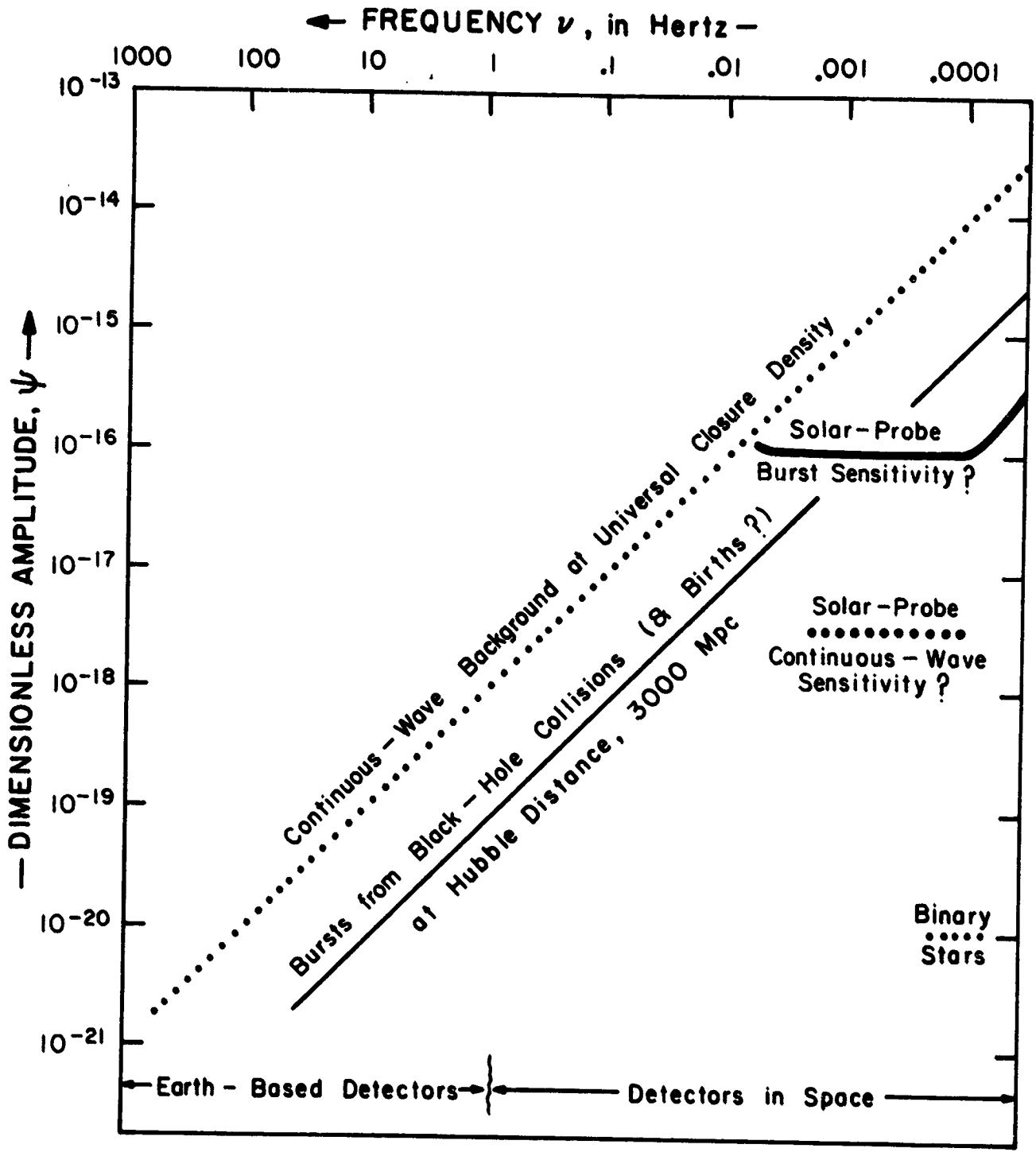


Figure 2. Comparison of estimated Solar-Probe sensitivity with strengths of possible gravitational-wave sources.

it is not inconceivable that waves near this upper limit were produced in the big bang, most astrophysicists would regard as unlikely any continuous-wave background strong enough for the Solar Probe to detect.[11] Again the search is a high-risk but high-possible-payoff enterprise.

My own guess is that, if the Solar Probe is optimally configured for gravitational-wave detection, the probability of successful detection is perhaps 20 per cent.

What will gravity-wave searchers do after the Solar Probe? Whether the Solar Probe succeeds or fails, the effort to use space technology for detecting waves with $\nu < 1$ Hz will continue. Because the risk of failure is so high at the sensitivities $\psi \gtrsim 10^{-17}$ reasonable for Doppler tracking, only if the Solar Probe succeeds will there be a big push for a mission dedicated solely to gravity-wave detection. If the Solar Probe fails to detect any waves, the push initially will be for further detection efforts that go "piggy-back" on other missions; and such missions with reasonable configurations for gravitational-wave detection may be few and far between.

On the other hand, in the more distant future — perhaps in the 1990's — it is likely that laser-interferometer detectors will be flown in space, and will achieve sensitivities $\psi \sim 10^{-21}$ at $\nu < 1$ Hz. Such sensitivities would likely be good enough to reveal waves from a variety of different sources.

In the meantime, I think we would be foolish to not make optimal use of an opportunity like the Solar Probe to search for the strongest gravitational waves that might be bathing the Earth.

This research was supported by the National Aeronautics and Space Administration [NGR 05-002-256 and a grant from PACE]; and by the Caltech/JPL President's Fund.

REFERENCES

1. For details of the curved-spacetime mathematical description of gravitational waves see, e.g., chapters 35-37 of C.W. Misner, K.S. Thorne, and J.A. Wheeler, Gravitation (San Francisco: W.H. Freeman, 1973).
2. See, e.g., D.M. Eardley, D.L. Lee, and A.P. Lightman, Phys. Rev. D, 3308 (1973).
3. See, e.g., R.D. Blandford and K.S. Thorne, in Gravitational Theories Since Einstein, eds. S.W. Hawking and W. Israel (Cambridge England: Cambridge Univ. Press, 1979) in press; also ref. 4..
4. K.S. Thorne, chapter 8 of Theoretical Principles in Astrophysics and Relativity, ed. N.R. Lebovitz, W.H. Reid, and P.O. Vandervoort (Chicago: Univ. Chicago Press, 1978).
5. M. Davis, R. Ruffini, and J. Tiomno, Phys. Rev. D, 5, 2932 (1972).
6. J. Weber, Phys. Rev., 117, 306 (1960), and Phys. Rev. Letters, 22, 1302 (1969). M. Lee, D. Gretz, S. Steppel, and J. Weber, Phys. Rev. D, 14, 893 (1976).
7. For reviews see, e.g. N. Rosen and G. Shaviv eds., General Relativity and Gravitation: Proceedings of GR7 (Jerusalem: Keter Publishing House, 1975); also V. De Sabbata and J. Weber eds., Topics in Theoretical and Experimental Gravitation Physics (London: Plenum Press, 1977); also refs. 8.
8. For reviews of current and future efforts see, e.g., J.A. Tyson and R.P. Giffard, Ann. Rev. Astron. Astroph., 16 (1978), in press; also V.B. Braginsky and V.N. Rudenko, Physics Reports, in press; also D. Douglass and V.B. Braginsky, in Gravitational Theories Since Einstein, eds. S.W. Hawking and W. Israel (Cambridge England: Cambridge Univ. Press, 1979) in press.
9. K.S. Thorne and V.B. Braginsky, Astrophys. J., 204, L1 (1976).
10. P.J. Young, J.A. Westphal, J. Kristian, C.P. Wilson, and F.P. Landauer, Astrophys. J., 221, 721 (1978); W.L.W. Sargent, P.J. Young, A. Boksenberg, K. Shortridge, C.R. Lynds, and F.D.A. Hartwick, Astrophys. J., 221, 731 (1978).
11. See, e.g., L.P. Grishchuk, in Proceedings of Eighth Texas Symposium on Relativistic Astrophysics, Ann. N.Y. Acad. Sci., 302, 439 (1977).

GRAVITATIONAL WAVE DETECTION WITH THE SOLAR PROBE. II. THE DOPPLER TRACKING METHOD

Frank B. Estabrook
Jet Propulsion Laboratory

We briefly describe how gravitational radiation affects Doppler tracking, and give the status of our analyses of the sensitivities achievable with the NASA-JPL Deep Space Net (DSN).

Using precision timekeeping, a Doppler observation can be thought of as a series of measurements, either of the phase $\varphi(t)$ of an electromagnetic wave train being received from a spacecraft, or of the "instantaneous" frequency $\nu = \dot{\phi}/2\pi$ (determined in practice by counting phase over an integration time τ), or of the fractional frequency shift $y = (\nu-\nu_o)/\nu_o$. For a round-trip observation, in which the uplink signal is coherently transponded and returned via a downlink to the same ground station, the phase of the received downlink signal at any instant t is measured with respect to that of the uplink signal then being transmitted. y(t) contains information about the clocks, the motions of the earth and transponder, the propagation medium, and passing gravitational waves.

If the ground timing speeds up momentarily, y will at that instant be observed to dip, while at a later time $2\ell/c$ (the round-trip-light-time, or RTLT, to the spacecraft and back) y will be observed to increase momentarily. Thus for timekeeping frequency noise, a round-trip Doppler measurement system is characterized by a 2-pulse response function or, to be precise, an anticorrelation at the RTLT. Tropospheric fluctuations, or motion of the earth antenna, are also characterized by a 2-pulse response, but for these it is a correlation at the RTLT. Solar wind fluctuations give correlations at intermediate periods. Quite different from all of these is the response function to a passing pulse of gravitational radiation. It is a three-pulse response, with both the spacings of the pulses and their relative heights depending on a single parameter θ. All these response functions are shown in Figure 1.

Gravitational radiation from distant sources has plane wave fronts and two possible polarizations, and travels at the speed of light. In Figure 2 a pulse is shown passing by the earth, at t_o, and past the spacecraft at $t_o + \ell/c \cos\theta$. The first event is echoed back at the earth at $t_o + 2\ell/c$, the second at $t_o + (1 + \cos\theta)\ell/c$. The spacing of the three pulses thus depends on the geometry. The relative sizes of the three pulses can be regarded as due to a superposition of buffeting and clock-slowing effects; in general relativity, perturbations of spacetime geometry have either the guise of strain, $\Delta\ell/\ell$, or

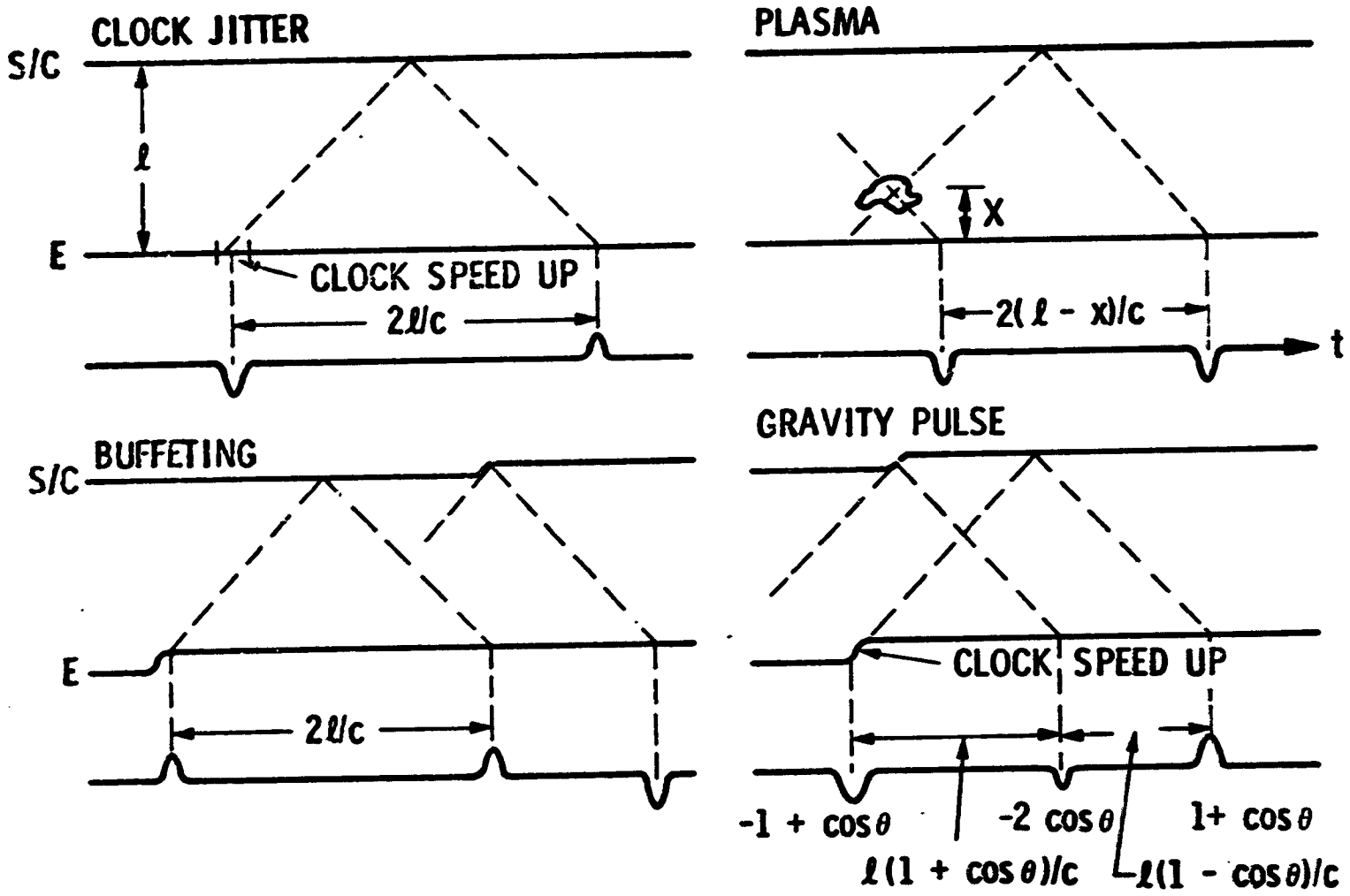


Figure 1. Doppler signals corresponding to different types of disturbance in the communication link.

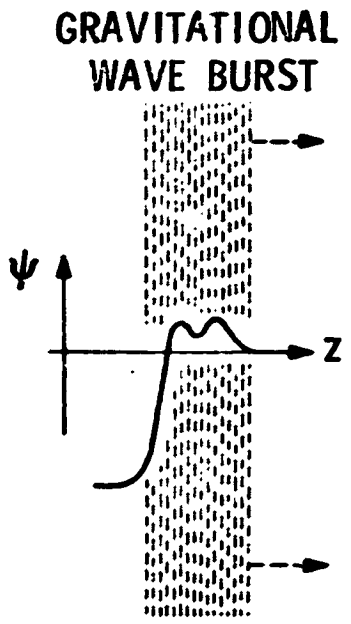
$$\frac{\Delta\nu}{\nu} = -\frac{1-\cos\theta}{2}\psi_R - \cos\theta\,\psi_T + \frac{1+\cos\theta}{2}\psi_E$$

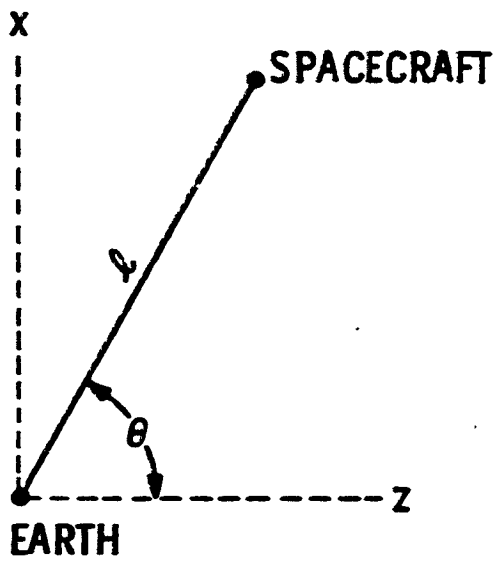
WHERE

$\psi_E = \psi(t_R - 2\ell/c)$

$\psi_T = \psi(t_R - \ell[1+\cos\theta]/c)$

$\psi_R = \psi(t_R)$





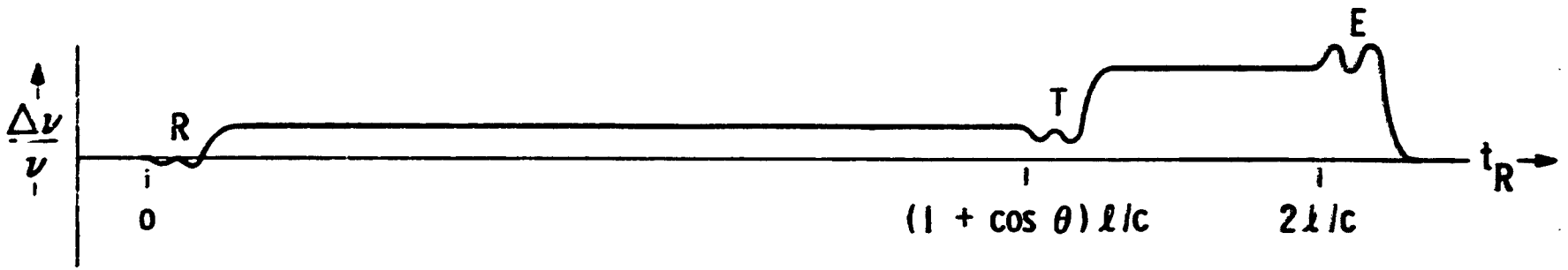


Figure 2. Gravitational Wave Interaction with Doppler Tracking.

of dimensionless gravitational potential, ψ/c^2, and differences of the latter affect the relative rates of separated clocks. The resulting Doppler shifts can be rigorously calculated, and are directly given by the dimensionless amplitude h (or ψ/c^2) of the passing wave, multiplied by the three pulse response function of Figure 1.

The three-pulse response demonstrates nicely both the transverse nature of gravitational radiation and the equivalence principle. As $\theta \to 0^\circ$ the first pulse vanishes and the second two pulses merge and cancel. There is no longitudinal effect! Also, for any θ, the sum of the pulses is zero--if the pulse duration approaches the RTLT, they blend and cancel. This imposes a long period (low frequency) limit to the gravitational radiation detectable with Doppler tracking. For pulse durations $\tau \geq 2\ \ell/c$, the tracking link rides up and down on the wave of propagating gravitational potential and the measurement becomes "local" and thus unobservable according to the principle of equivalence. Since a large RTLT is necessary for Doppler gravitational wave searches, the appropriate missions in the foreseeable future are Galileo, Solar Polar and Solar Probe.

We are looking into data reduction methods which take full advantage of these very characteristic and unique response functions to search for gravitational waves. In a subsequent paper, Robert Vessot shows how Doppler signals can be separated from other effects in real time with an on-board H-maser frequency standard and multiple link tracking, one-way, two-way and four-way. These developments must be actively supported and carefully studied in the next few years, and also closely coordinated with the advanced planning of the 1980-90 era Deep Space Net. A dramatic increase in the sensitivity achievable for our experiment is possible by the time of the Solar Probe mission.

Nevertheless, for this planning meeting we will be conservative and report achievable system sensitivities simply by considering Allan variances, $\sigma_y(\tau)$, for various system and propagation sources of phase fluctuation. Allan variance is additionally conservative for consideration of detection sensitivities for radiation with spectral content very different from that of the phase noise. In his paper, David Douglass discusses the much higher sensitivities possible for cw sources, which are better understood by using spectral plots of frequency noise $S_y(f)$. Plasma physicists are more used to the concept of phase noise

$W_\phi(f) = (\nu^2/f^2)\, S_y(f)$ because it is directly related to columnar electron content. We have adopted the Allan variance σ_y in order to communicate effectively with the clock and timekeeping community: σ_y is the rms value of two immediately successive measurements of y, each for integration time τ. It can be obtained by integrating the spectral descriptions W_ϕ or S_y with a filter which cuts out high and low frequency divergences--and indeed for our purposes may not filter the lows enough, and so be too conservative. Plasma, troposphere and H-maser phase fluctuations, however, do all tend to have the same low frequency spectral characteristics, and can at least be intercompared with this measure.

A precision tracking experiment is being performed this year with the Viking spacecraft at Mars, at the request of J. D. Anderson, J. W. Armstrong, F. B. Estabrook and R. Woo. Viking went through solar opposition on January 18, 1978, and data taken and analyzed by Armstrong and Woo have beautifully verified a model by Armstrong which predicts a dramatic lowering of the solar wind plasma fluctuation level near 180° elongation, when plasma inhomogeneities are convected nearly parallel to the radio link. From this and previous data, one finds a Kolmogoroff power spectrum. So, let $W_\phi = A\,(10^3 \times f)^{-2.6}$. Then Figure 3 (from Woo, Armstrong, and Estabrook, 1978) shows A, the spectral amplitude at 10^{-3} Hz, plotted against elongation or sun-earth-spacecraft angle. The points for elongations greater than 90° are from preliminary reductions of the new Viking data and corroborate the Armstrong model, indicated by the line in Figure 3. On the Allan variance Doppler sensitivity plot, Figure 4, we can thus show the fluctuation levels due to the interplanetary plasma to be an order of magnitude lower than in our previous plots of this sort (Wahlquist et al., 1977). It appears that for gravitational wave observations conducted at X-band acceptable levels of plasma phase fluctuation prevail for about a month centered on each solar opposition during the cruise phase of future missions to the outer planets.

We have commented that the low frequency limit of the Doppler gravitational wave detection scheme is set by the RTLT. From Figure 4 it is clear that the high frequency limit is set by the available timekeeping standards and system noise. The useful VLF band is .03 Hz - .0003 Hz, which lies in the VLF gravitational wave band defined by Press and Thorne to be .1 Hz - .0001 Hz.

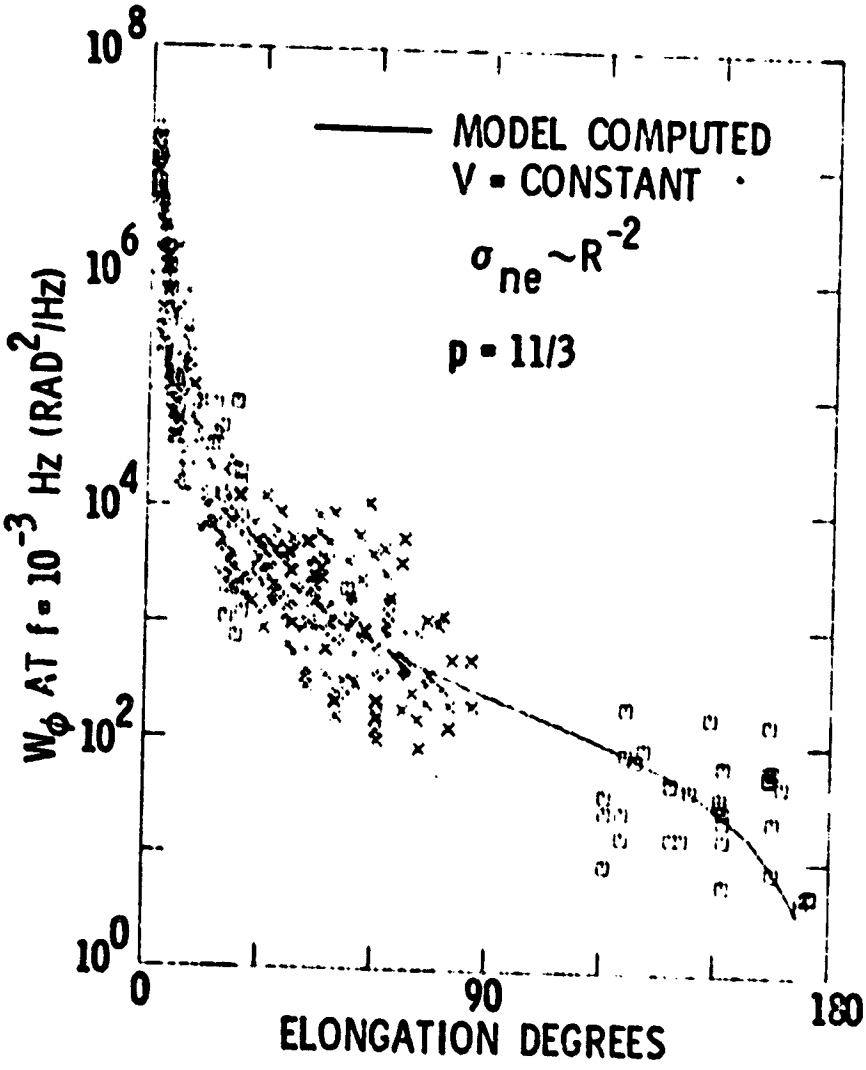


Figure 3 Spectral amplitude at 10^{-3} Hz versus elongation. From Woo and Armstrong (1978).

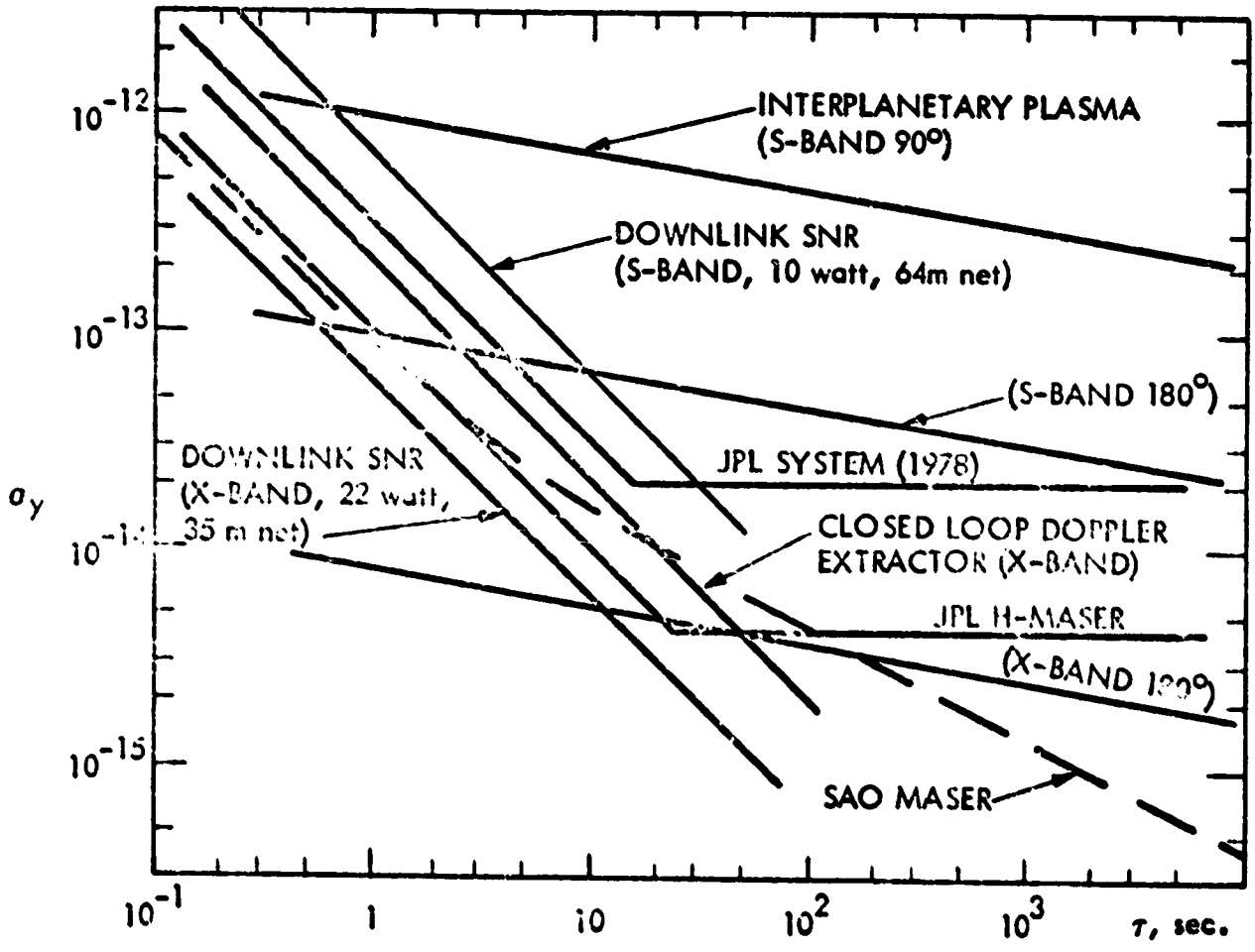


Figure 4. Doppler tracking sensitivity limits.

In Figure 4 we have not shown Doppler fluctuation levels caused by tropospheric or spacecraft buffeting. These are non-dispersive. The Viking data were obtained at opposition in both S-band and X-band, and are still dispersive--hence plasma related--at the level shown, $\Delta f/f = 3 \times 10^{-14}$ (set by the Viking S-band uplink). We have estimated spacecraft buffeting accelerations in this VLF frequency band as $\sim 3 \times 10^{-15}$. Available tropospheric data are discussed in the paper by Phil Callahan; it is clear that more study and experiments are needed, especially under nighttime desert conditions. We believe the ultimate solution to such propagation problems lies in on-board timekeeping and multilink tracking, solutions possible with the Solar Probe.

Returning again to Figure 4, we call attention to the plots of phase fluctuation associated with the downlink signal-to-noise ratio (SNR). The downlink is always the weak link in spacecraft communications, because of the low power available on board and despite the low noise temperature of ground maser receivers. For S-band downlink, antenna gains and the frequency dependence of S_y ($\sim \nu^{-2}(SNR)^{-1}$) conspire, and downlink phase fluctuations can be seriously large at the levels with which we are concerned. Although there will be a 20 watt downlink X-band, the minimal 1-watt S-band "beacons" presently planned for the Galileo and Solar Polar missions would, we believe, be marginal as auxiliary multilinks. Those missions will probably be able to conduct gravitational radiation demonstration experiments only at an Allan-variance-level of 3×10^{-14} set by the S-band <u>uplink</u> plasma fluctuation, unless X-band converter stages are added to allow round-trip X-band observations. In that case, these large RTLT missions can achieve a Doppler accuracy of 3×10^{-15} near solar opposition, with only an X-band system.

Pending more careful study of the use of multilink tracking to take advantage of the distinctive correlation signature of gravitational radiation, we estimate the achievable gravitational radiation strain sensitivity of the Solar Probe mission at $\leq 10^{-16}$. This great advance could be made possible by the incorporation of on-board H-maser timekeeping. We urge that the needed advanced planning of multilink Doppler for the mission be coordinated with the development plans of the DSN.

References

Wahlquist, H. D., Anderson, J. D., Estabrook, F. B., and Thorne, K. S., 1977, Experimental Gravitation, Atti dei Convegni Lincei, 34, 335-350.

Woo, R. and Armstrong, J., 1978, EOS, 59, 366.

Woo, R., Armstrong, J., and Estabrook, F. B., 1978 (in preparation).

This report presents one aspect of research done by the Jet Propulsion Laboratory for NASA under contract 7-100 and under a grant from the President's Fund of the California Institute of Technology.

D31

N78-33001

Transmission Media Effects on Precise Doppler Tracking

Philip S. Callahan

Jet Propulsion Laboratory
4800 Oak Grove Drive
Pasadena, California 91103

Abstract

The effects of the transmission media - the earth's troposphere and ionosphere, and the solar wind - on precise Doppler tracking are discussed. The charged particle effects can be largely removed by dual frequency observations. However, there are limitations to these corrections (besides system noise and/or finite integration times) including the effects of magnetic fields, diffraction, and differential refraction, all of which must be carefully evaluated. The Earth's troposphere can contribute an error of $\Delta f/f \approx 10^{-14}$.

I. Introduction

Investigations involving high accuracy Doppler tracking are likely to be an important part of Solar Probe science. In the past, it has been found that the media through which the radio signal passes can have important effects on the signal. I would like to reconsider these transmission media effects relative to the higher tracking accuracies discussed for relativity, quadrupole, and gravitational wave experiments. Doppler accuracies of 0.1 mm/sec ($\Delta v/c = 3 \times 10^{-13}$) have generally been assumed for the first two experiments, while gravitational wave detection requires accuracies of $\Delta f/f \sim v/c$ better than $\sim 10^{-15}$.

I consider a round trip S-X band system for correction of the effect of charged particles. The results show that the gravitational wave experiment is likely to be limited by the Earth's troposphere. The other experiments will require careful calibration, perhaps that provided by an accurate onboard clock which allows both one and two way Doppler measurements. Further studies of the effects discussed here are needed to determine optimal calibration strategies.

II. Solar Wind Plasma Effects

This paper deals with only the solar wind and the Earth's troposphere. Even though some of the plasma effects evaluated here occur in the ionosphere, it is usually of less importance because of the short path through it. In the absence of magnetic fields, a plasma is a dispersive medium with an index of refraction given by

$$n^2 = 1 - f_p^2/f^2$$

where the plasma frequency $f_p^2 = 2Ne^2/m$ or $f_p = 8.98 \times 10^3 N^{\frac{1}{2}}$, N being the number density of electrons. The group delay or phase advance caused by a plasma can be completely removed by observing at two frequencies. At present, spacecraft communications are carried out with an S band (2.1 GHz) uplink, while the spacecraft coherently retransmits S and X band (8.4 GHz) providing a one way measurement of the integrated electron content.

Since real plasmas are not static, and one is interested in knowing the round trip effect, one needs either two way dual frequency measurements or a clever way to obtain the round trip effect from the downlink observations. Simulations by S. C. Wu and F. B. Winn (Ref. 1) showed that errors of several meters could be made from approximating the round trip effect by doubling the downlink. They developed a technique in which two way data were used to estimate the plasma distribution, and hence the approximate timing of changes. This reduced errors by about a factor of 10, to 10 to 30 cm. Thus, 10-30 cm is the limiting error for one way dual frequency measurements of a changing plasma.

One might suppose that a round trip dual frequency system would eliminate all charged particle errors. While it does reduce them to a very low level, a number of effects remain. Among the effects in need of consideration are the gyrofrequency in the plasma, different paths through the plasma, and diffraction effects. All of these effects are small ($\leq 10^{-4}$) factors times the basic plasma effect which is f_p^2/f^2. The values of some plasma parameters are given in Table 1.

The effects of spacecraft and solar wind velocities on charged particle corrections were investigated some time ago by Mo (Ref. 2). He noted the well-known result that a medium whose index of refraction is like that of a simple plasma does not "drag" its refractive index with it. Thus, there is <u>not</u> a correction term of order $(v_{\text{solar wind}}/c)$. Furthermore, with a round trip dual frequency system, there is no direct contribution from the spacecraft velocity. (Table 2 shown as a viewgraph at the Workshop was <u>incorrect</u> in showing such a contribution). The effects which will remain after calibration are of order $(f_p/f)^2$ and $(v_{sc}/c)(f_p/f)$, where v_{sc} is the spacecraft velocity ($\sim$300 km sec^{-1}, $v_{sc}/c = 10^{-3}$, near the sun).

The solar wind has a magnetic field imbedded in it. The refractive index is then not the simple relation given above but

$$n^2 = \frac{1 \pm f_g/f - f_p^2/f^2}{1 \pm f_g/f}$$

TABLE 1

PLASMA PROPERTIES

R ($R_\odot$)	N_e (cm^{-3})	f_p (Hz)	f_g (Hz)	$(f_p/f_s)^2$	$(f_p/f_s)^2 \cdot R$ (m)	f_g/f_s	δN_{el} (100 km) (cm^{-3})
5	2.8×10^{4}	1.5×10^{6}	1.3×10^{5}	4.7×10^{-7}	1.6×10^{3}	6.0×10^{-5}	2.7×10^{1}
10	5.3×10^{3}	6.6×10^{5}	3.3×10^{4}	8.9×10^{-8}	6.2×10^{2}	1.5×10^{-5}	7.8
21.6	1.1×10^{3}	2.9×10^{5}	7.0×10^{3}	1.8×10^{-8}	2.7×10^{2}	3.2×10^{-6}	2.0
216	9.4	2.7×10^{4}	7.1×10^{1}	1.6×10^{-10}	2.3×10^{1}	3.2×10^{-8}	3.1×10^{-2}
Iono-sphere (R = 200 km)	2.0×10^{5}	4.0×10^{6}	8.4×10^{5}	3.3×10^{-6}	7.0×10^{-1}	3.8×10^{-4}	~2.0% ($\Delta t = 10^3$ sec)

$$f_p = 8.98 \times 10^3 \, N_3^{\frac{1}{2}} \text{ Hz}$$

$$f_g = 2.80 \times 10^6 \, B \text{ Hz}$$

$$n_p^2 = \frac{1 - f_p^2/f^2 \pm f_g/f}{1 \pm f_g/f}$$

$$f_s = 2.1 \text{ GHz}$$

for propagation along the magnetic field, where the gyrofrequency $\omega_g = eB/mc$, or $f_g = 2.80 \times 10^6$ B Hz. The gyrofrequdncy gives rise to Faraday rotation of linearly polarized waves. Spacecraft communications are usually carried out with circular polarization. Since the solar wind's magnetic field fluctuates, there will be some RMS phase noise from integrating the gyro-effect along the signal path. A further consequence of the gyrofrequency term in the phase velocity is that the solar wind will tend to drag the wave with it. This introduces a term of order $(v_{sw}/c)(f_g/f)$ integrated along the line of sight.

Very close to the sun, the large value of the radial derivative of the refractive index causes the S and X band ray paths to separate and the different ray paths will sample different plasma. Work is in progress to evaluate the magnitude of this effect for the Solar Probe and other missions. If it is significant, it may be necessary to use higher communications frequencies.

The final plasma effect to be considered is diffraction. So far the discussion has been in terms of geometrical optics. However, radio scintillation work has shown that diffraction effects can be important in investigating the solar wind (see, for example, Ref. 3). The Rytov approximation can be used to find the RMS phase difference between S and X band waves propagating in the solar wind.

Although all of these higher order processes also affect range data, they are generally not important. There is an additional factor which is unique to ranging because it requires a finite bandwidth, of the form $\Delta f/f$. For present ranging systems at S band $\Delta f/f \sim 10^{-3}$. Ordinarily, the overall effect $f_p^2 \Delta f/f^3$ is not important, but near the sun we have $\sim 2 \times 10^{-10}$ (5 $R_\odot$) to 4×10^{-11} (10 $R_\odot$) giving $\Delta R \sim 30$ m to 6 m. Since the effect is integrated along the line of sight and the range measurement is made by hardware, only statistical correction is possible.

The effects at 1 AU are all small. This bodes well for the attempts to detect gravitational radiation on long ray paths in the anti-solar direction. Diffraction needs to be evaluated carefully for longer paths where larger scale sizes (with much larger spectral amplitudes) become important. On the other hand, in the antisolar (radial) direction an additional helpful feature emerges. Since one is then looking along the solar wind velocity vector, the rate at which structures are converted across the line of sight decreases and the noise spectrum appears to be shifted to lower frequencies. This effect has recently been observed by Woo and Armstrong (Ref. 4; see also paper by Estabrook this volume).

III. Tropospheric Errors

Besides being affected by charged particles, the radio signal is slowed by the Earth's neutral atmosphere. The tropospheric effect is usually separated into dry and wet parts. The dry part is rather stable and is easily removed by measuring the barometric pressure at the tracking station. The delay caused by water vapor is small ($\sim$10 cm at zenith) but neither steady nor easily modelled. Water vapor radiometers, which will be installed at the Deep Space Stations in about 1980-81, will measure the line of sight delay to an accuracy of 2 cm. However, path length fluctuations of

$$\Delta L \cong .13\ (t_1/400\ \text{sec})^{0.75}\ \text{cm},\quad 10 \lesssim t_1 \lesssim 10^4\ \text{sec}$$

where t_1 is the time scale of the variations, can be expected. Such fluctuations give, in the terms usually used in gravitational wave detection studies

$$\Delta v/c \cong \Delta f/f \cong 1.1 \times 10^{-14} \left(\frac{400\ \text{sec}}{t_1}\right)^{0.25},$$

where t_1 should be associated with the sampling interval ($\sim 10^2$ - 10^3 sec). At 1 AU, the troposphere produces the dominant effect. The fluctuations are likely to be least at night and in the winter, but no data are available to evaluate the magnitude of the reduction.

Acknowledgement:

I would like to acknowledge helpful conversations with D. O. Muhleman, R. D. Reasenberg, and J. R. Jokipii, and words of caution from G. Colombo

References:

1. S. C. Wu and F. B. Winn, 1977, Deep Space Network Progress Report, 42-41, 57.

2. T. C. Mo, 1971, JPL Technical Memorandum, 391-153.

3. R. Woo, 1975, Astrophys. J., 201, 238.

4. R. Woo, 1978, EOS 59, 366.

A TIME-CORRELATED FOUR-LINK DOPPLER TRACKING SYSTEM

R. F. C. Vessot and M. W. Levine

Center for Astrophysics
Harvard College Observatory and Smithsonian Astrophysical Observatory
Cambridge, Massachusetts 02138

INTRODUCTION

The chief objective of a Doppler tracking system is to retrieve range-rate information from the Doppler shift of signals exchanged between the spacecraft and the ground station. Many perturbations occur in the phase of the tracking signal as it progresses from the master oscillator to the transmitter and then radiates into space and is later received and compared to an oscillator. In the two-way tracking mode the signal is returned from the tracked object by a phase-coherent transponder and is received and compared to the same oscillator that generated the transmitted signal. In the one-way mode the spacecraft and ground station each have an oscillator and one station transmits signals to the other to make frequency comparisons and thus obtain the Doppler signal. Each of these modes of operation requires a specific type of electronics system and will be affected by its own set of perturbations.

Specifically, we will consider phase perturbations that occur in the following components of the one and two-way systems:

1) One-way System (Signal from A to B): Station A Oscillator — transmitter A — antenna A — signal path — antenna B — receiver B — station B oscillator — recording system B.

2) Two-way System (Signal from A to B back to A): Station A oscillator — transmitter A — antenna A — signal path up — antenna B — transponder B — antenna B — signal path down — receiver A — Station A oscillator — recording system A.

There is also a three-way mode of operation where the signal path down, receiver, oscillator, and recording system of the two-way system are at a different location than A and B. We will not consider this situation in this paper.

Here we describe a symmetrical one- and two-way system in which we can identify the place of origin of the fluctuation and quantify it so as to recover the unperturbed phase at a given moment of time by time-correlating the phase fluctuations obtained in the one- and two-way data taken at both stations.

In describing the projected performance of such a system we look at the results of a system that was flown in the gravitational red-shift rocket probe experiment (1), which involved the tracking of a vehicle that included an S-band transponder and a hydrogen maser-controlled S-band transmitter. In this experiment both one-way and two-way data were simultaneously obtained at the ground station.

C-6

OSCILLATOR AND CLOCK PERFORMANCE

We begin by considering the behavior of clocks suitable for Doppler tracking. The clock or oscillator serves the following functions in Doppler tracking systems: (a) Transmitter master oscillator, (b) Receiver local oscillator, (c) Clock to keep time (or epoch) to relate data to a celestial coordinate frame, (d) Clock to maintain synchronization (in the sense of coordinate time), and (e) (in the case of the red-shift experiment) The spacecraft clock as the probing instrument.

In the Doppler system both range-rate and range data are available. The range rate is obtained by using the oscillator to measure the Doppler frequency and range differences can be inferred by counting Doppler cycles.

Alternatively, the signals can be modulated or coded so that the light time of the signals can be obtained from the time interval measured between transmitted and received signals.

The measurement accuracy of range rate depends on knowing the frequency of the signal and the drift of the oscillator between the transmission and reception times. The range accuracy in the modulation mode depends on the accuracy of timing the interval between the transmitted and received signals. In these cases, the roles of oscillator and clock are quite obvious. However, there is also a requirement for maintaining epoch time accurately so as to relate these measurements to celestial phenomena and obtain the trajectory or ephemeris of the vehicles in a known coordinate frame.

Oscillator synchronization and stability for the performance of time correlation of the data are also important if it is necessary to obtain angular information from widely separated stations in the conventional VLBI (very long baseline interferometry) mode of operation. As will be seen later, there are other ways in which correlation techniques can be used in tracking applications.

Clock performance is described in terms of the two-sample variance, or Allan variance (2, 3), which is essentially a means of predicting the 1σ probability of the difference between a frequency measurement averaged over the τ seconds in the immediate future and measurements made in the past (also averaged over τ seconds). It can be shown that the functional behavior of $\sigma(\tau)$ depends on the type of random process perturbing the oscillator (4) and that conclusions from such data are only valid if random processes are at work. Systematic frequency effects can lead to interesting data but these data must be used very cautiously.

Allan variance data for hydrogen masers recently delivered by SAO to the U.S. Naval Research Laboratory (5) are shown in Figure 1. Their stability lies well below the 10^{-14} level and follows the limits imposed by thermal noise for $\tau = 0.8$ sec to $\tau = 3000$ sec reaching 6×10^{-16} at averaging intervals of 1 hr. In this data set, taken over 3 weeks, slow linear drifts were observed that became significant for averaging times greater than 1 hr. These were caused by a laboratory air conditioner malfunction during the run. Such slow linear drifts, when represented by the algorithm used to calculate variances, give the data points to the right of the dotted line.

To the best of our knowledge, these data represent results from maser oscillators that are close to the state of the art at the time of writing this paper. The Doppler

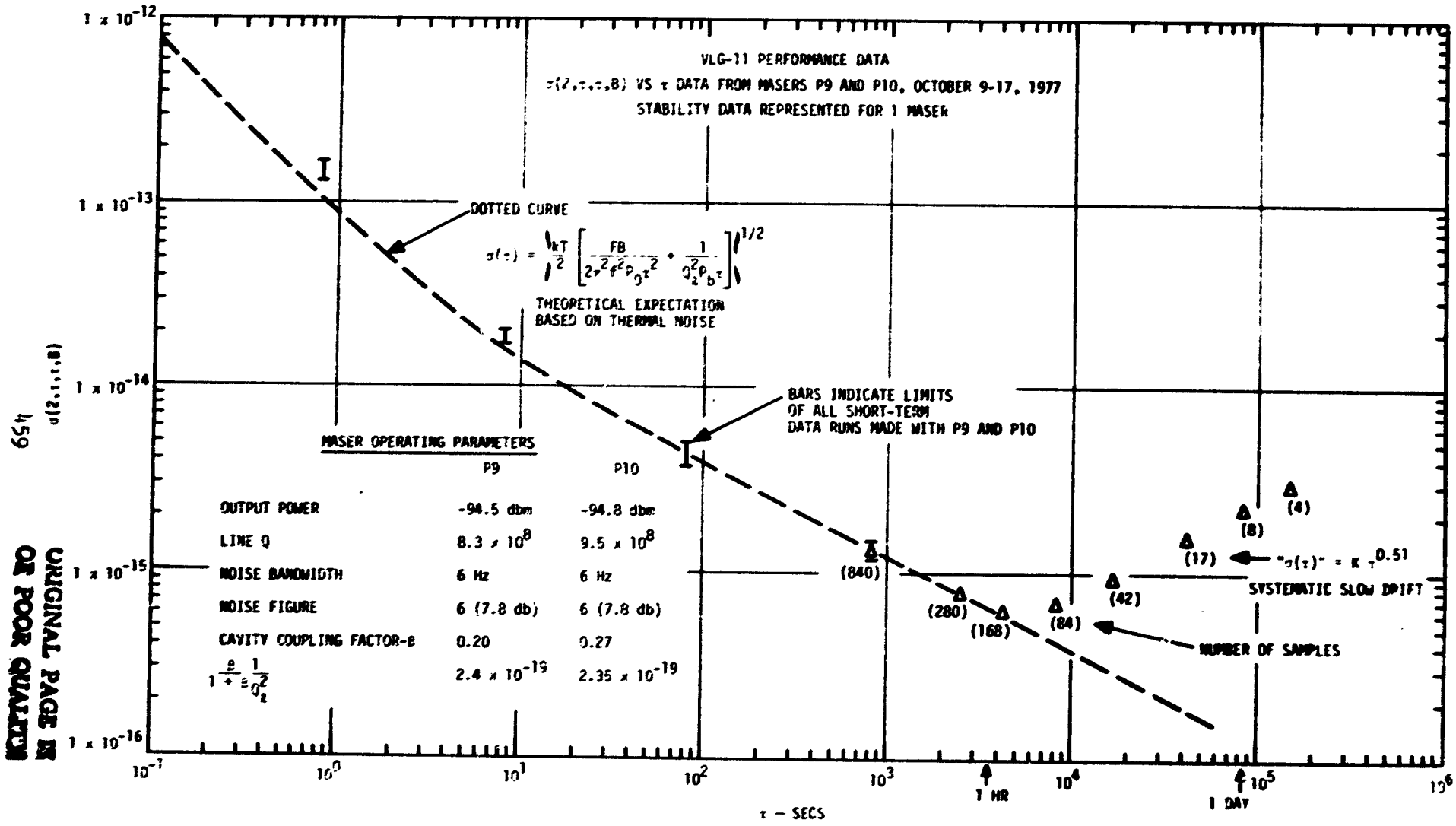


Figure 1. Allan variance data for an SAO hydrogen maser.

error $\Delta f/c$ can be related to the value of $\sigma(\tau)$ when it is associated with the light time of a two-way system. Oscillator or clock performance is a fundamental limit to all range and range-rate Doppler systems and we see from the data that this limit lies well below 1×10^{-14} for time intervals, τ, ranging from 50 sec on to well beyond 1 day.

ELECTRONICS AND ANTENNA SYSTEM HARDWARE

An estimate of the stability that can be obtained from the electronics and antenna systems of a tracking station can be illustrated by the performance of the modified U.S.B. system used in the June 1976 redshift rocket probe experiment. The schematic concept (Figure 2) shows the Doppler cancelling system developed for the experiment. This system was integrated into the existing U.S.B. Station at Merrit Island, Florida, and incorporated many modifications. Probably the most significant contribution to maintaining the phase stability of the transmitted signal was the use of a hydrogen maser to control the phase of the transmitter power amplifier. This was done by comparing the phase of the signal sampled at the transmitting antenna with the phase of a signal derived from the hydrogen maser and controlling the transmitter to maintain a constant phase relationship between the two signals. With this method of control the phase stability of the output power at the antenna was comparable to that of the H maser.

In all the following discussions we will consider that the transmitters are locked in this way to the maser oscillator and that the transmitter phase instability contributes no significant degradation to the signal delivered to the antenna.

The function of the redshift experiment Doppler cancelling system was to measure the Doppler frequency of the spacecraft using a transponder in a two-way (go-return) loop and simultaneously, using the same antenna, to measure the Doppler frequency of the maser-controlled spacecraft transmitter. The system then subtracted one-half the two-way Doppler cycles from the one-way Doppler, which also contained the redshift and second-order Doppler frequency shifts, to reveal these data directly.

Clearly Figure 2 is an oversimplification and Figure 3 is offered to describe the frequency offsets required to operate such a continuous wave system.

If we assume that the ray paths in all three links are essentially identical and that the light-time delay is small compared to the time averaging interval used to recover the data, we can cancel the dominant 1/f component of the ionospheric refractive index by choosing the frequency ratios shown in Figure 3 to satisfy the following equation (6):

$$\frac{P}{Q} = \frac{R}{S}\left[\frac{2}{1 + (N^2/M^2)}\right]^{1/2} \qquad (1)$$

In our case the transponder ratio was $M/N = (240/221)$ and we chose $R/S = 82/55$ and $P/Q = 76/49$, which fit the equation with a discrepancy of about 2×10^{-5}. The resulting possible residual effect, assuming a normal day-time ionosphere and a near-vertical trajectory going to 10,000-km altitude, is on the order of a few parts in 10^{15} of the carrier frequency at 2 GHz.

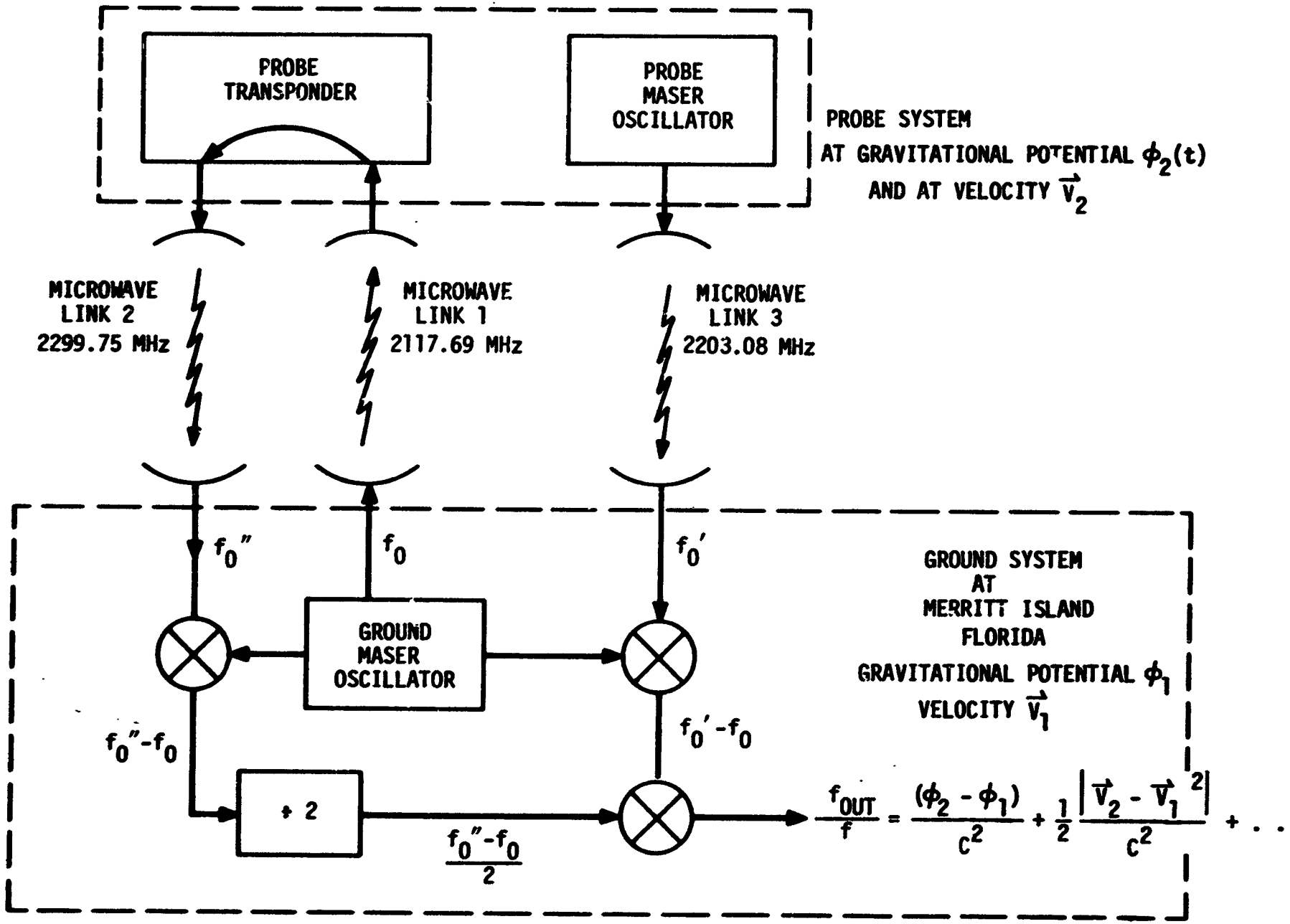


Figure 2. Redshift experiment Doppler cancelling system.

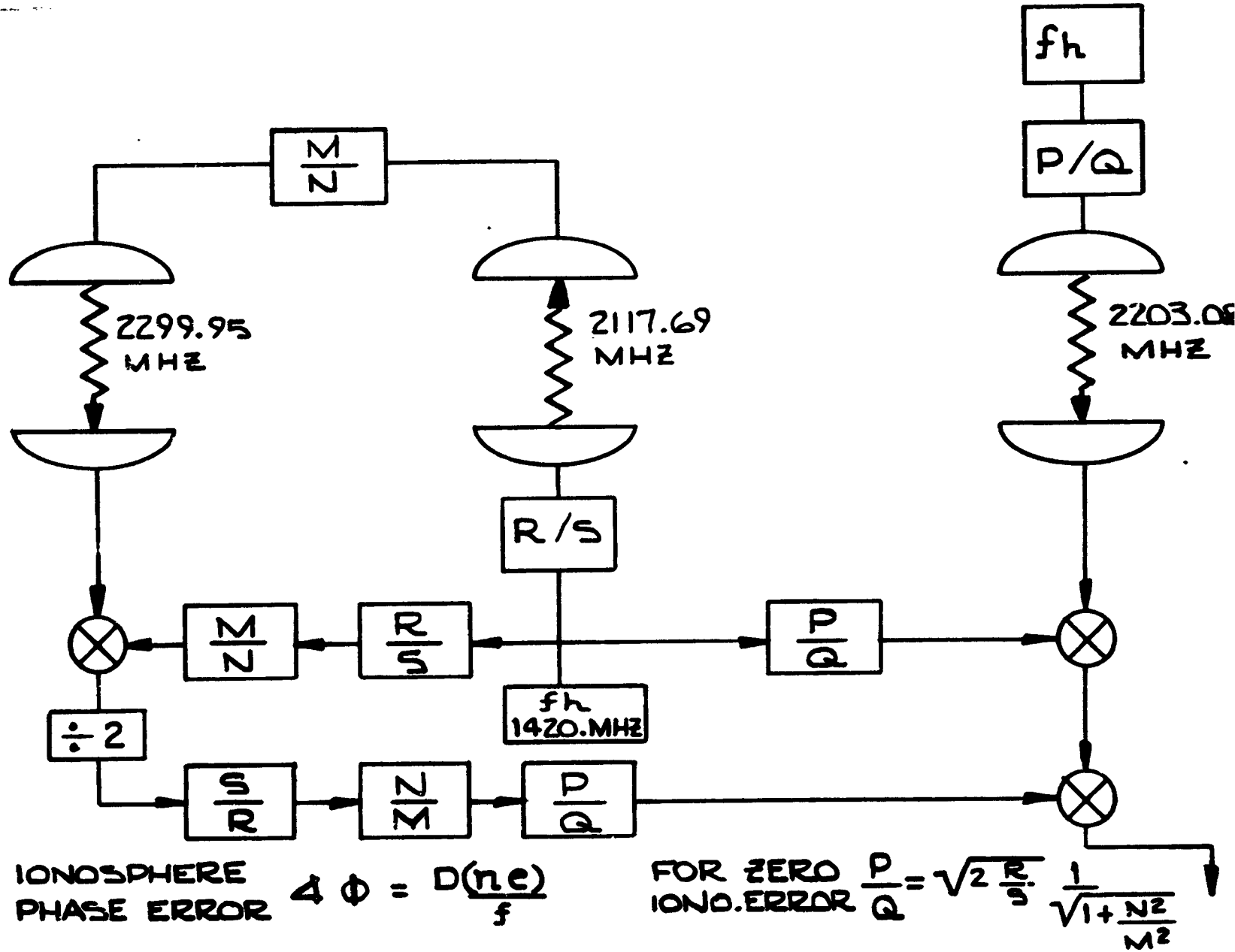


Figure 3. Redshift Doppler and ionospheric error cancelling system.

The system in Figure 3 is a phase-coherent device that, in real time, measured the Doppler cycles of the downlink signal controlled by the probe maser by a separate two-way microwave link and electronically subtracted these cycles in real time from the one-way link to reveal the Doppler cycles due to the relativistic frequency shifts of gravity and relative motion.

The short-term performance of this system during the mission has been measured by taking the Allan variance of the 0.84-Hz redshift beats for a 5-min sample of data centered at apogee where the redshift frequency variation was very slow. These data are shown in Figure 4.

For the time intervals observed, the stability data for the masers are essentially the same as in the laboratory. It is important to note that this comparison in space included 30,000 km of transmission path and three passes through the troposphere and ionosphere. It is clear that whatever tropospheric and ionospheric propagation anomalies may have occurred were very highly correlated.

It is also evident that the electronic hardware at the ground station and probe introduced no significant uncorrelated noise during the mission, which was confirmed before the experiment was launched by several months of end-to-end ground-based simulation tests of the system.

In addition to the above data from the analog cancellation system, the one-way and two-way Doppler data were separately recorded during the flight using the U.S.B. system's tracking data processor (TDP). The data were recorded in terms of Doppler cycle counts every 0.1 sec, alternating the one-way data with the two-way data. Later processing of these data gave the same results as were obtained from the analog system. The real time analog data and à posteriori computed data are shown in Figures 5 and 6. These data represent the residual phase still to be accounted for in the 2.6×10^3 radians of redshift data. (At the time of writing, work continues on the analysis of these data.)

The analog data were taken using a 20-Hz filter. The digital uplink and downlink data from the TDP were taken using the standard system. Both data sets are computed and plotted at 1-sec intervals. The undulations in the data are primarily due to an aliased representation of phase fluctuations from the rotating payload. The fluctuations result from the slight misalignment of the dipole antenna located on the payload axis. Figure 7 shows data plotted every 0.01 sec from the analog system; here we clearly see the rotation and nutation of the payload 30 min after separation from the fourth-stage rocket motor.

From these data we can conclude that tracking system hardware can provide stability at the 10^{-14} level or better for time averages on the order of 100 sec and longer.

EFFECTS OF SIGNAL PROPAGATION THROUGHOUT THE SYSTEM

To characterize the stability of the complete Doppler system including the antennas and ray paths, we must look at all the contributions to phase instability in detail. Figure 8 shows the light time diagram for one-way and two-way ray paths connecting an Earth station and a space probe going near the Sun. Here, because of the long

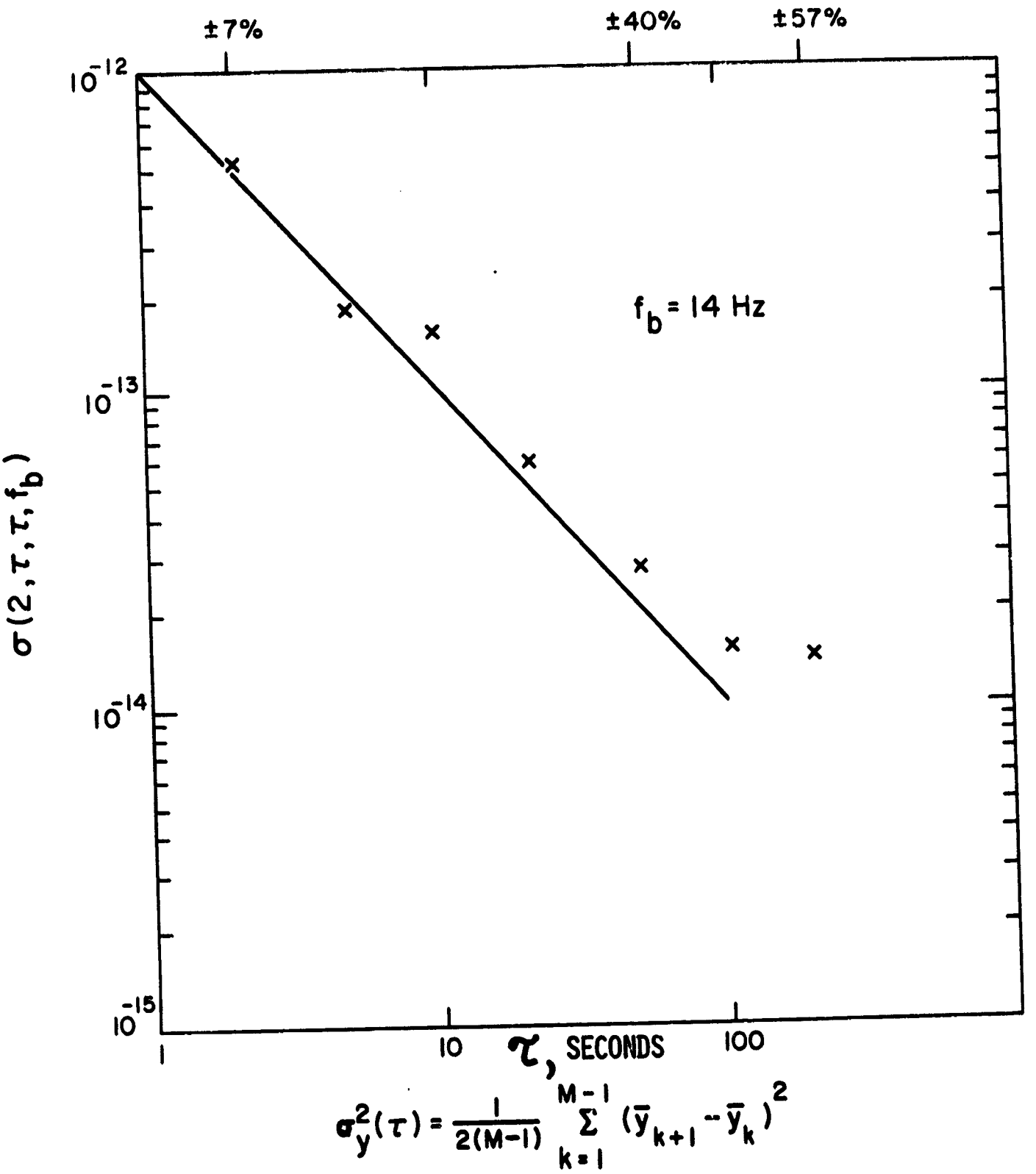


Figure 4. Allan variance of probe and ground maser comparison using 300 seconds of data near apogee.

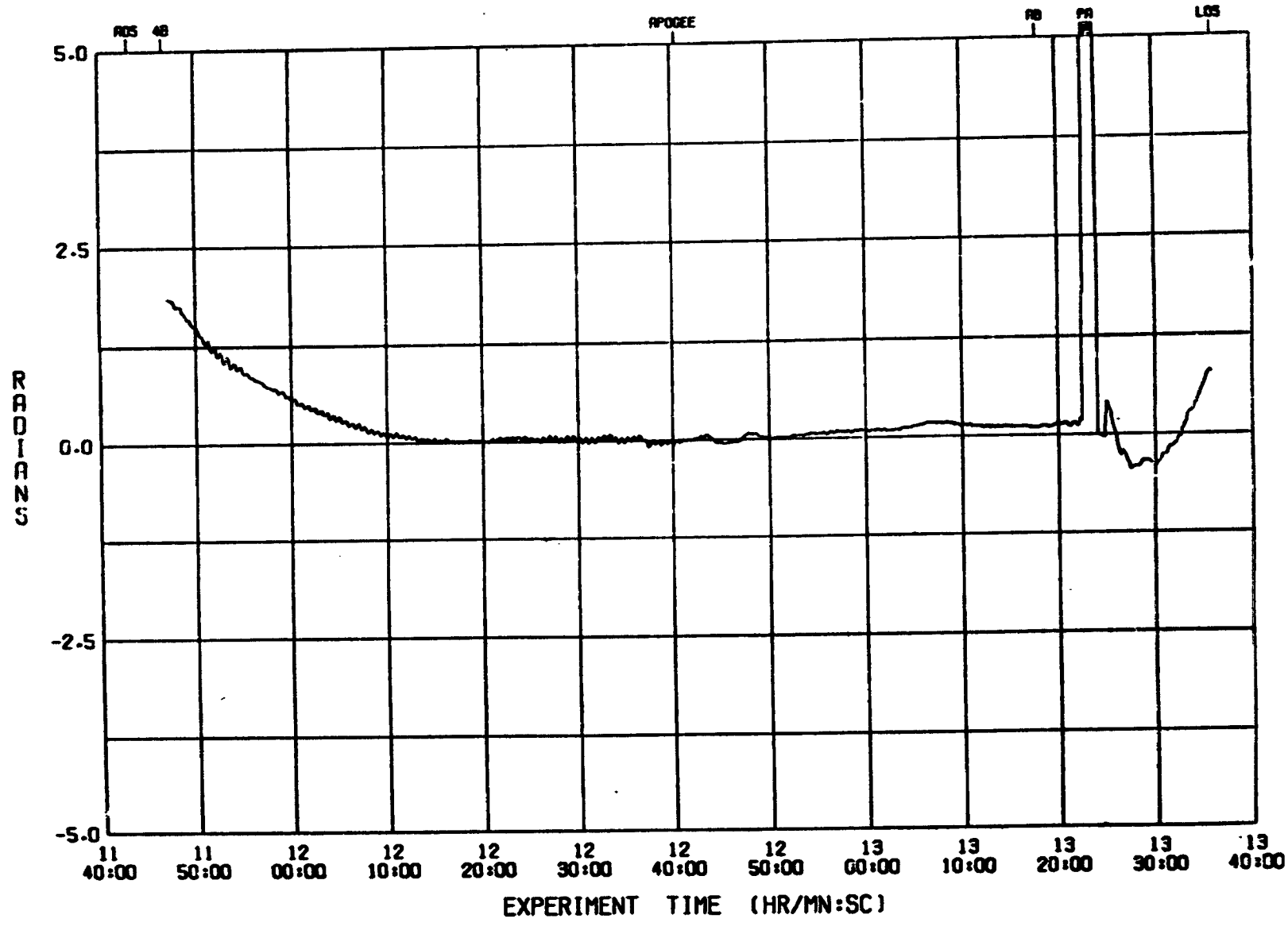


Figure 5. Redshift phase residuals from real-time analog data.

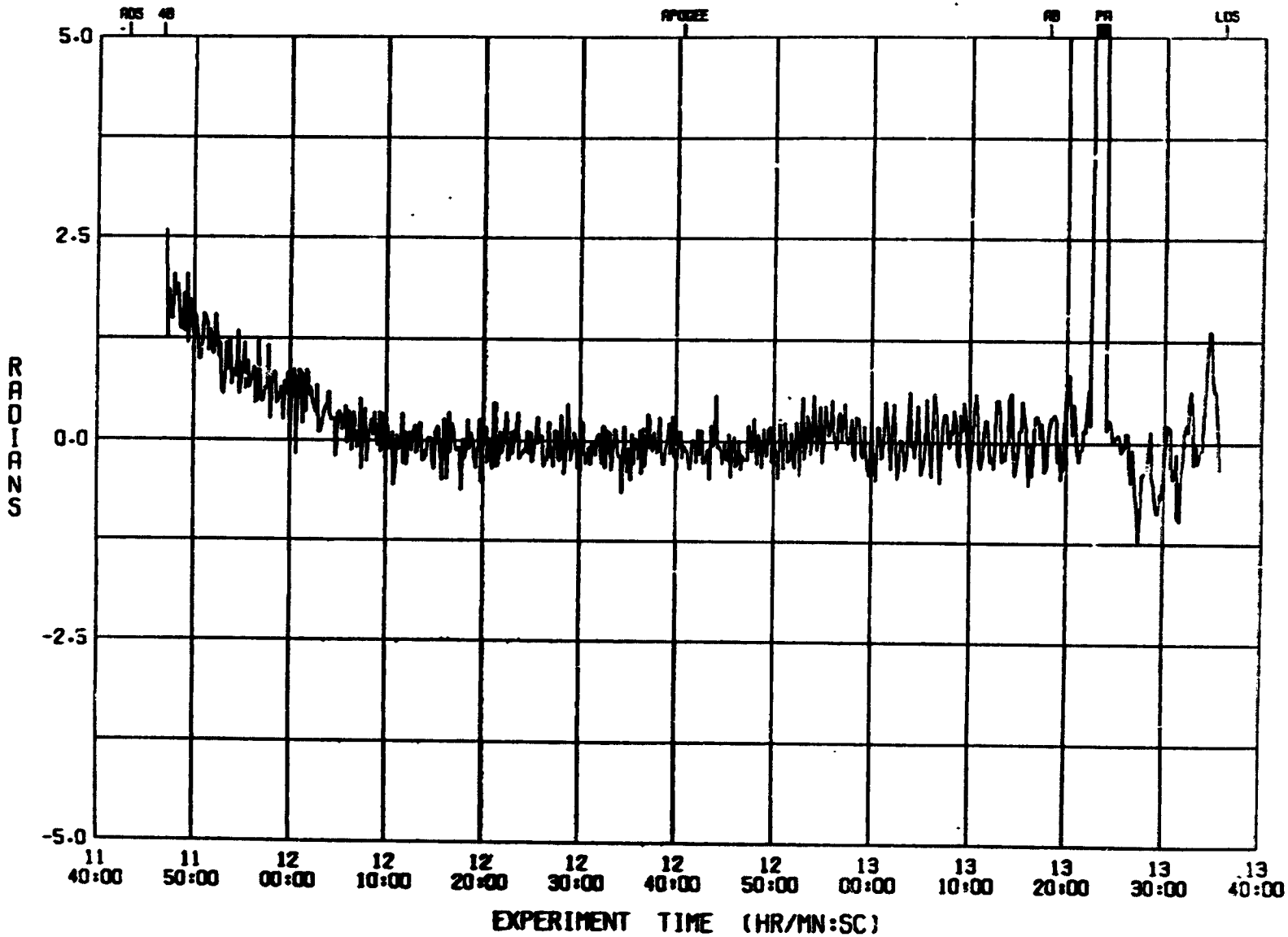


Figure 6. Redshift phase residuals from *a posteriori* computed data.

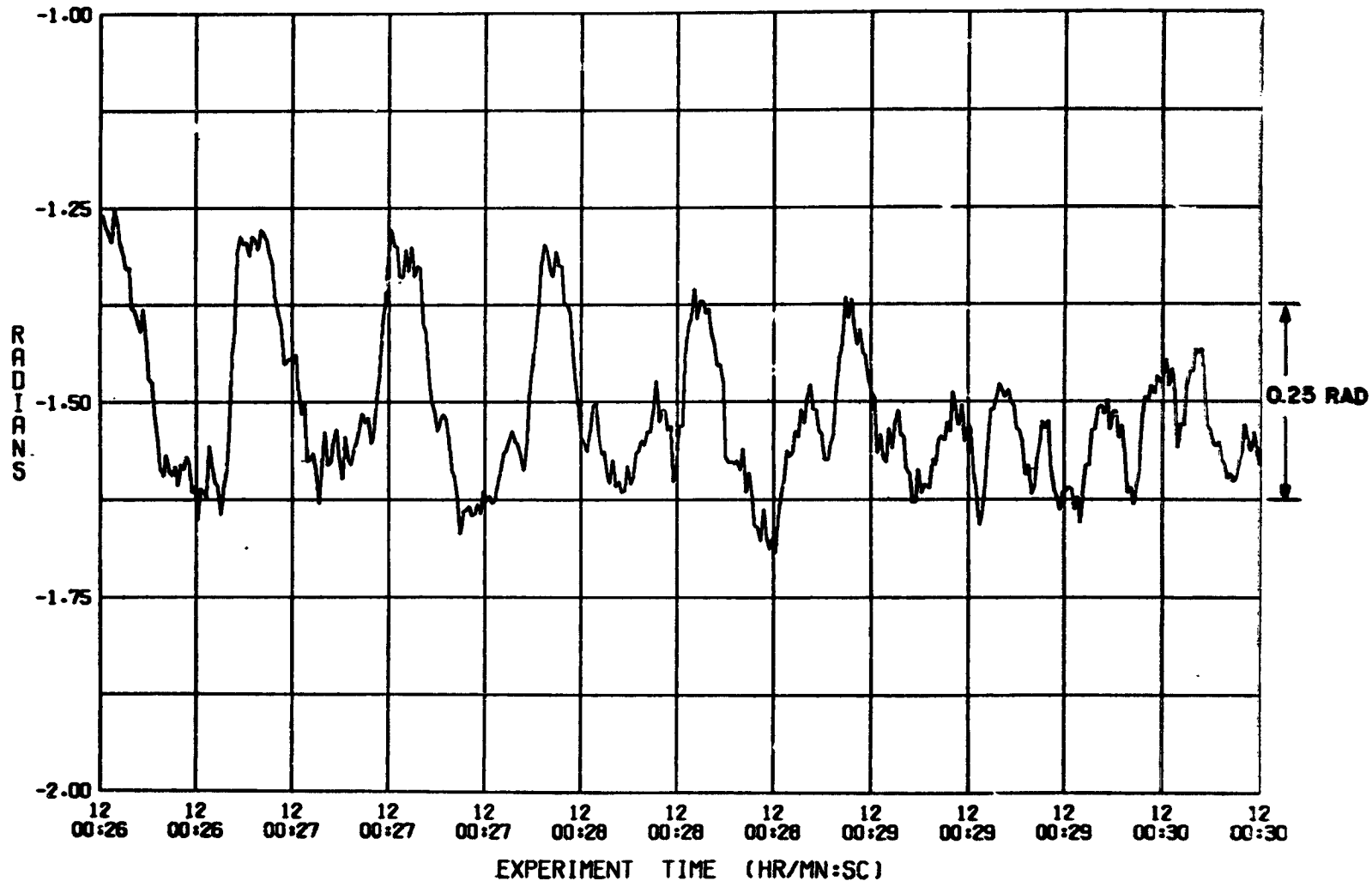


Figure 7. Redshift phase residuals from combination of data from 3 microwave links at 2.2 GHz between payload and ground station showing effects of payload nutation and rotation.

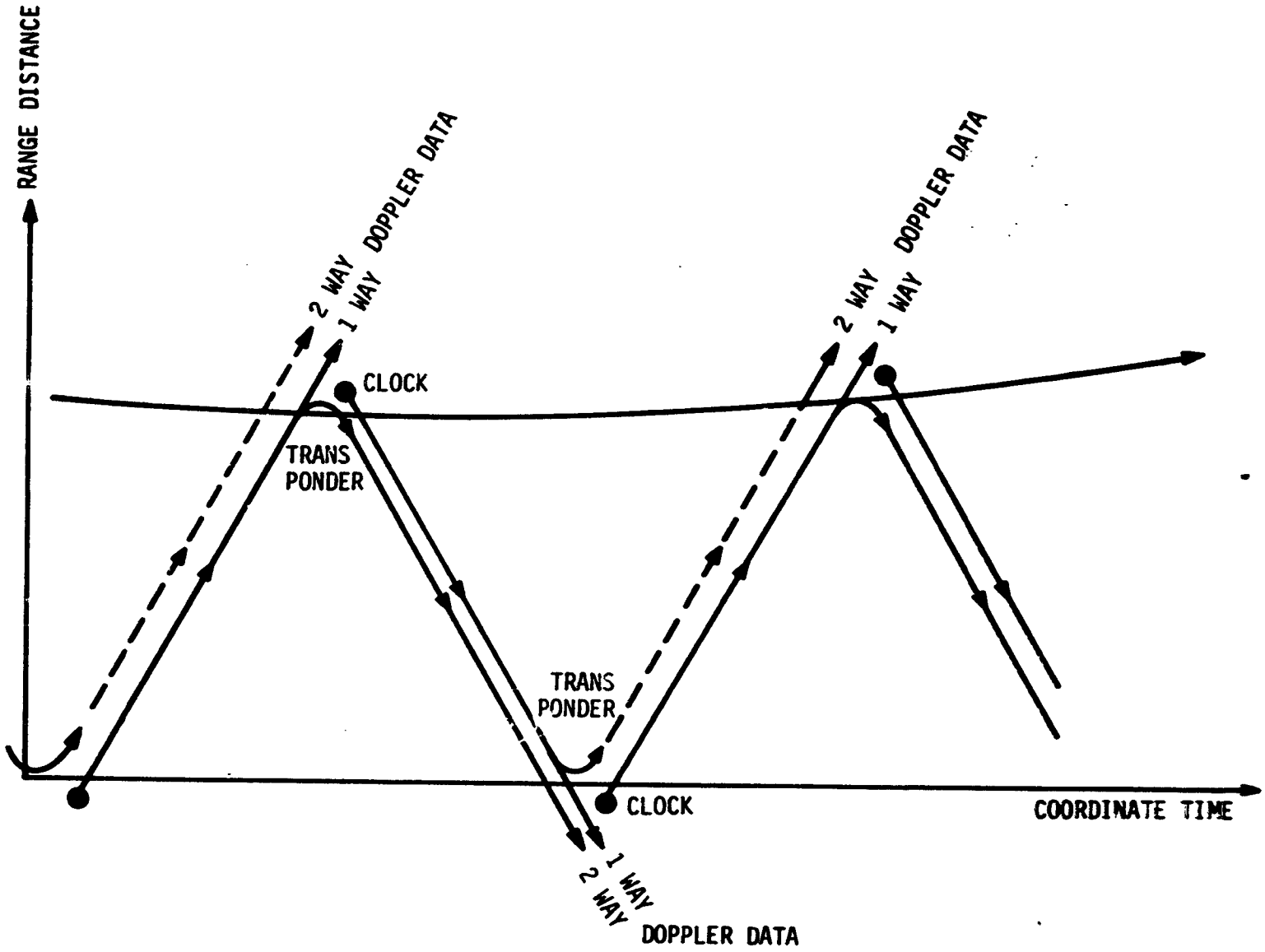


Figure 8. Light time diagram of ray paths in 4 link system. The dotted line shows link added to original gravitational redshift Doppler system.

($\tau \sim 15$ min) two-way light time for a signal transponded near the Sun, we assume that the Earth's tropospheric and ionospheric effects occur at the same time as the signals at the Earth antenna. The phase observed in a one-way Doppler system has the following contributions:

$$\Delta\Phi_1(t) = \phi_{p_c}\left(t-\frac{R}{c}\right) + \Phi_{p_c}\left(t-\frac{R}{c}\right) + \Phi_{p_A}\left(t-\frac{R}{c}\right) + \Phi_c(\downarrow) + \Phi_I(t) + \Phi_T(t) + \Phi_p(\downarrow)$$

$$+ \Phi_{G_A}(t) + \Phi_{G_{Rx}}(t) + \phi_{G_{Rx}}(t) - \Phi_{G_c}(t) - \phi_{G_c}(t) + \phi_{G_R} \quad (2)$$

Here,

$\phi_{p_c}(t)$ and $\phi_{g_c}(t)$	=	random phase noise of the probe and ground clocks, respectively.
$\Phi_{p_c}(t)$ and $\Phi_{g_c}(t)$	=	systematic phase variations of the probe and ground clocks, respectively.
$\Phi_{p_a}(t)$ and $\Phi_{g_a}(t)$	=	systematic phase contribution of the probe and ground antennas.
Φ_p	=	$\frac{2\pi f}{c}\int_p ds$ = the systematic phase change resulting from the ray path distance, $\downarrow$ represents Earthward propagation.
$\Phi_T(t)$ and $\Phi_I(t)$	=	systematic phase variation of the troposphere and the ionosphere, respectively.
Φ_c	=	the systematic phase variation of the solar corona, $\downarrow$ represents Earthward propagation direction.
$\Phi_{g_{Rx}}(t)$	=	the systematic phase variation in the ground receiver system.
$\phi_{g_{Rx}}(t)$	=	the random phase variation in the ground receiver system.
$\phi_{g_r}(t)$	=	the random phase contributed to the data by the ground station recording system.

For a two-way system, it is more complicated and we have the following:

$$\Delta\Phi_2(t) = \phi_{G_c}\left(t-\frac{R_1+R_2}{c}\right) + \Phi_{G_c}\left(t-\frac{R_1+R_2}{c}\right) + \Phi_I\left(t-\frac{R_1+R_2}{c}\right)$$

$$+ \Phi_T\left(t-\frac{R_1+R_2}{c}\right) + \Phi_c(\downarrow) + \Phi_p(\downarrow) + \Phi_{p_A}\left(t-\frac{R_2}{c}\right) + \Phi_{p_{TR}}\left(t-\frac{R_2}{c}\right)$$

$$+ \Phi_{p_{TRx}}\left(t-\frac{R_2}{c}\right) + \Phi_{p_a}\left(t-\frac{R_2}{c}\right) + \Phi_p(\downarrow) + \Phi_c(\downarrow) + \Phi_I(t) + \Phi_T(t)$$

$$+ \Phi_{G_A}(t) + \Phi_{G_{Rx}} + \phi_{G_{Rx}}(t) - \Phi_{G_c}(t) - \phi_{G_c}(t) + \phi_R(t) \quad (3)$$

Here, in addition to the phase perturbation terms identified earlier, two terms relate to the transponder, namely, $\Phi_{p_{TR}}(t)$, which is the systematic variation of the signal phase through the transponder, and $\phi_{p_{TRx}}(t)$, which is the random component

associated with the noise comprising the thermal and excess thermal noise of the transponder receiver. Note that this receiver is also the one-way clock link receiver described in the following system.

THE CORRELATED FOUR-LINK SYSTEM

From this list of one-way and two-way phase noises and the times at which these are reported at the various station recorders, we can identify common sources of systematic behavior described earlier for the redshift probe experiment. We now propose to go one step further and describe a system that is completely symmetrical in that we have one- and two-way data obtained both at the probe and at the Earth station. This is done by transponding the probe clock down-link signal back to the probe as shown in Figure 9.

One-way and two-way Doppler cycle data are recorded with respect to a proper time scale defined by the probe and ground clocks and later are appropriately corrected for relativistic effects of their relative time scale. These data are to be time correlated, and, as will be seen, the correlation process will permit the identification of systematic processes that occur either at the probe or at the Earth station. In addition to these, it is possible to identify fluctuations in propagation effects that occur between the stations. All the fluctuations discussed here occur on a time scale commensurate with or shorter than the light time, R/c.

We designate the four signal links as follows:

α Earth clock uplink
β Earth transponder uplink
γ Probe clock downlink
δ Probe transponder downlink.

Figure 9 can be configured as shown in Figure 10 to localize the origin of random and systematic phase variations. We can identify some time-related properties of this system and in the following discussions we will assume the following:

1. The clock controls the transmitter and also serves as receiver local oscillator.

2. The antenna serves all four links.

3. The receiver R. F. system serves both input links and the systematic R. F. phase variations are common to both signals.

4. The ϕ_c, ϕ_{Rx}^{γ}, ϕ_{Rx}^{δ}, ϕ_R^{γ}, ϕ_R^{δ} are random and uncorrelated.

   Note that the clock one-way receiver and the transponder receiver are the same device in this representation of the system.

5. The transmitter is phase locked to the clock. The microwave signal is sampled at the antenna feed so that the phase lock loop spans the complete microwave system.

6. Earth ionospheric and tropospheric phase variations are assumed to be equivalent to antenna motion since the variations occur with about 0.03 sec of the light time to the Earth antenna.

Now if we consider that the one-way Doppler phase accumulation designated by $\Delta\Phi_{G_1}(t)$ and the two-way phase designated by $\Delta\Phi_{G_2}(t)$ are simultaneously received, we

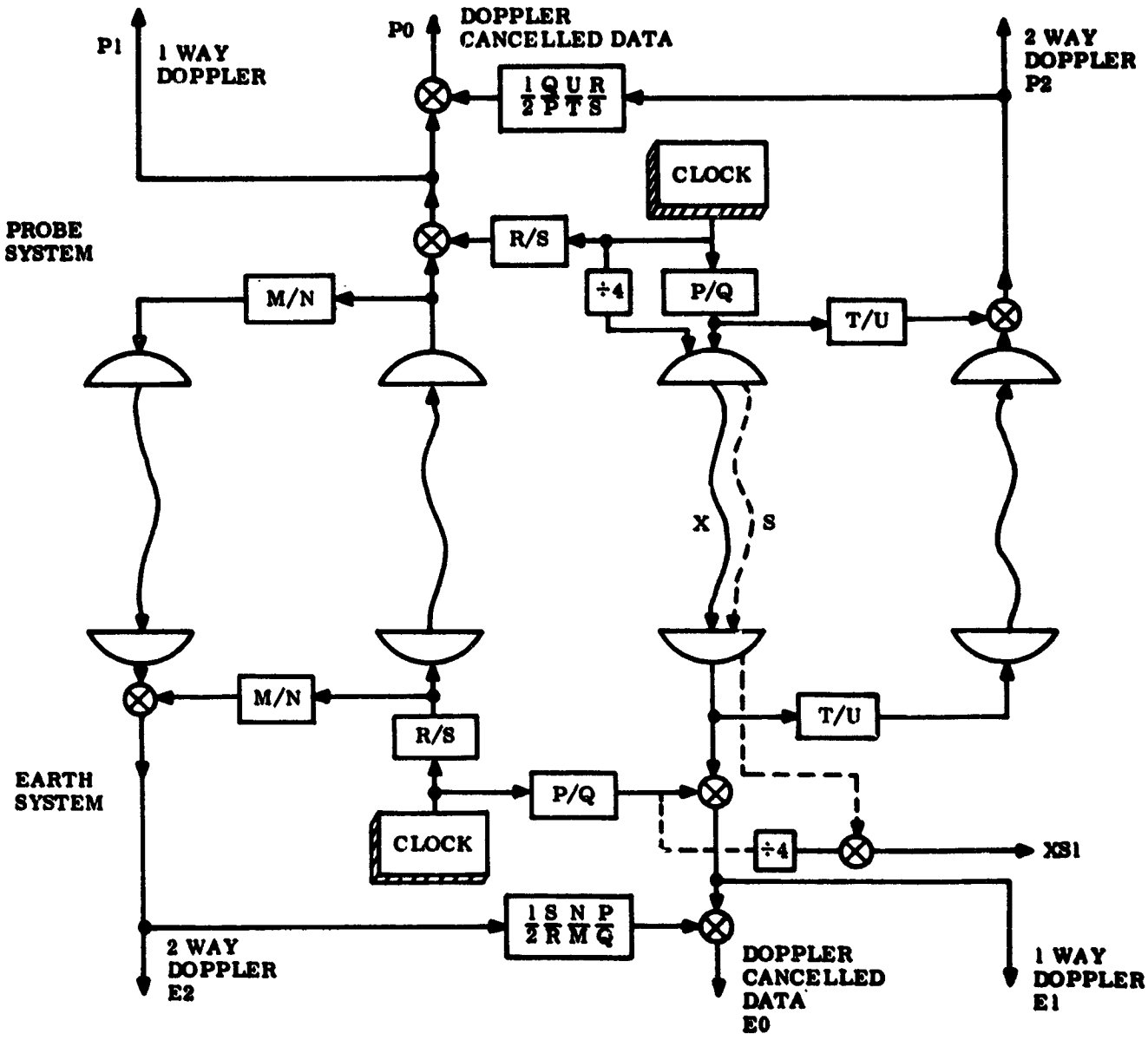
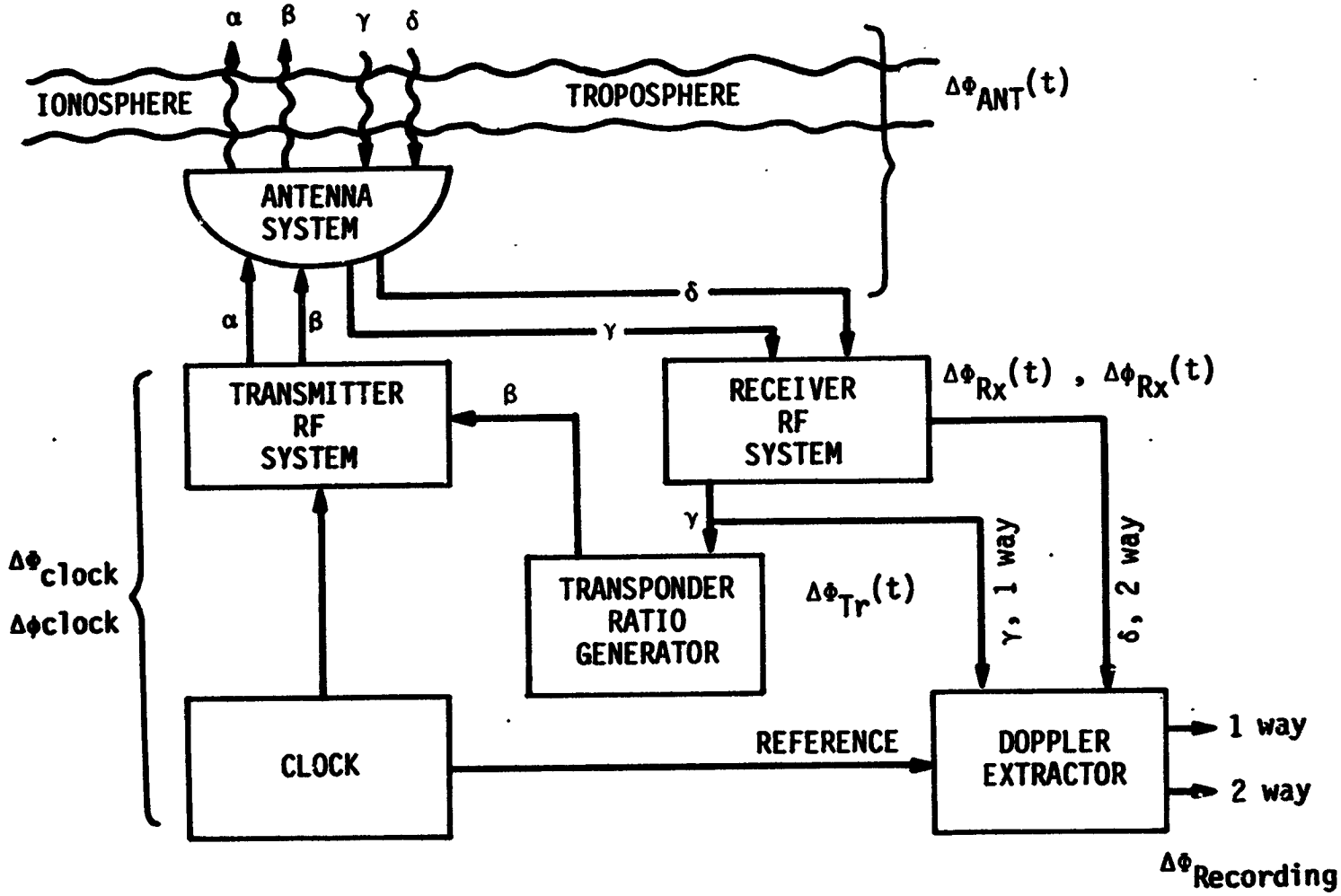


Figure 9. Dual one-way and two-way Doppler tracking and Doppler cancellation system similar to the system used in the redshift rocket probe experiment.



$\Delta\Phi$ systematic phase variation

$\Delta\phi$ random noise

Figure 10. Origin of systematic and random phase variations at terminals of the 4 link system.

can perform a real-time subtraction and obtain the one-way Doppler phase accumulated at the probe terminal delayed in its time representation by $\tau = R/c$.

The observed phase after subtraction is given by

$$\Delta\Phi_{G_2}(t) - \Delta\Phi_{G_1}(t) = \Delta\Phi_{p_1}\left(t - \frac{R}{c}\right) + \phi_{p_{TRx}}(t - R/c) - \phi^a_{p_{Rx}}(t - R/c) - \phi^a_{p_R}(t - R/c)$$
$$+ \Phi_{p_{TR}}\left(t - \frac{R}{c}\right) + \left[\phi^a_{G_{Rx}}(t) - \phi^{\gamma}_{G_{Rx}}(t) + \phi^a_{G_R}(t) - \phi^{\gamma}_{G_R}(t)\right] . \quad (4)$$

We see that the only systematic term that exists in this representation of the one-way data at the probe is the transponder systematic phase variation, $\Phi_{p_{TR}}$. The terms $\phi_{p_{TRx}}$ and $\phi_{p_{Rx}}$ are the same (as mentioned earlier). The term $\phi^a_{p_{Rx}}(t - R/c)$ is shown with negative sign to cancel the same term that belongs in $\Delta\Phi_p(t - R/c)$; this recording noise contribution is replaced by the two ϕ_{G_R} terms, which combine as random variables. The noise term of the probe receiver is replaced by $\phi^a_{G_{Rx}}(t) - \phi^{\gamma}_{G_{Rx}}(t)$, which also combine as random variables.

We see that we can obtain two representations of the one-way signal to the probe and these contain uncorrelated receiver and recording noise. Except for the transponder systematic phase term, all systematic effects are identical in the two representations of the one-way data.

In our test of the redshift, the phase stability of the transponder was of concern and considerable effort was devoted to keeping the temperature of the U.S.B. transponder as constant as possible by use of a newly invented ammonia boiler that acted as a heat sink for the transponder and clock transmitter box. In addition, precise phase calibrations were made of the transponder and clock transmitter as a function of temperature, and temperature data were telemetered to the Earth station for later use in making small a posteriori correction to the data.

With the exception of the transponder systematic term, it is possible to reduce the phase noise in the one-way data to the probe by combining the two data sets. From symmetry, we see that the same can be done in the corresponding set of one-way data obtained at the Earth station.

TIME CORRELATIONS IN THE DATA

From inspection of the one-way and two-way phase variations, we are led to some interesting correlations that will identify the origin of several types of phase variations occurring in times commensurate with $\tau \sim R/c$.

In the following representation of phase variations we assume that the systematic cycle counts given by the ray path integral $\Phi_p = \frac{2\pi f}{c}\int_p ds$ have been fitted and removed

from the data and that we are looking at the residual phase variations. The process of identifying and quantitatively removing in a systematic manner as many systematic variations as possible will allow us to iterate the process and obtain an improved representation of the ray path integral, which is the information we seek in Doppler tracking applications.

The identification of the origin of phase variations is very important in the search for gravity waves, and, as will be seen later, the four-link system provides a uniquely determined set of 10 pulses if a gravity wave with a period less than $\tau \sim R/c$ traverses the Earth-probe system.

Oscillator shifts have two effects: (1) they cause a spurious phase shift in the observed data, and (2) they can cause a timing error in the correlation process. We presume that the Earth station has many clocks so that the systematic behavior of the station oscillator is well understood from measurements with an ensemble of oscillators compared on Earth via satellite, Loran, and traveling clock techniques. Clearly this is not possible for the spacecraft; other means of comparison are available, however.

Short-term, $\tau \sim R/c$, variations in the behavior of the oscillator can be observed by data at time intervals associated with the ray path distances in the various links. For example, if the Earth station oscillator were to suffer a momentary shift upward in phase at $t = t_0$, we see from Figure 11 that the phases of the one-way and two-way Doppler signals are decreased at the same time and by the same amount. At time $t = t_0 + R/c$ the probe receives the signal and sees a momentary increase in the Doppler phase. Later at time $t = t_0 + 2\,R/c$ the previously transmitted momentary increase in transmitted phase is observed by the now restabilized local oscillator.

The clock perturbation occurs as a well defined set of phase variations spaced in time by the one-way light time R/c and the two-way light time $(R_1 + R_2)/c$. Note that there is no signature in the two-way probe Doppler data and that the sign of the effect depends on whether the perturbed oscillator serves as transmitter or as local oscillator. In this and all other examples we assume that in the heterodyne receiver the local oscillator frequency lies below that of the carrier.

Spacecraft antenna motion or buffeting of the whole spacecraft produces a set of correlated events in the four-link data set as shown in Figure 12. Here we assume that the spacecraft antenna suddenly moved away from the Earth at coordinate time t_0, so as to cause a momentary decrease in the observed Doppler frequency in the one and two-way signals received by the probe. The one-way data received at the Earth station at time $t_0 + R_0/c$ show the same momentary decrease in frequency and the Earth station two-way data will show twice this shift also at $t_0 + R_0/c$.

The antenna motion effects show us a different set of phase variations from those of the clock. These are spaced at time intervals R_1/c and $(R_1 + R_2)/c$. The two-way Earth data show double the amplitude of the one-way data and the phase of the signature is preserved in all links. We note that if a drag-free system is on board and operating, we can distinguish spacecraft motion of gravitational origin from perturbations from some other force. This is because the proof mass of the drag compensation system responds to gravity in the same way as the payload. This allows us to identify whether the disturbance is of gravitational origin or is due to external factors such as fluctuating solar wind or light intensity.

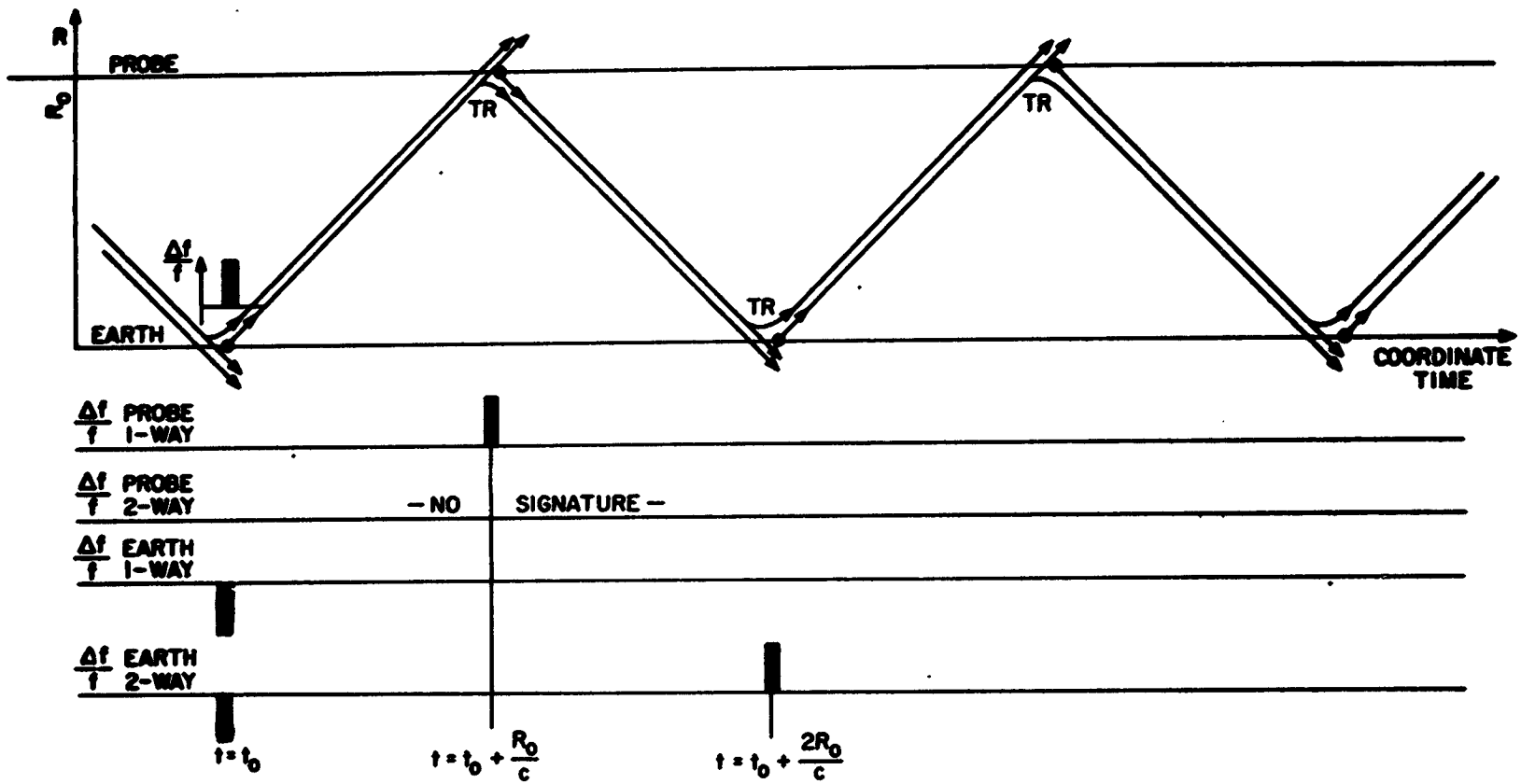


Figure 11. Effect of momentary frequency offset of ground station clock at $t = t_o$ as observed at the 4-link correlated system.

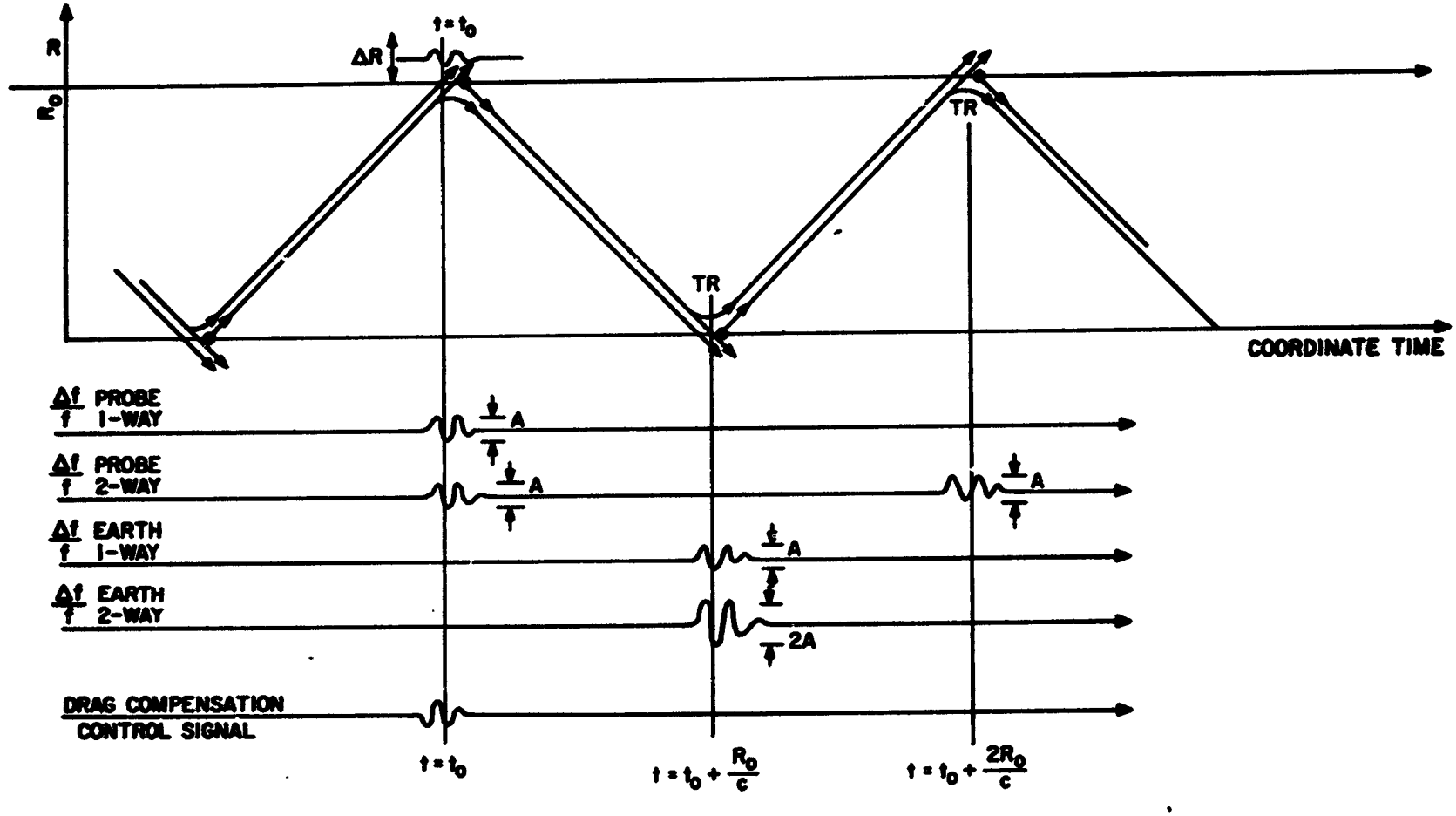


Figure 12. Effect of spacecraft buffeting by solar-wind pulse or fluctuating solar radiation pressure signature for burst arriving at $t = t_o$.

SOLAR CORONA EFFECTS ON SIGNAL PROPAGATION

The most serious problem for Doppler tracking of the probe at high accuracy and for the detection of gravity waves results from the ionization in the solar corona. Several speakers at this conference have described various versions of this effect, which can cause large, unpredictable, and rapid variations in the Doppler signal as the ray passes through regions of space where its columnar electron density varies rapidly with time.

In our approach to this problem we avoid to the greatest possible extent any statistical representation of the solar coronal behavior. This is not to deny the usefulness of such representations made from sets of long-term data. However, for the 19 hours or so that the probe is within the latus rectum of its trajectory, we believe that a statistical representation of fluctuations that can occur in time intervals on the order of 15 min is very risky.

If we had properly well-behaved noise processes, then the redundant data might only provide an improvement by $\sqrt{2}$. For the solar mission, the noise on the signal is not well behaved, and the extra data links can give us information relating to the origin of the noise bursts and assist us in establishing the cause of the phase irregularities seen in the data. The value of this information is particularly high when we search for gravity waves or small scale irregularities in the solar gravitational potential and goes far beyond the $\sqrt{2}$ factor assigned to a purely statistical interpretation of the data.

We believe that it may be possible by time-correlation methods to remove a significant part of the coronal phase fluctuations occurring in time intervals $\tau \sim R_0/c$, and, for the purpose of gravity wave detection, to identify and follow in space-time the behavior of blobs of ionization that cross the ray paths between the probe and Earth station.

The effects of electrons and the magnetic fields in which they are situated are described by the Appleton-Hartree equation, which, when developed in a power series expansion in terms of frequency, gives us considerable insight into what the problems are. Using the expansion method given by Tucker and Fannin (7), we have the following expression for the phase of the signal

$$\Phi(t) = -\frac{2\pi f}{c}\int_g ds - \frac{2\pi f}{c}\int_g ds\, M(\vec{\xi}, t) + \frac{K_1}{f}\int_g ds\, N(\vec{\xi}, t)$$
$$+ \frac{K_2}{f^2}\int_g ds\, O(\vec{\xi}, \vec{H}, N, t) + \frac{K_3}{f^3}\left\{\int_g ds\, P(\vec{\xi}, H^2, N^2, t) - Q\left[\int_g ds\, n(\vec{\xi}, t) - \int_p ds\, n(\vec{\xi}, t)\right]\right\} \quad (5)$$

Here, $\vec{\xi}$ is the position vector, and g(t) is the nondispersed line-of-sight, or geometric, ray path. M is the nondispersive refractive index given in terms of position, $\vec{\xi}$, and time, t. N is the electron density at $\vec{\xi}$ and t. O is a function of $\vec{\xi}$, t, magnetic field H, and N. P is a function of $\vec{\xi}$, H^2, N^2, and N; n is the refractive index of the electron cloud at $\vec{\xi}$ and t; and P is the actual, refracted ray path.

From this approach we learn that the chief effect in the phase delay varies as $f^{-1} \times$(columnar electron density). The term that depends on f^{-2} involves the magnetic field and describes the rotation of the plane of polarization. This is a nonreciprocal effect in that the rotation depends on the relative direction of the wave propagation vector and the magnetic field. The term in f^{-3} is complicated because it applies to higher order delay terms involving both the electron density and the square of the magnetic field in addition to the path length differential from ray bending.

This last effect can be serious at the very high electron density encountered for ray paths going near the Sun. At this time it may be prudent to keep in mind the possibility that the ray paths going past the Sun or originating near the Sun can be deflected substantially and that the deflection will depend on the signal frequency.

The columnar electron density is measured by dual frequency techniques. In Figure 9 we show an S-band signal that is phase coherent with the X-band probe clock down-link signal. The output of this signal labeled XS1 gives the one-way S-band Doppler information, which, when compared to the X-band output, E1, will provide the phase difference

$$\Delta\Phi_{XS}(t) \doteq \left(\frac{1}{f_1} - \frac{1}{f_2}\right) K_1 \int_P ds\, N(\vec{\xi}, t)$$

and a measure of $\int_P ds\, N(\vec{\xi},t)$, provided that P is the same for the S- and X-band rays.

We expect that dense, rapidly time-varying electron clouds in the ray path will deflect the rays at the two widely separate frequencies sufficiently to prevent an accurate short-term measurement of the electron density in the actual ray path whose Doppler frequency is to be corrected. However, for large scale, relatively uniform electron clouds, two-frequency data are essential for establishing the average level or bias in the phase data.

The effects of the Earth's ionospheric and tropospheric fluctuations, as mentioned earlier, are included in the fluctuations attributed to the Earth antenna. The ionospheric phase bias is included in the columnar electron density measurement made by use of the two-frequency method.

The tropospheric bias is not so easily accounted for and must be included in the interpretation of the Doppler data to obtain accurate range information. Several models (8) are available that include the effect of barometric pressure and humidity. It may even be necessary to make measurements of the tropospheric delay during the tracking of the Solar Probe by observing signals from Earth satellites that provide ray paths that are nearly collinear to those from the probe. The global positioning system now under deployment (ultimately involving 24 Earth satellites each including atomic clocks that transmit time signals) may be very useful for this purpose.

Correlation and pattern recognition can be used to cope with the short-term ray path variations. However the problem is not so simple as in the cases previously described where the disturbances are located at one end or the other of the transmission path. Figure 13 attempts to describe the effect of time-varying blobs of ionization crossing the ray path. As pictured in the figure, the scale of these disturbances is small compared to the light time, yet large enough to corrupt the accuracy of the data.

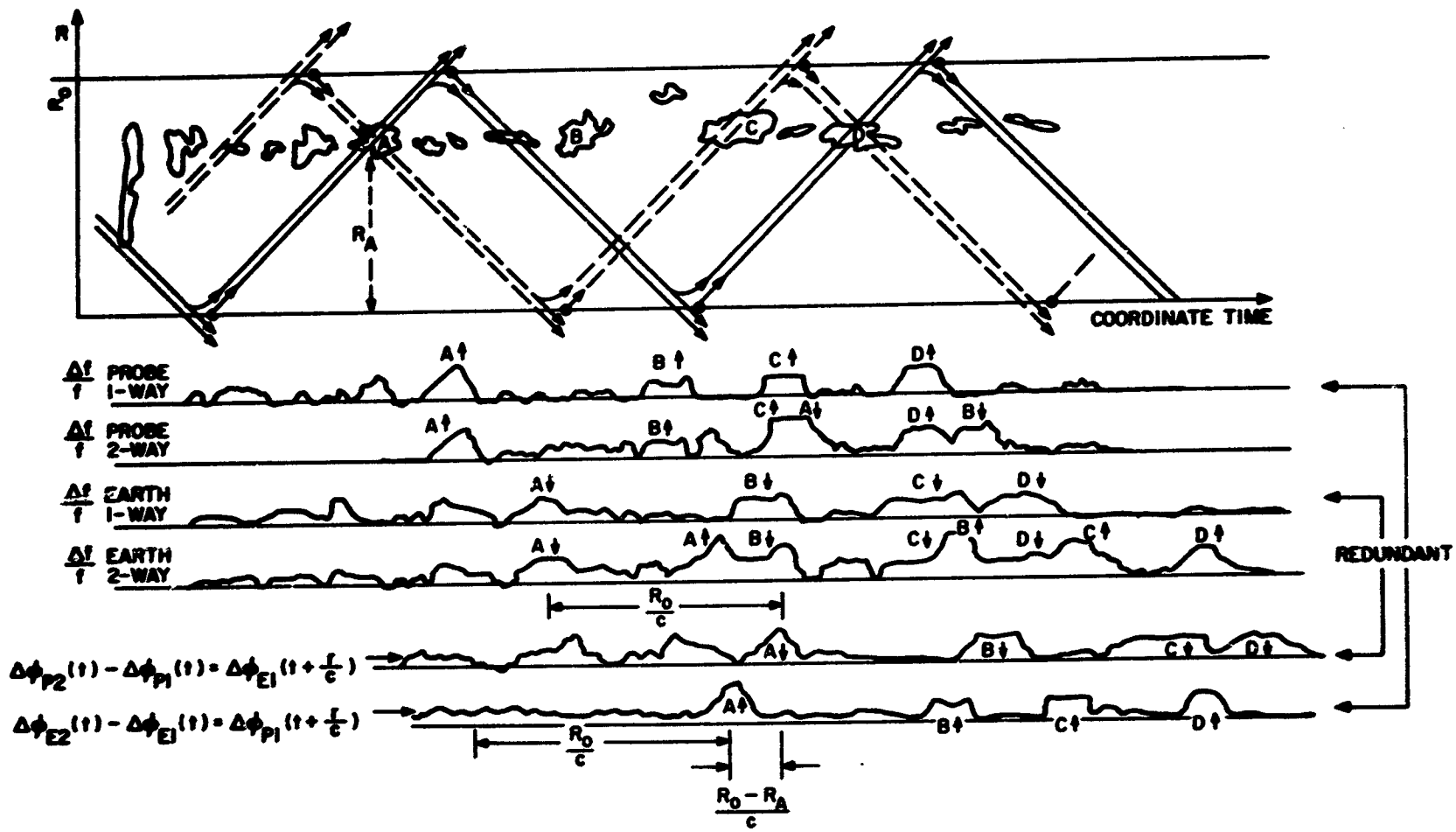


Figure 13. Illustration of correlation technique to remove effect of phase fluctuation in propagation. Redundant data are used to remove uncorrelated noise and correlated $\uparrow$ and $\downarrow$ fluctuations are used to determine the $\frac{R_o - R_A}{c}$ position of the propagation anomaly in the R,t plane.

The time of closest approach of the ray to the Sun is R_0/c, where the blobs are assumed to be the most densely localized. Here, owing to the rapid motion of the probe, spatial variations in the electron density will be relatively rapid, and the discrepancy of multi-frequency columnar density measurements over short intervals (say 10 min) can be expected to be significant as will also be the changes in the rotation of the plane of polarization of the signals. Since in the projected Solar Probe Mission the spacecraft antenna is not spinning, we can obtain, at least at the ground station, some measure of the rotation of the linearly polarized signals from the probe and thus obtain a measure of the magnetic field in the disturbance. This information would be used to correct the phase of the Doppler signals for Faraday rotation effects included in the 1/f term of Equation 5.

Figure 13 describes the correlation of phase fluctuation in rays that cross a region of space time where there is a time-varying electron density. It is clear that if the cloud has a very asymmetrical shape, the phase delay in the up and down directions may not be similar and the phase signature represented in the uplink and downlink data will differ.

This illustration also shows the way in which redundant one-way data can be obtained so that the random-level uncorrelated fluctuations can be reduced. By cross-correlation techniques, using the probe and Earth one-way data, we can estimate the delay $(R_0 - R_a)/c$ associated with the position of the most prominent blobs in the ray path and make an estimate of the magnitude of the contribution to the phase delay.

The information on magnitude and position of the origin of the phase fluctuations can be very helpful in allowing us to differentiate propagation fluctuations from fluctuations that result from the gravitational behavior of the probe. It is also very likely that the behavior of the short-term variations in the highly ionized regions of the ray path will be of interest to those who study the solar corona.

The system in Figure 9 includes a real-time Doppler cancellation system, which subtracts one-half the two-way Doppler signal from the one-way Doppler signal. In the short-range rocket experiment, this system effectively cancelled the gross propagation effects and revealed the relativistic terms including the effect of acceleration of the ground station owing to the Earth's rotation. For the solar probe scheme, the system will serve to cancel the gross Doppler frequency at about 10 MHz and provide a means for monitoring in real time both the one-way and two-way signal links. It is of interest to note that the gravitational effect of the Sun, besides causing a shift toward lower frequencies as the clock falls toward the Sun, will cause an additional shift of comparable magnitude to lower frequencies resulting from the second-order Doppler effect of the relative motion of the Earth and probe clocks. At the 8-GHz carrier frequency, this amounts to about 8.5 KHz as the probe reaches perihelion.

While the system described above is meant to obtain as much information as possible using the Doppler technique, one must remember that there is also information available in making range measurements by modulation of the carrier signals and time correlation of the transmitted and received signals. Range data so obtained can be used to obtain the relatively slowly varying effects due, for instance, to the solar quadrupole moment, J_2. The operation of the four-link system described in no way precludes the use of such modulation and the gathering of additional information in the range mode.

The goal of this system is to provide a means of obtaining range and range-rate measurements that do not depend on long-term averages of fluctuations and the consequent limitation in the number of uncorrelated data points. This is particularly important for a short-term mission going very close to the Sun.

GRAVITY WAVES AND SOLAR OSCILLATIONS

The structure of the Sun can provide a number of gravitational signatures. The obvious objective is the measurement of J_2. In addition there is the possibility of observing the gravitational effect of 2- to 6-hr, oscillations of the Sun that have been reported from optical measurements (9). Furthermore, as K. Nordtvedt has observed (10), given tracking data at the 10^{-14} level and very accurate drag compensation, we can observe the effect of frame dragging owing to the Sun's angular momentum. If we are sufficiently confident in the theory that describes this effect we can use it to measure the solar angular momentum. This, in conjunction with the shape of the Sun's static gravitational potential will go a long way toward unraveling the mysteries of the Sun's interior.

Finally we look at the quest for gravity waves and the possibility of detecting the Doppler shifts of signals between spacecraft as described by Thorne, Wahlquist, Estabrook, Davies, and Braginsky using a transponded microwave tracking signal (11, 12, 13). Figure 14 shows the behavior of the four-link system for a gravity wave propagating at a 60° angle to the Earth-probe vector. In it we see the usual Earth two-way signature given in earlier references. In addition, we have other manifestations of the gravity wave showing as two pulses in the one-way systems and a reciprocal set of data taken aboard the probe that have a unique time relation with the Earth data. Figures 15 and 16 show the results of gravity waves crossing the Earth probe vector at 90° and at 30° and illustrate the amplitude time relationship of signals in the four Doppler links.

THE APPLICATION OF THE TECHNIQUE TO THE SOLAR PROBE MISSION

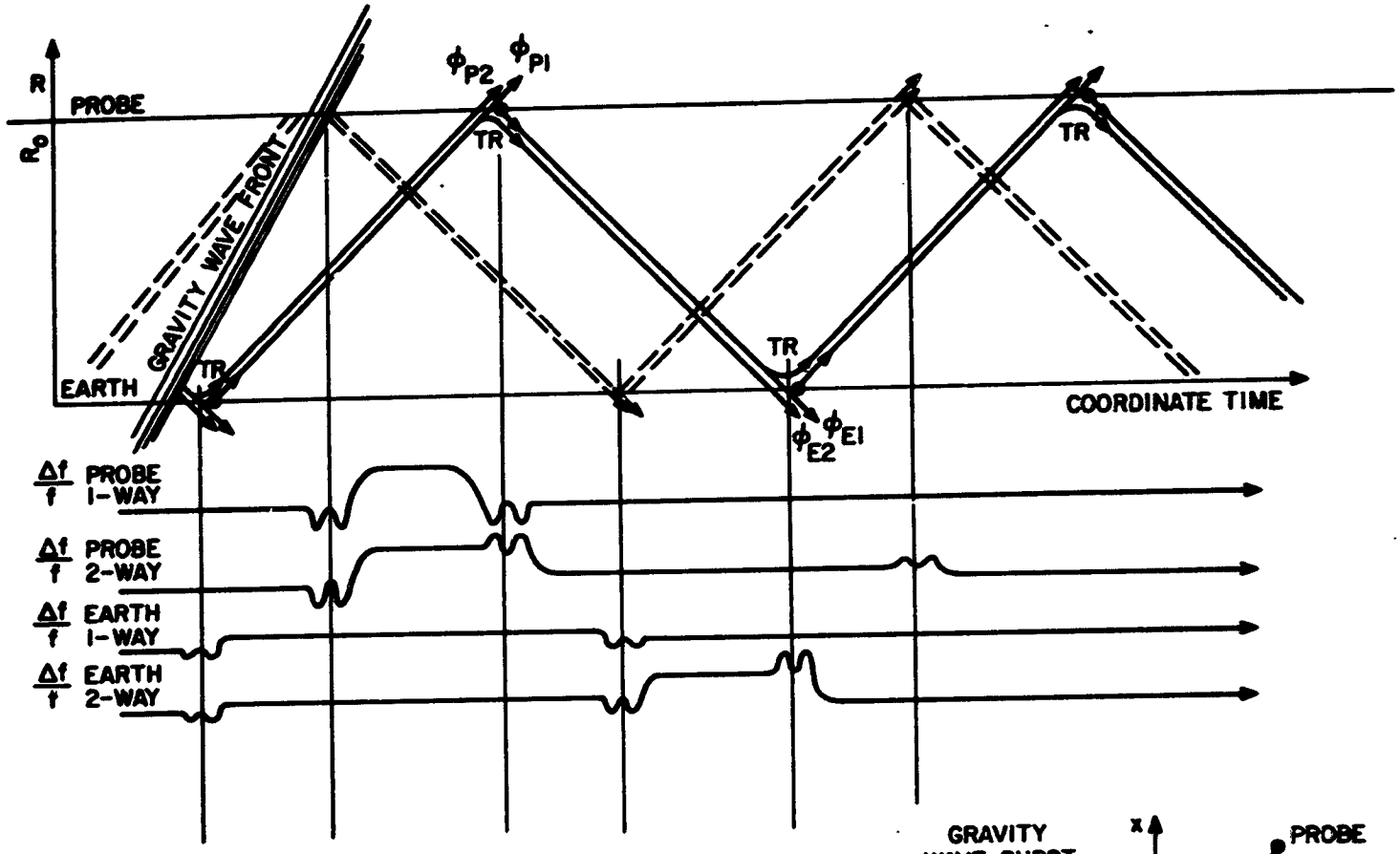
There are obviously a large number of objectives for a mission that would go near the Sun and it is likely that these objectives will not all be compatible in a single spacecraft. It seems clear to us that the measurement of J_2 of the Sun will necessitate a drag compensation system and that the measurement will require accounting for the relativistic parameters β and γ. At present, there does not appear to exist a clear consensus as to the relation between tracking accuracy (either in position or velocity), the level of drag compensation, and the statistical methods used to evaluate the results of such a mission. We expect that a clarification of this problem will soon result from further study.

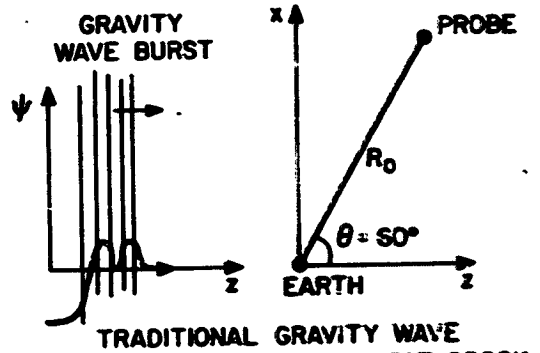
HARDWARE CONSIDERATIONS

Summary of Telecommunications Requirements

The Solar Probe requires telecommunications support for the following:

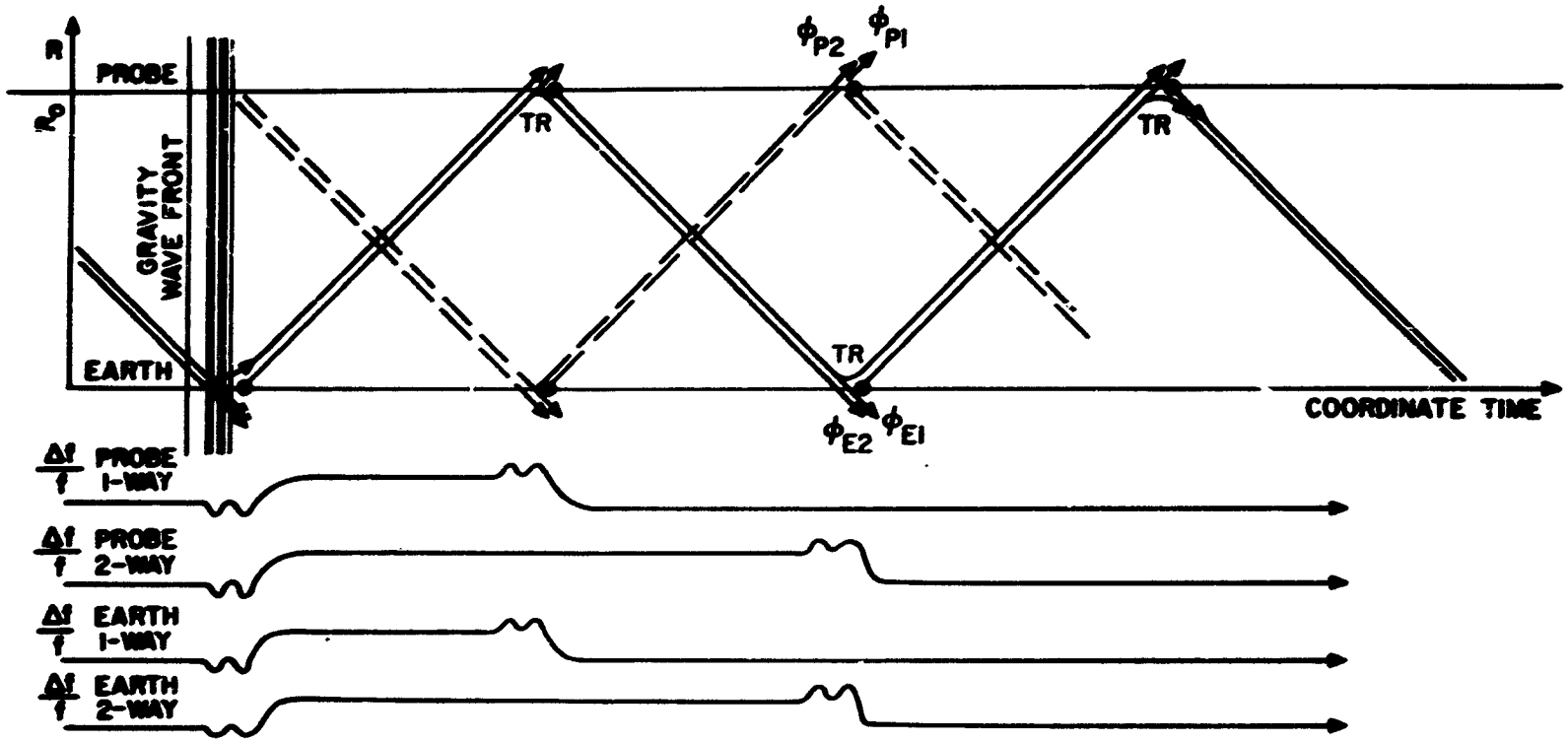
(a) Routine spacecraft tracking and navigation,
(b) Command functions,
(c) Telemetry.





TRADITIONAL GRAVITY WAVE
AFTER THORNE, WAHLQUIST, ESTABROOK, BRAGINSKY (1975-76)

Figure 14. Illustration of gravity-wave signature for a 4-link correlated system θ = 60°.



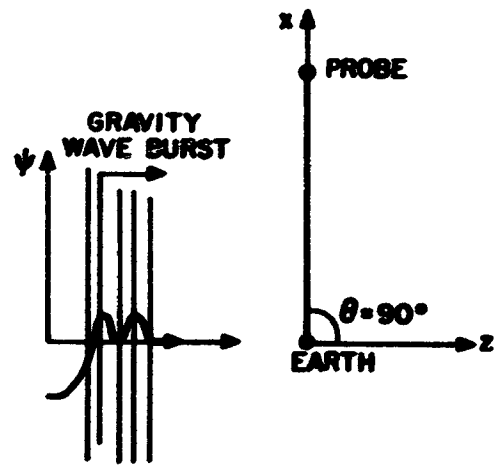


Figure 15. Illustration of gravity-wave signature for a 4-link correlated system $\theta = 90°$.

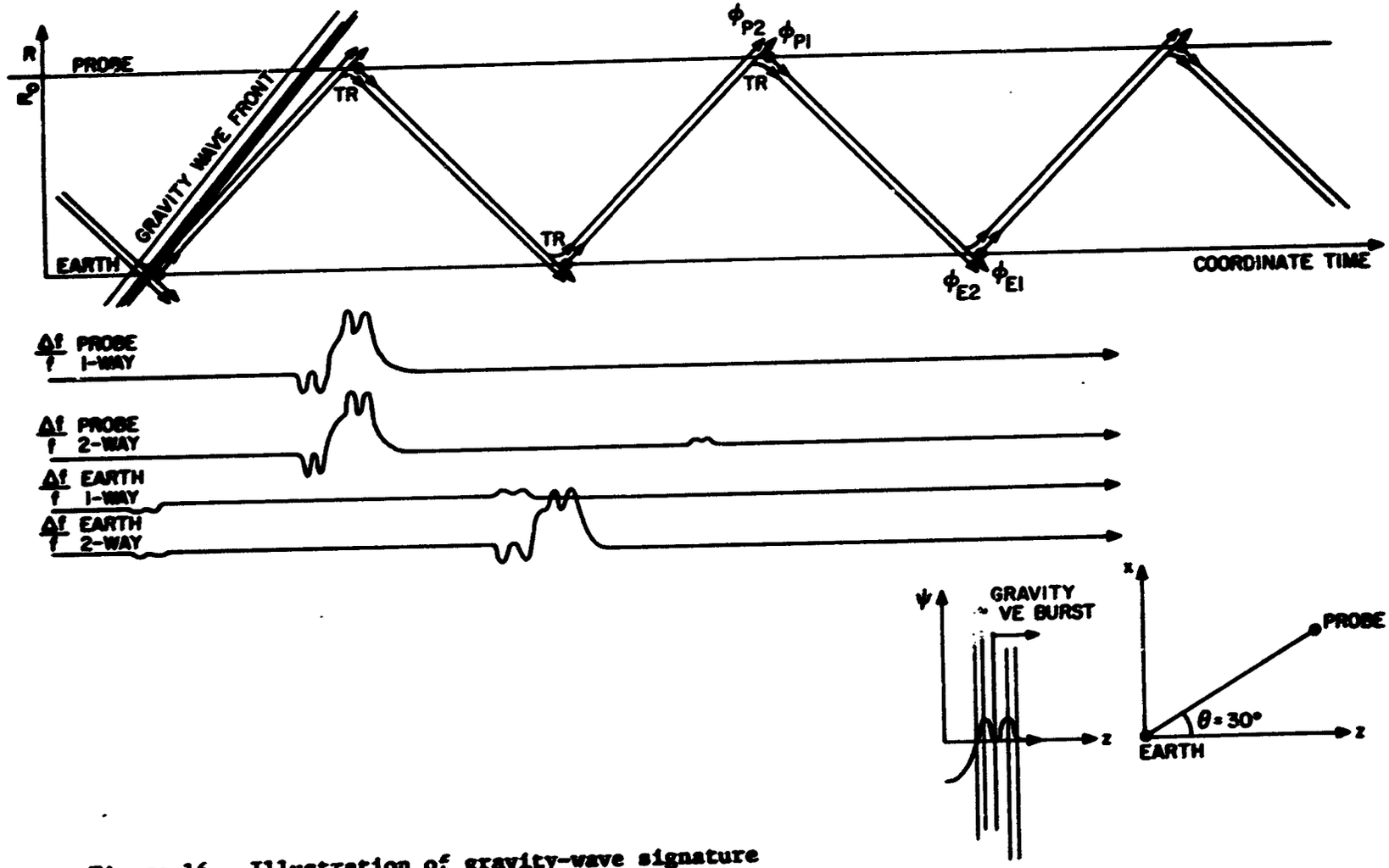


Figure 16. Illustration of gravity-wave signature for a 4-link correlated system $\theta = 30°$.

In addition, the time-correlated tracking experiment requires:

(d) Precision X-band two-way tracking,
(e) Measurement of columnar electron density.

For the sake of clarity, and because the spacecraft systems have not yet been defined, we will consider that functions a, b, and e are performed by S- and X-band dual wavelength communications systems similar to that used on the Viking missions. This type of system as shown in a simplified block diagram in Figure 17 is not considered to be a part of the present experimental payload.

Function d, together with the on board maser clock, is the heart of the gravity experiment. The two-way tracking concept shown in simplified form in Figure 18 consists of two two-way Doppler tracking loops, one originating and terminating at the ground station and the second originating and terminating at the spacecraft. Inherent in this configuration is the ability to make comparisons between the spacecraft maser clock and the ground masers. In one case the maser frequency is transmitted from the ground to the spacecraft, and in the other case from the spacecraft to the ground terminal.

It should be noted that it is possible to merge the spacecraft and the experiment tracking systems to provide substantial weight and power savings. The configuration described here is a worst-case situation employing separate and distinct hardware to perform each task.

Overall System Design

From Figure 18 it can be seen that the four-link Doppler tracking system is implemented by two Doppler tracking loops, the uploop and the downloop. Each loop can be considered a conventional, two-way Doppler tracking system in which the apparent range rate of the spacecraft is determined by comparing the transmitted frequency with the transponded received frequency returned to the same terminal. The Doppler extractor performs the range-rate computation, correcting for the frequency offset of the repeater.

Each of the tracking loops serves a second purpose in addition to providing range-rate information. One leg of each loop is a clock link, providing phase comparison between the spacecraft maser and the ground maser. The comparisons are made directly to provide the one-way Doppler data and as well as in a Doppler-cancelled manner to permit phase monitoring at the much lower Doppler rate associated with the redshift and second-order Doppler.

The uploop originates at the ground station clock uplink transmitter, the output frequency of which is controlled by the ground maser, and continues to the receiver of the spacecraft repeater. The local oscillator of this receiver is phase-locked to the received carrier frequency. The local oscillator signal, in turn, is used for a clock comparison and also controls the output frequency of the spacecraft repeater transmitter. The repeater transmitter signal is received by the ground station uploop receiver and the local oscillator of this receiver is phase-locked to its received carrier frequency. The local oscillator frequency is compared with the output frequency of clock uplink transmitter to determine the apparent range rate of the spacecraft with respect to the ground station.

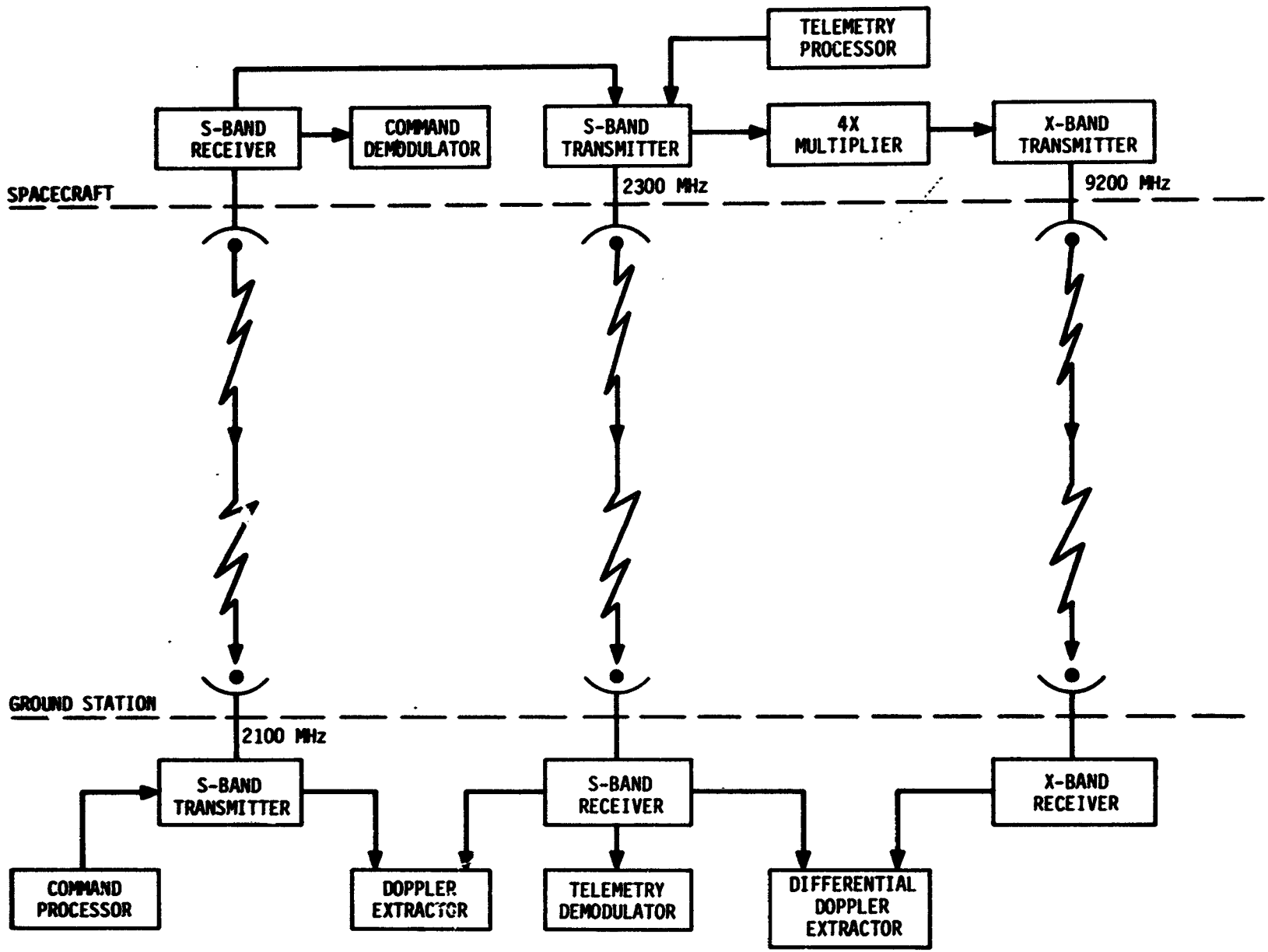


Figure 17. Unified S- and X-band communications system.

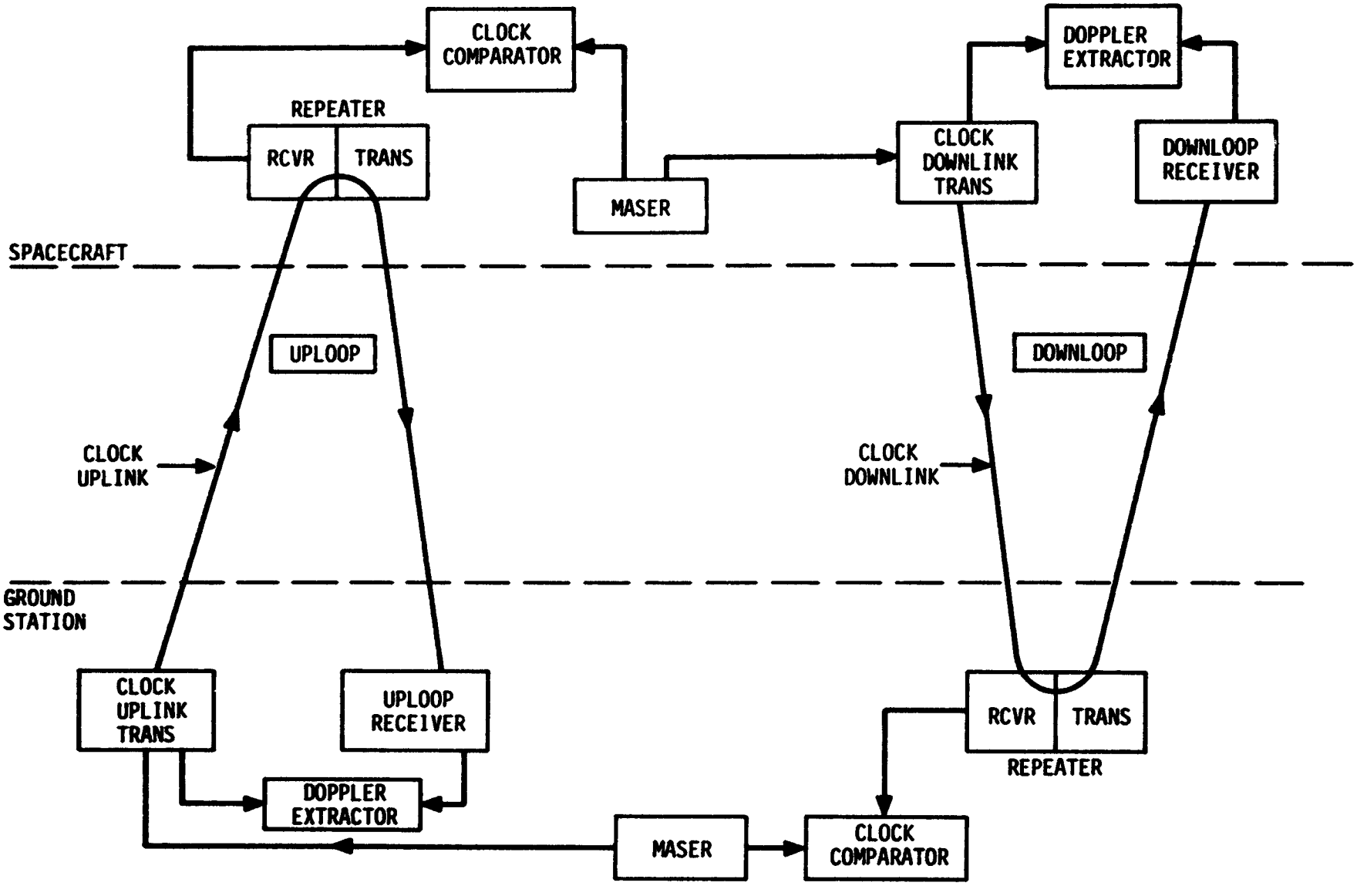


Figure 18. Four-link tracking system.

The downloop originates at the maser-controlled clock downlink transmitter in the spacecraft and continues to the ground station repeater receiver in the ground station. The local oscillator of the repeater receiver phase-locks to the received frequency and this frequency is compared with the ground station maser frequency to make a second clock comparison. The repeater receiver local oscillator also controls the output frequency of the ground station repeater transmitter. The output of the repeater transmitter is received back on the spacecraft and the local oscillator of the downloop receiver locks to the received carrier frequency. A second range-rate measurement is made by comparing the frequency with the spacecraft clock downlink frequency.

For short-range applications the two clock comparisons and the two range-rate measurements shown in Figure 18 are clearly redundant. However, as shown earlier for deep-space probes, the two set of measurements provide data that are not redundant and that contain information about the clocks, the electronics, and the transmission medium, which cannot be extracted in any other way.

The transmitter-receiver pairs in the spacecraft and the ground stations have deliberately not been labeled as "transponders." As will be shown in the next section, the proposed devices are much simpler, and inherently more phase stable, than the conventional multiple-use transponder used in the U.S.B. system.

Spacecraft System

Figure 19 is a block diagram of the spacecraft terminal for the proposed gravity experiment. The maser oscillator generates a low-level (-100 dBm) signal at 1420.405 MHz; the maser receiver provides standard signals at usable levels by phase-locking a 100-MHz crystal oscillator to the maser output. The first local oscillator frequency in the maser receiver is 1400 MHz. And multiplying this frequency by 6 in the frequency multiplier provides an extremely stable driving signal for the clock downlink transmitter. The output of the transmitter, 2 W at 8400 MHz, is coupled to the spacecraft antenna through the multiplexer.

The 8400-MHz signal from the maser receiver is also fed to the spacecraft clock comparator. A sample of the 8500-MHz local oscillator from the clock uplink receiver is the other input to the comparator. The output of the clock comparator is a measure of the apparent one-way range rate, including first- and second-order relativistic effects. These data are digitized and transferred to the spacecraft telemetry system.

The downloop repeater receiver completes the downloop; Figure 20 is a simplified block diagram of this receiver. The nominal 100-MHz VCXO is multiplied by 89 to 8900 MHz in the frequency multiplier chain and then mixed with the incoming 8850-MHz signal to generate a 50 MHz IF. The VCXO frequency is divided by 2 to produce a 50-MHz phase detector reference; the output of the phase detector, in turn, locks the VCXO to the incoming carrier.

A sample of the 8900-MHz first LO frequency from the downloop receiver is fed to the Doppler extractor to be compared with the clock downlink frequency. The Doppler extractor computes the apparent range rate of the vehicle from the two-way Doppler shift of the received carrier, digitizes the data, and transfers the data to the spacecraft telemetry system.

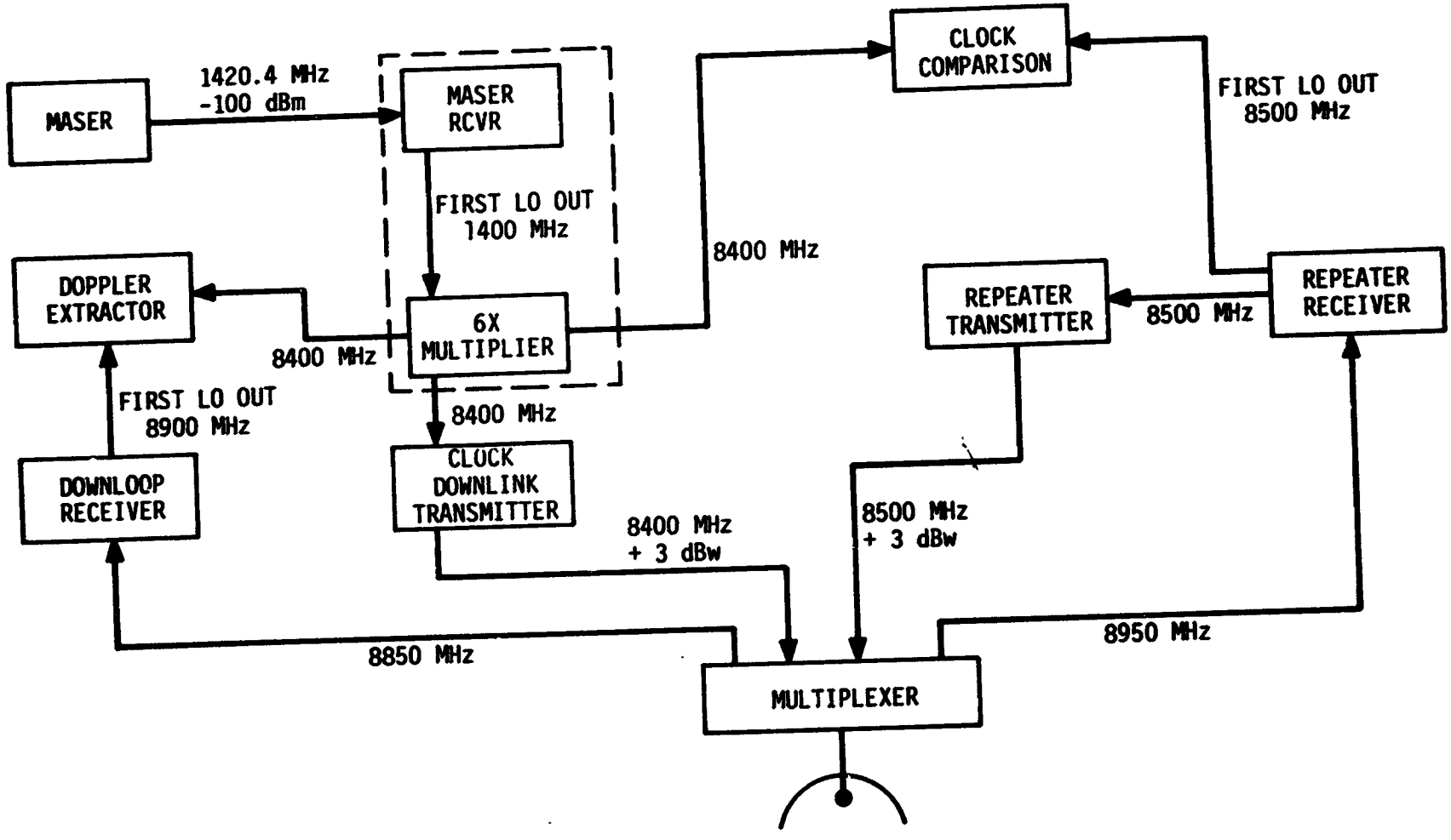


Figure 19. Spacecraft X-band terminal.

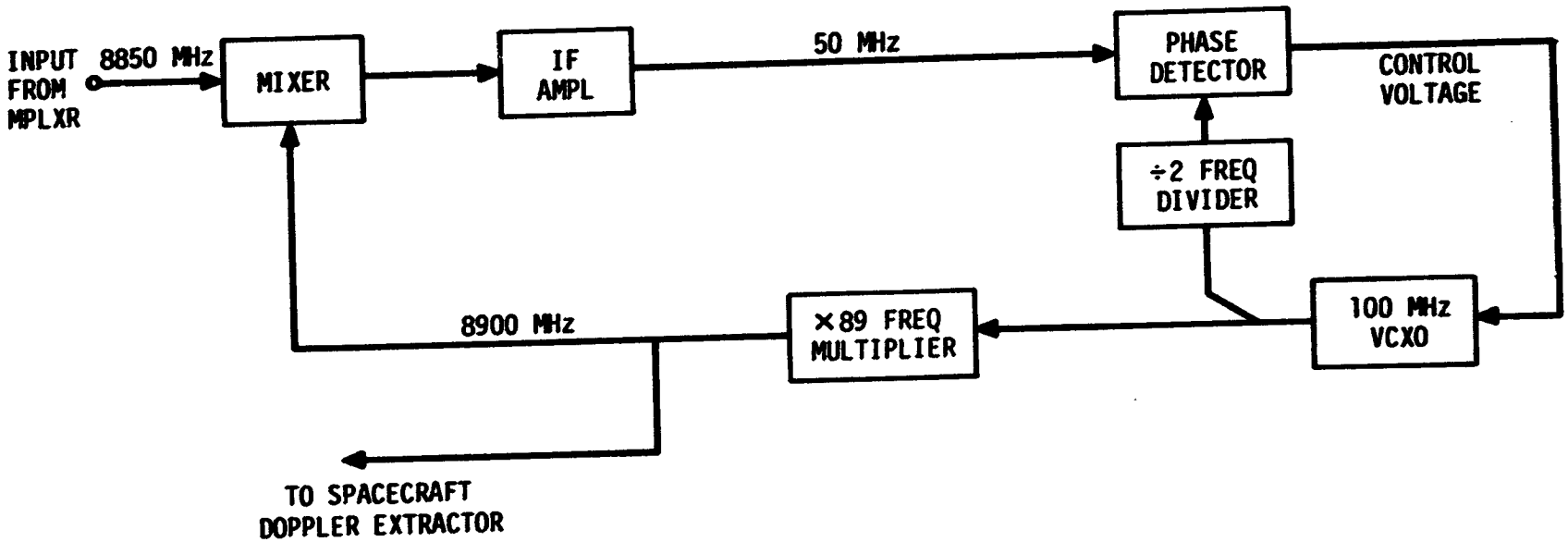


Figure 20. Block diagram of the spacecraft downloop receiver.

The repeater receiver, shown in block diagram form in Figure 21, is also coupled to the antenna through the multiplexer. The receiver local oscillator is phase-locked to the incoming uplink carrier, which, in turn, is referenced to the ground maser. The first local oscillator frequency is 8500 MHz for a nominal uplink frequency of 8950 MHz. The 8500-MHz signal is generated by multiplying the 100-MHz output of the voltage-controlled crystal oscillator (VCXO) by 85 in the frequency multiplier chain. The 8500-MHz output of the frequency multiplier chain is split, isolated, and amplified to feed both the receiver first mixer and the input to the repeater transmitter.

The first IF frequency is 450 MHz and mixing with the second LO frequency of 500 MHz (also generated by the multiplier chain) produces a 50 MHz second IF. The 50-MHz phase reference for the phase detector is generated by dividing the VCXO frequency by 2 in the frequency divider. The DC output of the phase detector is fed back to the electronic tuning port of the VCXO to steer the oscillator into phase coherence with the received signal.

The repeater transmitter amplifies the 8500-MHz signal from the clock uplink receiver to a 2-W level; this signal is then coupled to the common antenna through the spacecraft multiplexer.

Note that this proposed design for the uploop repeater minimizes the phase noise and phase drift contributions from the clock uplink receiver and the repeater downlink transmitter. Virtually all components of the loop are within the phase-lock loop; phase errors in the IF amplifiers, frequency multipliers, and phase detectors are corrected by the gain of the loop. No special environmental conditioning for the repeater loop components should be necessary in order to satisfy the phase stability requirements of the experiment.

Ground System

Figure 22 is a block diagram of the ground terminal for the proposed relativity experiment. The ground station is functionally identical to the spacecraft terminal described in the previous section, except for the transmitter and receiver frequencies. Note that the clock uplink transmitter output frequency is determined by a programmable digital synthesizer, referenced to the ground maser, to permit frequency sweeping of the transmitter during the acquisition phase.

From all appearances the combination of probe dynamics and possible solar eruptions can bring about very high rates of change of Doppler frequency. Therefore the phase-lock loop bandwidths in the spacecraft and ground station terminals must be made wide enough to permit accurate tracking and avoid loss of lock. Wideband loops are incompatible with high signal-to-noise ratios in the loop bandwidth. One approach to resolving the two mutually contradictory requirements of wide bandwidth for accurate tracking of rapidly accelerating spacecraft and narrow bandwidth for high signal-to-noise in the loop is illustrated in Figure 22. The clock uplink transmitter frequency and the repeater receiver frequency are determined by frequency synthesizers, which in turn are computer-controlled. The station computer is programmed with the predicted vehicle trajectory as a function of real time (epoch) so that the transmitter and receiver frequencies are stepped to minimize the loop stress in the spacecraft repeater and the ground station repeater receivers. The concept can easily be extended to reduce the loop stress in all four system receivers.

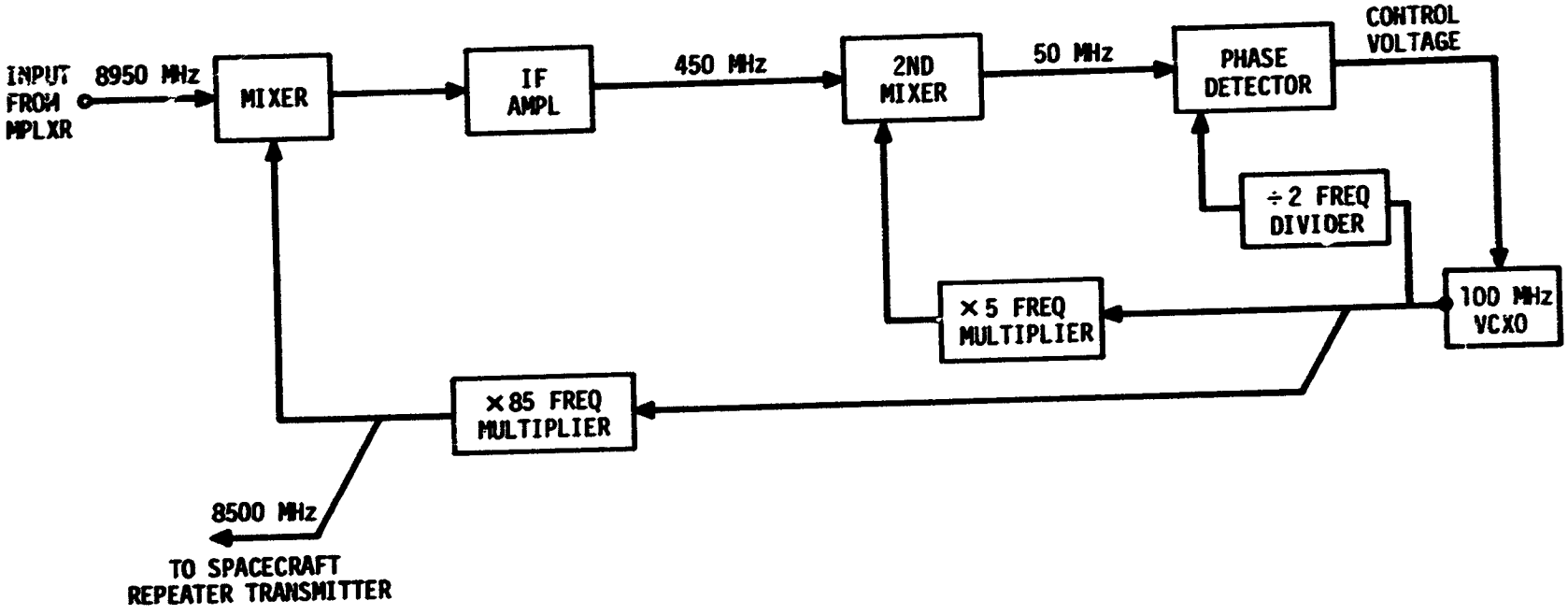


Figure 21. Block diagram of spacecraft repeater transmitter.

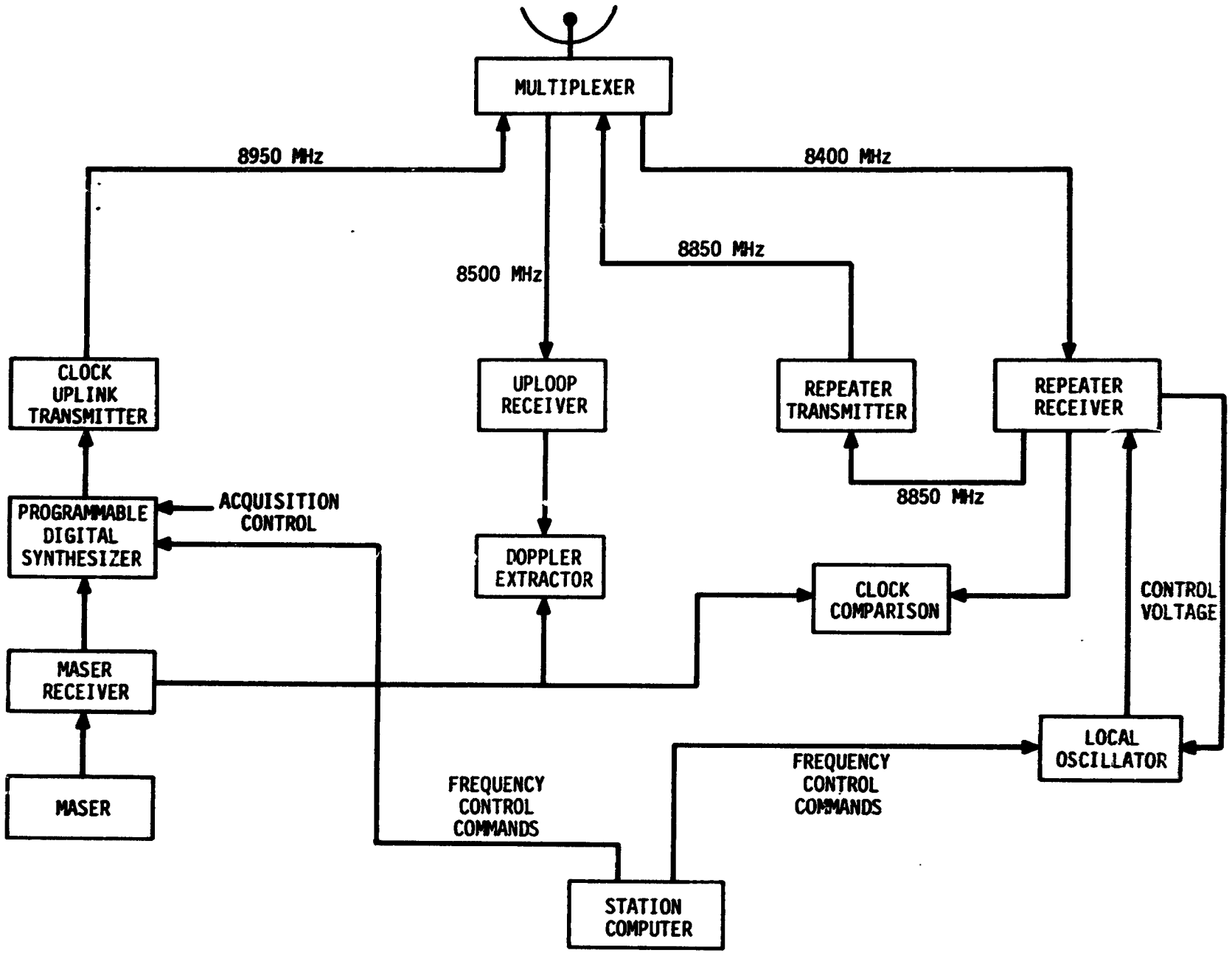


Figure 22. Block diagram of ground station X-band terminal.

This combination of open-loop and closed-loop phase tracking can yield almost optimum performance, but the synthesizers must be designed to maintain phase coherence and phase continuity throughout the frequency switching cycle. In addition, the computer must keep track of each cycle of phase throughout the critical phases of the mission, factoring in the variations in the rate-of-change of phase caused by the open-loop tracking system.

Signal Margins

It is possible at this time for us to provide only crude estimates of signal-to-noise ratios, and therefore of received carrier fractional frequency stability. In the following tables we have assumed the use of the 26-m DSN antennas with very-low-noise (25 K) receivers at X-band (8800 MHz). We have further assumed that relatively narrow loop band widths are adequate, either because of the geometry of the trajectory at solar encounter, or because open-loop tracking has reduced the loop stress. To estimate the contribution of solar noise at encounter, we have assumed a geometry that places the spacecraft at an apparent angular separation of 4 solar radii from the center of the solar disk (2.0°). A solar noise model that assumes a noise temperature of 10^4 K within the solar disk, which drops abruptly to cosmic background levels at the limb, has been used to estimate the effective noise temperature of the 26-m antennas at solar encounter. The computed value is shown in Table 1, which lists the components of the signal margin estimate at a range of 1 AU for the clock downlink. No allowance has been made for solar atmosphere scattering and polarization losses.

Table 1 also shows a similar calculation for the range at the superior conjunction of Jupiter (range is 9.3×10^8 km). The system parameters are similar to the first case, except for a lower effective receiver temperature and a narrower loop bandwidth.

Tracking Operations

Tracking operations for the solar probe can utilize existing DSN practice and technique. Since the critical phase of the solar encounter is approximately 19 hours in duration, it will be necessary to hand off control between DSN tracking stations. In order to accomplish the transition without loss of phase information, it is necessary to provide simultaneous tracking of the spacecraft by both ground stations before hand off. One station will be tracking in a two-way mode, the second station in the receive-only three-way mode. At this point, the situation is very similar to a VLBI experiment; the frequency (phase) of the onboard maser is compared to each of the ground masers and the two ground masers are intercompared through the clock links to the spacecraft. At the appropriate time, the second station will assume uplink control, with phase continuity in the transition ensured by the continuous, unbroken, phase comparisons of the three masers.

CONCLUSIONS

The concept of simultaneous one-way and two-way tracking by using similar terminals at both ends of the interval of space time lying between a pair of spacecraft or between Earth and a spacecraft provides a method to extract a very large amount of information about the structure of space time and the propagation medium. Such information, when obtained from a deep-space mission such as a Solar Probe, will be

TABLE 1. Signal margins

Parameter	Solar encounter		Superior conjunction	
	Level	Notes	Level	Notes
Frequency		8800 MHz		8800 MHz
Range		1.5×10^8 km		9.3×10^8 km
Antenna Diameter				
Transmitter		1.0 m		1.0 m
Receiver		26 m		26 m
Transmitter Power	+33 dBm	2.0 m	+33 dBm	2.00 w
Receiver Effective Temperature		50 K		30 K
Space Loss	247.9 dB		290.0 dB	
Antenna Gain				
Transmitter	40.1 dB		40.1 dB	
Receiver	64.9 dB		64.9 dB	
Loop Bandwidth		50 Hz		10 Hz
Receiver Input Level	-136.9 dBm		-152.7	
Carrier/Noise Ratio	27.0 dB		21.2 dB	
Frequency Stability Referenced to Carrier Frequency	8.0×10^{-15}	100 sec avg.	1.6×10^{-14}	100 sec avg.

of great interest in the search for gravity waves, particularly during the very long drift phase before solar encounter. During the solar encounter and while the drag-free system is operating, we would use the system to "sleuth" out and determine the origin of the many systematic variations expected in the Doppler phase and reduce substantially our dependence on *à priori* models and data averaging and filtering processes. This will permit us to measure Doppler shifts that are of gravitational origin, such as the solar quadrupole moment, solar oscillations and relativistic gravitational effects.

We believe that the system we have described is completely feasible and does not significantly stretch the state of the art of clocks and microwave techniques.

REFERENCES

[1] R. F. C. Vessot, and M. W. Levine, "A Preliminary Report on the Gravitational Red-Shift Rocket Probe Experiment," in Experimental Gravitation, Accademia Nazionale dei Lincei, Rome, 1977, pp. 372-391.

[2] David W. Allan, "Statistics of Atomic Frequency Standards," Proc. IEEE, vol. 54, #2, pp. 221-230, Feb. 1966.

[3] J. A. Barnes, A. R. Chi, L. S. Cutler, D. J. Healey, D. B. Leeson, T. E. McGunigal, J. A. Mullen, W. L. Smith, R. L. Sydnor, R. F. Vessot and G. M. R. Winkler, "Characterization of Frequency Stability," IEEE Trans. on Instrumentation and Measurement, vol. IM-20, #2, pp. 105-120, May 1971.

[4] R. F. C. Vessot, L. F. Mueller and J. Vanier, "The Specification of Oscillator Characteristics from Measurement Made in the Frequency Domain," Proc. IEEE, vol. 54, #2, pp. 199-207, Feb. 1966.

[5] R. F. C. Vessot, M. W. Levine and E. M. Mattison, "Comparison of the Theoretical and Observed Hydrogen Maser Stability Limitation due to Thermal Noise and the Prospect for Improvement by Low-Temperature Operation," Proceedings of the Ninth Annual Precise Time and Time Interval Applications and Planning Meeting, NASA-Tech, Memo 78104, pp. 549-569, March 1978.

[6] R. F. C. Vessot and M. W. Levine, "Performance Data of Space and Ground Hydrogen Masers and Ionospheric Studies for High-Accuracy Comparisons between Space and Ground Clocks," Proceedings of the 27th Annual Frequency Control Symposium, 1973.

[7] A. J. Tucker and B. M. Fannin, "Analysis of Ionospheric Contributions to the Doppler Shift of C. W. Signals from Artificial Earth Satellites," Journal of Geophysical Research, Space Physics, vol. 78, No. 13, pp. 4325-4334, July 1968.

[8] J. Saastamoinen, "Contributions to the Theory of Atmospheric Refraction," Bulletin Geodesique, No. 105, p. 9279, Sept. 1972, No. 106, p. 9383, Dec. 1972.

[9] A. B. Severny, V. A. Kotov and T. T. Tsap, "Observations of Solar Pulsations," Nature, vol. 259, pp. 87-89, Jan. 1976.

J. R. Brookes, G. R. Isaak and H. B. VanDer Raay, "Observations of Free Oscillations of the Sun," Nature, vol. 259, pp. 92-95, Jan. 1976.

[10] K. Nordtvedt, Jr., "A Study of One and Two-Way Doppler Tracking of a Clock on an Arrow Towards the Sun," in Experimental Gravitation, Accademia Nazionale dei Lincei, Roma, pp. 247-256, 1977.

[11]A. J. Anderson, "Detection of Gravitational Waves by Spacecraft Doppler Data," in Experimental Gravitation, ibid, pp. 235-246.

[12]H. D. Wahlquist, J. D. Anderson, F. B. Estabrook and K. S. Thorne, "Recent J. P. L. Work on Gravity Wave Detection and Solar System Relativity Experiments," in Experimental Gravitation, ibid., pp. 335-350.

[13]S. Thorne and V. B. Braginsky, "Gravitational Wave Bursts from the Nuclei of Distant Galaxies and Quasars: Proposal for Detection Using Doppler Tracking of Interplanetary Spacecraft," Astrophysical Journal, vol. 204, pp. L1-L6, Feb. 1976.

D39

DETECTION OF GRAVITATIONAL RADIATION AND OSCILLATIONS OF THE SUN VIA DOPPLER TRACKING OF SPACECRAFT

D. H. Douglass
Department of Physics and Astronomy
University of Rochester
Rochester, New York 14627

ABSTRACT

The magnitude of Doppler signals produced by gravitational wave bursts, continuous gravitational waves, and oscillations of the Sun interacting with a spacecraft are considered. Expressions are worked out for the appropriate noise entering each measurement. The noise sources considered are the Doppler extractor, fluctuations in the solar wind and the troposphere, and fluctuations in the reference oscillator. With a single (X-band) frequency tracking system it is suggested that: gravitational bursts with $h \gtrsim 10^{-15}$ and duration $t \gtrsim 10^{2}$ sec can be detected; continuous gravitational waves with $h \gtrsim 10^{-17}$ can also be measured; and the oscillating solar quadrupole moment (dimensionless) can be determined if its value is of the order of 10^{-7}. It is also suggested that the tropospheric fluctuations can be measured in an auxiliary experiment and subtracted from the Doppler signal.

Introduction

My original purpose in presenting this paper at the Solar Probe Workshop was to call attention to the possibility of detecting continuous gravitational radiation by Doppler tracking of spacecraft. The possibility of detecting gravitational radiation by Doppler tracking has, of course, been considered previously in a number of papers (Estabrook and Wahlquist, 1975; J. D. Anderson, 1974; Wahlquist et al., 1976; A. J. Anderson, 1971 and 1977; Thorne and Braginsky, 1976) and by Thorne and Estabrook at this meeting. With the exception of A. J. Anderson (1977), the general thrust of these studies was directed toward the possibility of detecting bursts or isotropic incoherent waves. Thus my remarks are intended to complement those of Thorne and Estabrook. The apparent lack of enthusiasm for continuous waves is probably based on the fact that the sensitivity level that can be easily achieved is not good enough to detect any presently known source. However, the candidate pool for potential sources (binary, rotating, and vibrating stars) is very large. This sensitivity, which is "not good enough" can, in fact, be at the 10^{-17} level of dimensionless amplitude h. To my knowledge, there has never been a systematic search for continuous gravitational radiation in the frequency band available to the Doppler tracking system (10^{-4} to 10^{-1} Hz). I would guess, though, that if the amplitude were as large as 10^{-8}, that it could not have escaped detection by influencing some other experiment. (Again, I exclude A. J. Anderson these last remarks. I shall discuss his work later in the paper.) Astronomy and astrophysics are constantly full of surprises - estimates of the magnitudes of quantities are very often in error by many orders of magnitude. Whenever a new apparatus is developed which is many orders of magnitude better than its predecessor, one should turn it on and find out what signals it sees. The Doppler tracking systems proposed for future space flights could be such instruments.

In the processes of calculating the magnitude of the Doppler signal produced by continuous gravitational waves and estimating the associated noise, I was also forced to consider the problem of the detection of gravitational burst signals. I show below that the "Allan variance", which is commonly used to estimate the noise amplitude that competes

with the gravitational burst amplitude, overestimates the noise amplitude by a factor which can be as large as 7.5. This, I believe, is good news since the estimates of the strength of the burst sources are only about an order of magnitude below what was thought to be the noise in the Doppler tracking system.

I make the pedagogical point in this paper that since the Doppler tracking system measures phase, all measurable quantities should also be expressed in terms of the phase, and I have done so. It follows that all sources of noise should also be described in terms of phase; this includes the reference oscillator. This means that the (now discredited) "Allan variance" definition of burst noise is not needed.

In this paper, I explicitly define the phase noise associated with the detection of burst and continuous wave signals. Quantitative expressions for these definitions are worked out for the noise from the Doppler extractor, the solar wind, the troposphere, and the reference oscillator. At this Workshop and elsewhere, Vessot has suggested that the fluctuations in the solar wind can be measured and subtracted from the Doppler signal. In a similar way, I suggest that the tropospheric fluctuations can be measured in an auxiliary experiment and also subtracted away. This is discussed below.

Lastly, after my talk at this Workshop, Ian Roxburgh, Kenneth Nordtvedt and I discussed the possibility of detecting the gravitational radiation from the Sun if it is vibrating with an amplitude as large as is implied by the measurements of Hill and others. We decided that this would not be feasible because the spacecraft would not be in the wave zone of the radiation. At this point in the conversation, it became obvious to us that a vibrating Sun would produce "ordinary" time varying Newtonian fields which would produce forces on the spacecraft which might be large enough to be detected. A colleague of mine, Warren Johnson, and I proceeded to work out the order of magnitude of the effect. Our rough calculations indicate that if the dimensionless oscillating quadrupole moment J_2' of the Sun is as large as 10^{-7} that it can be detected. This very exciting prospect is discussed below.

I. Signals from the Doppler Tracking System

A. The Doppler Tracking System

In the simplest form of a Doppler tracking system, an electromagnetic signal of frequency ν_o is transmitted from the Earth to a spacecraft which transponds it back to Earth. The received signal $\nu(t)$ will differ from ν_o due to the motion of the spacecraft and will consequently vary with time. The difference frequency $f(t)$ is defined as

$$f(t) = \nu(t) - \nu_o \tag{1.1}$$

The signal $\nu(t)$ is usually detected in a circuit whose output yields the phase $\phi(t)$

$$\phi(t) = 2\pi \int f(t)dt \tag{1.2}$$

The relative velocity $v(t)$ of the spacecraft is related to $f(t)$ via the Doppler effect

$$\frac{v(t)}{c} = \frac{f(t)}{\nu_o} \tag{1.3}$$

Thus the phase can be expressed as

$$\phi(t) = \frac{2\pi\nu_o}{c} \int v(t)dt \tag{1.4}$$

Also, the range $r(t)$ is given to within an additive constant (which can be known) by

$$r(t) = \int v(t)dt \tag{1.5}$$

Thus one can express the phase directly in terms of the range

$$\phi(t) = \frac{2\pi\nu_o}{c} r(t) \tag{1.6}$$

B. Gravitational Wave Signals

Estabrook and Wahlquist (1975), among others, have calculated the influence of a gravitational wave of dimensionless amplitude $h(t)$ on the Doppler tracking signal. They find that, for optimum polarization and orientation,

$$\frac{f(t)}{\nu_o} \approx h(t) \tag{1.7}$$

[There is a factor $\sin(\frac{\pi r}{\lambda_g})$ multiplying the right hand side of (1.7) which I have set equal to 1; r is distance to the spacecraft and λ_g is the wavelength of the gravitational wave.] Thus the influence of the gravitational wave on the phase is given by

$$\phi(t) = 2\pi\nu_o \int h(t)dt \tag{1.8}$$

There are two kinds of gravitational waves - bursts and continuous waves.

1. Gravitational Wave Bursts

These are the signals from gravitationally collapsing objects that Thorne and Estabrook discussed at this workshop (see Thorne and Braginsky, 1976). These signals are characterized by a change in amplitude h_b which occurs during a characteristic time τ_b. Douglass and Braginsky (1978) give expressions which yield the order of magnitude of τ_b in terms of the mass M

$$\tau_b \approx \frac{4GM}{c^3} = (2 \times 10^{-5} \text{ sec})(\frac{M}{M_\odot}) \tag{1.9}$$

and the amplitude in terms of τ_b

$$h_b \approx \frac{c\tau_b}{R} E^{\frac{1}{2}} = 1.0 \times 10^{-15} \left(\frac{\tau_b}{10^3 \text{sec}}\right)\left(\frac{10^9 \text{pc}}{R}\right)\left(\frac{E}{0.01}\right)^{\frac{1}{2}} \tag{1.10}$$

where R is the distance to the object and E is the efficiency for converting mass into gravitational radiation. These simple expressions yield estimates that are comparable to those of Thorne. (For example, if $M \approx 10^7 M_\odot$, then $\tau_b \approx 2 \times 10^2$ sec. This object, placed at $R \approx 3 \times 10^9$ pc, produces an $h_b \approx 7 \times 10^{-17}$ for an efficiency factor $E = 0.01$.) The relation between the burst amplitude h_b and the resulting change in phase ϕ_b is obtained from (1.8)

$$h_b \sim \frac{\phi_b}{2\pi\nu_o\tau_b} \tag{1.11}$$

The smallest h_b that can be measured is determined by the phase noise entering into the measurement of ϕ_b. This will be considered below.

2. Continuous Gravitational Waves

Continuous gravitational radiation can be produced by vibrating and rotating stars and by binary stars. Since 50% of the stars in our galaxy are thought to be binary, there is potentially a very large number of sources. Approximately 3000 eclipsing binaries are known, several hundred of which are known well enough to have had their orbital elements determined. Both neutron stars and white dwarf stars can be excited into modes of oscillation which would be expected to relax by emitting gravitational radiation. A continuous wave source would be characterized by an amplitude h_c and a frequency f_c

$$h(t) = h_c e^{i2\pi f_c t} \tag{1.12}$$

where f_c is expected to be in the very low frequency band (10^{-4} to 10^{-1} Hz), which is the band in which the Doppler tracking system is most sensitive. The largest h_c among known sources is $\sim 10^{-20}$. (See Douglass and Braginsky (1978) for a comprehensive review.) One can only speculate about unknown sources. However, as mentioned above, 50% of the stars in our galaxy are binary, but only ~3000 have been well studied; and, of course, these were not selected on the basis of their gravitational properties. The continuous waves will produce a phase signal

$$\phi(t) = \phi_c e^{i2\pi f_c t} \tag{1.13}$$

The relation between ϕ_c and h_c is obtained from (1.08)

$$h_c = \frac{f_c}{\nu_o} \phi_c \tag{1.14}$$

The smallest h_c that one can detect is again determined by the phase noise entering the measurement of ϕ_c. This will obviously be different from the estimate of phase noise for bursts. This calculation will be done below.

C. Signals Induced by Oscillations of the Sun

One of the important topics of this Workshop is the possibility of measuring the static quadrupole moment of the Sun. The discussions suggest that, with a probe that passes within 4 $R_\odot$ of the Sun, values of the dimensionless quadrupole parameter J_2 of the order of 10^{-7} could be measured.

For a probe that comes this close to the Sun, one might expect that if the Sun is vibrating in modes with fractional amplitudes of 10^{-5} as the work of Hill et al. (1976) and Severny et al. (1976) suggests, then an oscillation in the position of the spacecraft at the vibrational frequency of the Sun might be detected. The lowest modes which would couple to a spacecraft are the quadrupole modes. Following a calculation of Johnson and Douglass (1978), the potential for the Sun for this case can be expressed as

$$V(r_s,\theta,\phi,t) \approx \frac{GM_\odot}{r_s}\left[1 + J_2' e^{i\omega_2 t}\left(\frac{R_\odot}{r_s}\right)^2\left(\frac{3\cos^2\theta-1}{2}\right)\right] \tag{1.15}$$

where r_s is the distance from the Sun to the spacecraft, $M_\odot$ is the mass of the Sun, $R_\odot$ is the radius of the Sun, ω_2 is the vibrational frequency of the Sun, and J_2' is the dimensionless amplitude of the oscillating quadrupole moment. The oscillating potential will modulate the position of the spacecraft as

$$r(t) = r_2 e^{i\omega_2 t} \tag{1.16}$$

which in turn produces an oscillation in the Doppler signal of

$$\phi(t) = \phi_2 e^{i\omega_2 t} \tag{1.17}$$

The amplitude of oscillation at the distance r_s is easily related to J_2'

$$r_2 = \frac{3GM_\odot}{R_\odot^2}\,\frac{1}{\omega_2^2}\left(\frac{R_\odot}{r_s}\right)^4 J_2' \tag{1.18}$$

where $\theta = 0$ has been chosen as a reference direction. The order of magnitude of some of the reported vibrational periods that have been reported (Severny et al., 1976) are $\sim$2 hrs. So for $r_s = 4\ R_\odot$

$$r_2 = (40\text{ cm})\left(\frac{J_2'}{10^{-7}}\right)\left(\frac{\tau_2}{2\text{ hrs}}\right)^2\left(\frac{4R_\odot}{r_s}\right)^4 \tag{1.19}$$

where $\tau_2 = (2\pi)/\omega_2$.

Johnson and Douglass (1978) have estimated the order of magnitude of J_2'

$$J_2' \sim 10^{-2} \left(\frac{\delta R_\odot}{R_\odot}\right) \tag{1.20}$$

where $\delta R_\odot$ is the magnitude of the displacement of the vibrating Sun at its surface. This calculation was for a polytrope (n = 3) model of the Sun and follows the method of Cowling (1941). If the observations of Hill et al. (1976), Severny et al. (1976), and others are interpreted as implying $R_\odot/R_\odot \sim 10^{-5}$, then $J_2' \sim 10^{-7}$, which is comparable in magnitude to values expected for the static Sun. The phase amplitude ϕ_2 and displacement amplitude r_2 are related according to Equation (1.6)

$$r_2 = \frac{c}{2\pi\nu_o}\phi_2 \tag{1.21}$$

The phase noise associated with the measurement ϕ_2 will be estimated below.

II. Noise in the Doppler Tracking System

In order to determine the detectability of the signals discussed above, one must estimate the phase noise associated with that measurement. One of the objectives of this paper is to show that the phase noise depends on what measurement is being made. The various sources of phase noise include: the Doppler extractor, fluctuations of the solar wind, fluctuations in the troposphere, noise in the transponder and receiver, and noise in the reference oscillator which is used in the comparison measurement of the Doppler signal. There are perhaps other sources that I have forgotten. I will not discuss all of the sources of noise. Instead I will work with those most commonly discussed so as to illustrate the orders of magnitude of the effects and to show the differences among the various types of measurements.

A. The Doppler Extractor

The phase noise due to the Doppler extractor ϕ_{nd} is frequently expressed in terms of a fraction of a cycle of phase

$$\phi_{nd} = \eta 2\pi \tag{2.1}$$

The fraction η for present Doppler systems is $\sim 10^{-2}$. There exists a system called Occultation Data Assembly (ODA) which should have $\eta \approx 10^{-4}$. There is no reason why values of $\eta \approx 10^{-6}$ could not be achieved if that is deemed necessary.

B. Phase Noise which is Characterized by a Spectral Density

1. The Spectral Density

With the exception of the Doppler extractor, almost all the rest of the phase noise can be characterized by a spectral density of phase fluctuations $S_\phi(f)$. (The Doppler extractor can be described by a spectral density if necessary. This is discussed below.) Furthermore, the spectral density frequently follows a power law in f.

$$S_\phi(f) = \frac{a}{f^\beta} \tag{2.2}$$

The index β is characteristic of the particular noise process. For example, both the solar-wind and tropospheric fluctuations are described quite well by an index close to $\beta = 2.6$.

2. Definition of Phase Noise

As mentioned above, the amount of phase noise depends on the measurement. I will now proceed to define and evaluate the phase noise associated with different measurements assuming the spectral density given by (2.2).

a. Bursts. If one is detecting a burst signal characterized by a change in phase which occurs over a time interval τ_b (for example, the gravitational wave burst signal given by (1.11)), then the appropriate phase noise ϕ_{nb} against which to compare the signal is given by

$$\phi_{nb}^2 = \int_{f_\ell}^{f_u} S_\phi(f)\,df \tag{2.3}$$

The lower frequency limit is given by

$$f_\ell \approx \tau_b^{-1}$$

and the upper frequency limit should be at least 10 times the lower frequency in order to resolve the burst. For the power law expression, Equation (2.3) becomes

$$\phi_{nb}^2 = \begin{cases} \frac{a}{1-\beta}\left(f_u^{1-\beta} - f_\ell^{1-\beta}\right) & , \beta < 1 \quad (2.4a) \\ a\,\ell n\left(\frac{f_u}{f_\ell}\right) & , \beta = 1 \quad (2.4b) \\ \frac{a}{\beta-1}\left(\frac{1}{f_\ell^{\beta-1}} - \frac{1}{f_u^{\beta-1}}\right) & , \beta > 1 \quad (2.4c) \end{cases}$$

For $\beta > 1$, the upper frequency can be chosen to be infinite. Doing this and using $f_\ell = \tau_b^{-1}$, one obtains

$$\phi_{nb}^2 = \frac{a\tau_b^{\beta-1}}{\beta-1}, \quad \beta > 1 \qquad (2.5)$$

b. Continuous Waves. For a continuous wave signal at the frequency f_c, the appropriate phase noise ϕ_{nc} is given by

$$\phi_{nc}^2 = \int_{f_c - \frac{\Delta f}{2}}^{f_c + \frac{\Delta f}{2}} S_\phi(f)df \qquad (2.6)$$

where Δf is the bandwidth of the measurement. Assuming $\Delta f < f_c$

$$\phi_{nc}^2 \approx S_\phi(f_c)\Delta f \qquad (2.7)$$

If one integrates for a time τ, then

$$\Delta f \approx \frac{1}{2\tau} \qquad (2.8)$$

and

$$\phi_{nc}^2 \approx \frac{S_\phi(f_c)}{2\tau} \qquad (2.9)$$

For the spectral density given by (2.2), this becomes

$$\phi_{nc}^2 = \frac{a}{2\tau f_c^{\beta}} \qquad (2.10)$$

c. The Allan Variance. The Allan variance $\sigma_a^2(\tau)$ is an expression used by the frequency and time community to compare clocks and

oscillators. Since Estabrook and others have used this expression to estimate the noise that competes with a gravitational burst signal, I am compelled to discuss it. I shall show below that this expression overestimates the amplitude of the noise for this measurement by a factor between 4 and 8. The exact factor depends on the source of the noise.

The Allan variance is the variance of a set of successive fractional frequency differences spaced in time by the interval τ made on an oscillator of frequency ν. It can be shown that in terms of $S_\phi(f)$ this measurement is

$$\sigma_a^2(\tau) = \frac{\phi_{na}^2}{(2\pi\nu_o\tau)^2} \tag{2.11}$$

where

$$\phi_{na}^2 = 8\int_o^\infty df S_\phi(f)\left[\sin^2(\pi f\tau) - \frac{\sin^2(2\pi f\tau)}{4}\right] \tag{2.12}$$

Using (2.2) one obtains for $\beta \geq 1$

$$\phi_{na}^2 = \frac{4\pi^2 a\, g_\beta}{\tau^{1-\beta}} \tag{2.13}$$

and

$$\sigma_a^2 = \frac{a\, g_\beta}{\nu_o^2 \tau^{3-\beta}} \tag{2.14}$$

The constant g_β is tabulated in Table 1 for integer values of β.

Table 1. g_β as a function of β.

β	g_β
1	$\dfrac{3[2 + \ell n(2\pi f_u\tau)] - \ell n2}{4\pi^2}$
2	1/2
3	2 ℓn2
4	$(4\pi^2)/6$

Values of g_β for noninteger β can be interpolated from the table.

C. Phase Noise Due to Solar Wind Fluctuations

Fluctuations in the solar wind produce fluctuations in the phase of electromagnetic signals which propagate through it. Woo et al. (1976) have reported measurements of the fluctuation spectrum at S-band ($\nu_o = 2.4 \times 10^9$ Hz) from the Mariner 10 spacecraft. These measurements were made at a range of 1.6 AU when the closest approach distance R_c of the line-of-sight to the Sun was 43 $R_\odot$(=0.23 AU) and the solar elongation angle α (the Sun-Earth-spacecraft angle) was about 12°. They found

$$S_\phi = \frac{a_{sw}}{f^{2.6}} \tag{2.15}$$

where

$$a_{sw} \approx 1.0 \times 10^{-3}$$

Woo (1975) has considered the general problem of phase fluctuations as a function of ν_o and R_c, and he indicates that

$$a_{sw} \propto R_c^{-3.2} \nu_o^{-2} \tag{2.16}$$

Assuming these dependences and normalizing to X-band (9×10^9 Hz) one obtains

$$a_{sw} = 4.2 \times 10^{-7} \left(\frac{1 \text{ AU}}{R_c}\right)^{3.2} \left(\frac{9 \times 10^9 \text{ Hz}}{\nu_o}\right)^2 \frac{w(\alpha)}{w(12°)} \tag{2.17}$$

where I have added a function $w(\alpha)$ to allow for variation with elongation angle α. In his talk, Frank Estabrook gave the happy news that Woo and Armstrong's (1978) Viking observations showed a dramatic decrease in spectral density at large angles. From the graph shown by Estabrook, I estimate $w(\alpha)/w(12°)$ to be about 10^{-3} for $\alpha = 180°$. The various phase noises are easily worked out for this case:

bursts $$\phi_{nb} = 0.78\, a_{sw}^{1/2}\, b^{0.8} \tag{2.18}$$

continuous waves $$\phi_{nc} = 0.5\, a_{sw}^{1/2}\, \tau^{-1/2} f_c^{-1.3} \tag{2.19}$$

Allan variance $$\phi_{na} = 6.0\, a_{sw}^{1/2}\, \tau^{0.8} \tag{2.20}$$

It should be noted that the Allan variance phase noise (2.20) is a factor of 7.5 larger than that for the bursts (2.18). This result is independent of a_{sw}.

It should be noted that this is the noise for a one-way Doppler ranging system. Vessot has reported at this Workshop and elsewhere on a scheme involving a clock on a spacecraft with two-way Doppler ranging and with multiple frequency capability. He makes the case that with this scheme much of the phase noise due to the solar wind can be removed.

D. Phase Noise Due to Fluctuations in the Troposphere

Phase fluctuations in the troposphere have been observed by Thompson et al. (1975) and others. They report measurements at 9.5×10^9 Hz and 22.2×10^9 Hz and show that the spectrum is the same at both frequencies. Averaging over the nine samples in their Table IV, I find

$$S_\phi = \frac{a_{tr}}{f^{2.6}} \tag{2.21}$$

$$a_{tr} = 3.4 \times 10^{-6}\left(\frac{\ell o}{65\ \text{km}}\right)^2 \tag{2.22}$$

These measurements were made over a distance of 65 km, so I used a scaling factor to estimate the coefficient for a different path length ℓo. The first point to note is that the exponent of f is the same as for the solar wind. This is probably no accident since in both cases the fluctuations are caused by turbulence in a fluid medium. (Kolmogorov theory predicts an exponent of 8/3.) The second point is that the relations for the phase noise in Section C above can be used for this case simply by substituting the value of a for the troposphere.

The observed phase fluctuations seem to be correlated with the presence of water vapor. Also the coefficient a was found to be highly variable from sample to sample. Of the nine values reported by Thompson et al. (1975), the smallest value of a was 7.3×10^{-8} and the largest was 8.0×10^{-6}. This suggests that the receiver should be placed in a dry location and experiments should be performed during seasons of low water vapor. A most important observation by this group was that the fluctuation-spectra at 9.5×10^9 Hz and 22.2×10^9 Hz showed nearly 100% coherence up to fluctuation frequencies of 10^{-1} Hz. Since the frequencies of interest

for the signals discussed in this paper are considerably below this value, then tropospheric fluctuations can in principle be measured with an auxiliary experiment and subtracted away. One possible scheme would be to measure the tropospheric fluctuations by means of a signal from an overhead satellite with an antenna located next to the Doppler tracking antenna.

E. Phase Noise Due to the Reference Oscillator

The reference oscillator (local oscillator) will contribute phase noise which is indistinguishable from that from other sources, so this noise must be considered along with all the rest. One fundamental source of noise is the kT "noise energy" associated with every (classical) oscillator. This noise leads to a spectral density of phase fluctuations of

$$S_\phi = \frac{a_{\ell o}}{f^2} \tag{2.23}$$

$$a_{\ell o} = \frac{\nu_o^2 kT}{2Q^2 P} \tag{2.24}$$

where T is temperature, Q is the quality factor of the oscillator, P is the power dissipated, and ν_o is the frequency at which the oscillator is used. To estimate the magnitude of $a_{\ell o}$, we assume an oscillator with $T = 300K$, $Q = 2 \times 10^9$, and $P = 10^{-12}$ watts; then $a_{\ell o} = 5 \times 10^{-28}\ \nu_o^2$. If this oscillator is used at X-band ($\nu_o = 9 \times 10^9$ Hz) (one may have to multiply or divide the basic oscillator frequency to achieve this), then $a_{\ell o} \approx 4.3 \times 10^{-8}$. These values are close to that of the hydrogen maser (see Vessot, these proceedings). To improve the performance of the hydrogen maser, one could cool it. Another possibility is the superconducting cavity stabilized oscillator, which may have higher Q and P and (certainly) lower T.

The various phase noises defined above can be worked out as follows:

bursts $$\phi_{nb} = (a_{\ell o}\tau_b)^{\frac{1}{2}} \tag{2.25}$$

continuous waves $$\phi_{nc} = \left(\frac{a_{\ell o}}{2\tau}\right)^{\frac{1}{2}} \frac{1}{f_c} \tag{2.26}$$

Allan variance $$\phi_{na} = \sqrt{2}\,\pi\,(a_{\ell o}\tau)^{\frac{1}{2}} \tag{2.27}$$

Note that the Allan variance phase noise is a factor $\sqrt{2}\,\pi$ (=4.4) larger than that for the burst definition. Again, this factor is independent of the value of a.

Oscillators can be no better in their noise performance than that given by (2.23) and (2.24). They are almost always worse because of noisy electronics, amplifiers, vibration, aging effects, etc. (These effects can sometimes be characterized by power law spectral densities.) To my knowledge, only Vessot has reached the thermal limit of an oscillator (the hydrogen maser) and in this case only over a restricted range of frequency. Clearly there are many practical difficult problems in building a stable oscillator. These problems, however, are not considered fundamental, so I have ignored them in this paper.

III. Detection of Signals in the Presence of Phase Noise

Having estimated the magnitude of the signals due to gravitational radiation and due to vibrations of the Sun and also having defined the noise associated with the measurement of those signals, we are now in a position to comment on the signal detectability.

A. Burst Signals

The burst gravitational wave signals of magnitude h_b were related to the phase by Equation (1.11). The condition for detectability is then

$$h_b > \frac{\phi_{nb}}{2\pi\nu_o\tau_b} \tag{3.1}$$

where ϕ_{nb} is the phase noise for this measurement. For the various sources enumerated above this is

$$\phi_{nb}^2 > 4\pi^2\eta^2 + a_{\ell o}\tau_b + 0.78\,(a_{sw} + a_{tr})\tau_b^{1.6} \qquad (3.2)$$

where the first term is from the Doppler extractor, the second from the reference oscillator, and the third from the solar wind and troposphere combined. It is explicitly noted that the burst definition of phase noise has been used in the second and third terms and not the Allan variance. (The differences are factors of 20 and 56, respectively.) Putting (3.2) into (3.1) one finds

$$h_b > \frac{1}{2\pi\nu_o}\left|\frac{4\pi^2\eta^2}{\tau_b^2} + \frac{a_{\ell o}}{\tau_b} + \frac{0.78(a_{sw} + a_{tr})}{\tau_b^{0.4}}\right|^{\frac{1}{2}} \qquad (3.3)$$

or, taking each term separately,

$$h_b > \begin{cases} 1.0 \times 10^{-15}\left(\frac{\eta}{10^{-3}}\right)\left(\frac{10^{10}\ \text{Hz}}{\nu_o}\right)\left(\frac{100\ \text{sec}}{\tau_b}\right) & (3.4a) \\ 1.6 \times 10^{-15}\left(\frac{a_{\ell o}}{10^{-6}}\right)^{\frac{1}{2}}\left(\frac{10^{10}\ \text{Hz}}{\nu_o}\right)\left(\frac{100\ \text{sec}}{\tau_b}\right)^{\frac{1}{2}} & (3.4b) \\ 5.6 \times 10^{-16}\left(\frac{a_{sw} + a_{tr}}{10^{-8}}\right)^{\frac{1}{2}}\left(\frac{10^{10}\ \text{Hz}}{\nu_o}\right)\left(\frac{100\ \text{sec}}{\tau_b}\right)^{0.2} & (3.4c) \end{cases}$$

where it has been assumed that a Doppler extractor with $\eta \sim 10^{-3}$ is feasible, that an X-band, thermally-limited oscillator is possible, and that the solar wind and tropospheric constants are of the of the order of 10^{-8} or below (this can be achieved for elongation angles of 180° and under favorable tropospheric conditions). These estimates are more optimistic by nearly a factor of 10 than those of Estabrook, primarily because of the use of the burst definition of noise. Also it should be pointed out again that these calculations are for a single frequency tracking system. Vessot's scheme with onboard clocks and multiple frequencies promises to take out much of the solar wind noise and I believe that the tropospheric noise can be handled in a similar way by measuring this noise with an auxiliary experiment. Thus gravitational wave burst amplitudes much smaller than those implied by (3.4) may be detectable.

Another point I wish to make is that it is incomplete to try to specify the h_b sensitivity without also specifying the associated τ_b. (The only exception to this statement is the case for which the dominant noise is $S_\phi \propto f^{-3}$, so that ϕ_{nb}/τ_b is independent of τ_b. Some sources appear to have a portion of their spectrum described by such a function.)

B. Continuous Wave Signals

1. Gravitational Waves

The continuous wave signal h_c is related to the phase via Equation (1.14). The condition for detectability is that

$$h_c > \frac{f_c}{\nu_o}\phi_{nc} \tag{3.5}$$

where ϕ_{nc} is the phase noise associated with this measurement. Collection of results from the various sources enumerated above yields

$$\phi_{nc}^2 > 4\pi^2\eta^2\frac{1}{2\tau f_{sam}} + \frac{a_{\ell o}}{2\tau}\frac{1}{f_c^2} + \frac{(a_{sw}+a_{tr})}{2\tau}\left(\frac{1}{f_c^{2.6}}\right) \tag{3.6}$$

where the factor $\left(\frac{1}{2\tau f_{sam}}\right)$ in the Doppler extractor term comes from assuming that the noise is spread equally over the frequency band from 0 to the sampling frequency f_{sam}. Here τ is the integration time. Comparing h_c to the phase noise term by term we find

$$h_c > \left\{\begin{array}{l} 3.2\times10^{-19}\left(\frac{\eta}{10^{-3}}\right)\left(\frac{0.1}{f_{sam}}\right)^{1/2}\left(\frac{f_c}{10^{-3}\,Hz}\right)^{1/2} \quad (3.7a)\\ 1.6\times10^{-17}\left(\frac{a_{\ell o}}{10^{-6}}\right)^{1/2} \quad (3.7b)\\ 1.2\times10^{-17}\left(\frac{a_{sw}+a_{tr}}{10^{-8}}\right)^{1/2}\left(\frac{10^{-3}\,Hz}{f_c}\right)^{0.3} \quad (3.7c)\end{array}\right\}\left(\frac{10^{10}\,Hz}{\nu_o}\right)\left(\frac{2\times10^7\,sec}{\tau}\right)^{1/2}$$

where an integration time of 2×10^7 sec and a gravitational wave frequency of 10^{-3} Hz were used as reference values. Continuous gravitation radiation is thus detectable at the 10^{-17} level with a single frequency Doppler tracking system. Although this amplitude is several orders of magnitude larger than any known sources, one should look anyway because, as was mentioned above, there are so many potential sources. Also, it seems plausible to me that if one can reduce the phase noise from solar wind and the troposphere by special techniques for the gravitational wave burst case, then one should also be able to do it in this case. This possibility should be studied.

I will finish this discussion on continuous waves by recalling a calculation of A. J. Anderson (1977). He analyzed Doppler residuals from the Pioneer-) spacecraft for any continuous signals. This was a single frequency S-band ($\nu_o = 2.3 \times 10^9$ Hz) tracking system. His analysis of the phase noise showed a flat spectrum with a value $S_\phi \approx 4.0 \times 10^{-3}$ $\text{rad}^2\text{Hz}^{-1}$. This means that he could have detected gravitational radiation of frequency f_c whose amplitude h_c was (using Equations (1.14) and (2.8))

$$h_c > \frac{f_c}{\nu_o} \left(\frac{S_\phi(f_c)}{2\tau} \right)^{\frac{1}{2}} \tag{3.8}$$

He found a signal at $f_c = 7.75 \times 10^{-3}$ Hz and he probably used $\tau \approx 2000$ sec; so,

$$h_c > 3.3 \times 10^{-15} \tag{3.9}$$

He interpreted his signal as being due to the rotation of the spacecraft with an amplitude of 0.24 cm. If interpreted as a continuous gravitational wave signal, the amplitude would be 2.5×10^{-14}. Looking at the data, the signal appears to be about 10 times the noise.

2. Oscillations of the Sun

The amplitude of vibration r_2 of a spacecraft caused by quadrupole oscillations of the Sun was estimated above for a distance of 4 $R_\odot$. The relationship between r_2 and the induced oscillating phase signal is given by (1.20). The detectability condition is

$$r_2 > \frac{c}{2\pi\nu_o} \phi_{nc} \tag{3.10}$$

ϕ_{nc} is the narrow-band phase noise from the various sources which have already been collected in (3.6) above. However, close to the Sun the solar wind dominates so we have

$$r_2 > \frac{c}{2\pi\nu_o}\left(\frac{a_{sw}}{2\tau f_2^{2.6}}\right)^{\frac{1}{2}} \tag{3.11}$$

At the distance of 4 $R_\odot$, a_{sw} is estimated from (2.16) to be $\sim$0.09. Thus (3.11) becomes

$$r_2 > (23\text{ cm})\left(\frac{10^{10}\text{Hz}}{\nu_o}\right)\left(\frac{a_{sw}}{0.09}\right)^{\frac{1}{2}}\left(\frac{\tau_2}{2\text{ hr}}\right)^{1.3}\left(\frac{2\times 10^5\text{sec}}{\tau}\right)^{\frac{1}{2}}$$

where $\tau_2 \sim 2$ hrs has been assumed for the oscillation period of the Sun. Hill and Candell (1978) have suggested that the Sun's oscillations are coherent for several days, so I have chosen an integration time of 2×10^5 sec ($\sim$2 days), which is the maximum time that the spacecraft would be expected to be in the vicinity of 4 $R_\odot$. The signal predicted by (1.19), assuming $J_2' \approx 10^{-7}$, is seen to be comparable to noise estimated by (3.9). Clearly it is an exciting prospect to contemplate measuring J_2' directly with a spacecraft. The closeness of the estimated signal to the estimated noise means that more accurate estimates of both are needed. We are working on this.

I wish to acknowledge helpful discussions with Warren Johnson, Kenneth Nordtvedt and Ian Roxburgh. This research was supported in part by the National Science Foundation, USA.

REFERENCES

Anderson, A. J. (1971), Nature 229, 547.

Anderson, A. J. (1977), Proceedings of the International Symposium on Experimental Gravitation, Pavia, Italy, Sept. 1976.

Anderson, J. D. (1974), Proc. of International School of Physics, "Enrico Fermi" Course LVI. Experimental Gravitation, ed. by B. Bertotti, Academic Press.

Cowling, T. G. (1941), Mon. Not. Roy. Ast. Soc., 101, 367.

Douglass, D. H. and Braginsky, V. B. (1978), Einstein Centenary Volume, ed. by S. Hawking and W. J. Israel, Cambridge University Press.

Estabrook, F. B. and Wahlquist, H. D. (1975), GRG 6, 439.

Hill, H. A. and Candell, T. P. (1978), submitted to Mon. Not. Roy. Ast. Soc.

Hill, H. A., Stebbins, R. T. and Brown, T. M. (1976), Atomic Masses and Fundamental Constants, 5, ed. by J. H. Saunders and A. H. Wapstra, Plenum, New York.

Johnson, W. and Douglass, D. H. (1978), to be published.

Severny, A. B., Kotov, V. A. and Tsap, T. T. (1976), Nature 256, 87.

Thompson, M. C., Wood, L. E., Janes, H. B. and Smith, D. (1975), Nov. IEEE Trans. on Antennas and Propagation, Vol. AP-23, 190.

Thorne, K. S. and Braginsky, V. G. (1976), Ap. J. 204, L1.

Wahlquist, H. D., Anderson, J. D., Estabrook, F. B. and Thorne, K. S. (1976), Proceedings of the International Symposium on Experimental Gravitation, Pavia, Italy, Sept. 1976.

Woo, R., and Armstrong, J. (1978), EOS, 59, 366.

Woo, R., Yang, F.-C., Yip, K. W., and Kendall, W. B. (1976), Ap. J. 210, 568.

Woo, R. (1975), Ap. J. 201, 238.

SECTION IX.

SPACECRAFT AND TRAJECTORIES

SOLAR PROBE STUDY

James E. Randolph
Jet Propulsion Laboratory

Introduction and Background

Past studies of the Solar Probe identified both low-thrust (Ref. 1) and ballistic (Refs. 2 and 3) modes of delivery of a spacecraft to the Sun. The concept discussed here was first suggested by Colombo et al. (Ref. 2) in early 1976. This was followed by a feasibility study at JPL (see Anderson et al., Ref. 3) which was focussed on a spin-stabilized spacecraft on a rectilinear trajectory as a minimum, first mission. As the Solar Probe technology has evolved, new mission modes have been introduced. New trajectory options are under consideration (see Bender, Yen, this volume) and new spacecraft designs have been developed (see Bender et al., Ref. 4).

The major objective of t. . present study at JPL is to identify realistic options for a Solar Probe Mission consistent with known technical, fiscal, and programmatic constraints. A second objective is to develop a program plan for NASA which includes identification of necessary research and development activities. A baseline mission and a "strawman" spacecraft design have been selected and used to determine the feasibility of meeting the mission requirements and the sensitivity to variations in those requirements.

Mission Options

Table 1 summarizes some of the Solar Probe mission options in four major cate.ories: delivery mode, period of the final orbit, spacecraft stabilization, and auxilary propulsion. The attributes of each option are qualitatively summarized by three measures of mission performance: scientific return, cost and risk. The lowest cost, baseline mission is a ballistic Jupiter gravity assist (JGA) which places a 3-axis stabilized spacecraft in a four-year heliocentric orbit. However, the scientific return is considered to be greatest for a spinning probe in a final orbit with a one year period, which would allow a second pass close to the Sun within the probable life of the spacecraft. The scientific return from a multiple-spacecraft mission would be very great.

TABLE 1: SOLAR PROBE MISSION OPTIONS

		Delivery Mode		Final Orbit Period		Spacecraft Stabilization		Auxiliary Propulsion			No.of Passes	Attributes			
Mission Option	Multi-S/C Option	Ballistic (JGA)	Low Thrust (Ion Drive)	Long Period (4.2 Years)	Short Period (1 Year)	Body(3 axis)	Spin	Ion Propulsion Module (IPM)	Large Aphelion ΔV (ΔV-EGA)	Drag Chute	(Assumed 5 year Lifetime)	Science Return	Cost	Risk	
1		X		X		X					1	Low	Low	Low	Least expensive, one pass, "fast delivery"
2			X		X	X		X			2	Med	High	Med	IPM Major Cost, "Free" Short Period
3		X		X			X		X		1	Med	Med	Med	Spinner Launch Mass Higher, May Require ΔV-EGA, one pass
4			X		X		X	X			2	High	High	High	Short Period Spinner, Dual Attitude Control and Power Subsystem Capability Required
5		X			X	X			X	X	2	Med	Med	Med	Chute Major development cost. May require ΔV-EGA
6		X			X		X		X	X	2	High	High	Med	SP Spinner, Spinner/Chute Compatibility, ΔV-EGA
7	Relay Probe		X		X	X		X			2	High	High	Med	Corona "Entry" probe, Possibl Launch Mass Problem
8	Dual Launch	X		X		X					2	High	High	Low	1 Rectilinear, 1 flyby

Delivery Mode

With present launch vehicle (Shuttle/IUS) capability, a significant velocity change (ΔV) is required following launch from Earth to deliver a spacecraft close to the Sun. A sufficiently large velocity increment can be attained using a gravity assist from Jupiter and/or ion drive propulsion. Table 2 summarizes the delivery modes in terms of launched mass, delivered mass, etc. The ΔV-Earth Gravity Assist (ΔV-EGA) ballistic option (row 2) can significantly increase the mass delivered on a JGA trajectory. Ion drive trajectories are characterized by the number of "loops" around the Sun before reaching the final orbit. Type A or single-loop trajectories require a large ion propulsion system and have a low delivered mass, but a short flight time. Type B or dual-loop trajectories can deliver more mass with a smaller propulsion system, but at the cost of a longer flight time and thermal problems during the perihelion pass between loops. The ballistic and ion-drive trajectories are discussed in greater detail in the papers by D. F. Bender and C. W. Yen, respectively.

The footnote in Table 2 refers to a drag chute (a concept due to G. Colombo) which, if deployed after the Jupiter swingby, could use solar radiation pressure over a long period of time to reduce the spacecraft velocity and aphelion distance. The estimated mass of such a parachute is 200 kg.

In addition to the orbital period, other "open" parameters at this time are the perihelion distance, the inclination, and the nodal angle between the line of apsides and the Sun-Earth line. These parameters can be chosen to optimize the scientific return from the mission. As part of the "strawman" design, we have assumed perihelion at 4 R_s, an inclination of 90^o, and a nodal angle of 45^o.

Spacecraft Design and Thermal Control

Early in the study, it was recognized that the method of stabilizing the attitude of the spacecraft has a major impact on the configuration and mass of the spacecraft. It is estimated that a spin-stabilized spacecraft would be about 200 kg heavier than a 3-axis stabilized spacecraft. The current estimate of the Shuttle/IUS performance precludes the use of a spin-stabilized or dual-spin spacecraft on a simple JGA trajectory. Thus, our strawman spacecraft is 3-axis stabilized. Since other delivery modes would allow this greater mass and since the fields and particles experiments would

TABLE 2: SOLAR PROBE DELIVERY MODE SUMMARY

	ORBIT TYPE	LAUNCH C_3 (KM^2/SEC^2)	LAUNCH MASS (KG)	DELIVERED MASS (KG)	FLIGHT TIME (YEARS)	FINAL PERIOD (YEARS)
1.	BALLISTIC (JGA)	117	840	800	3.00	4.2
2.	BALLISTIC (ΔV-EGA, JGA)	33	3300	1500	5.80	4.2*
3.	ION DRIVE (TYPE A)	7.5	5590	740	2.05	1.0
4.	ION DRIVE (TYPE B)	22	4115	900	3.85	1.0

*LARGER PAYLOAD COULD SUPPORT A "DRAG CHUTE" TO REDUCE THE FINAL PERIOD TO 1 YEAR.

benefit if the spacecraft were spinning, the design of a dual-spin spacecraft is also underway and will be completed within the next few months. If an ion-drive option were selected, the spacecraft design would have to be integrated with the ion propulsion system consisting of thrusters, large solar arrays, etc.

A spacecraft configuration has been conceived which is consistent with the baseline mission (i.e., a JGA ballistic trajectory resulting in a 4 R_s perihelion and a 4.2-year final period). Figure 1 illustrates some of the characteristics of this design. The structure is dominated by the "roof" thermal shield and the secondary shields which cast a shadow over the rest of the spacecraft. During perihelion passage, the spacecraft must rotate through 180^o in about 14 hours to keep the shield pointed at the Sun. Estimates of the temperature profile expected in the main part of the spacecraft behind the thermal shield are given in Figure 2. The baseline perihelion distance of 4 solar radii was chosen on the basis of this figure. Just below the heat shields are two large radiators for the Selenide Isotope Generators (SIG's) which supply power (450 w) to the spacecraft.

Some of the scientific objectives of the Solar Probe Mission can best be met if the spacecraft follows a purely ballistic or drag-free trajectory. In the strawman payload, the drag-compensation sensor is sandwiched between the SIG radiators. This sensor must be isolated from the remainder of the spacecraft to minimize biases from asymmetries and changes in the self-gravity of the configuration.

A large-diameter, high-gain antenna is suspended from the rear of the spacecraft by multi-degree-of-freedom linkage which allows the complex deployment and articulation needed to point at the Earth through the perihelion passage.

Figure 3 is a schematic drawing of a possible set of scientific instruments mounted on the electronics compartment.

Telecommunication and Data Handling

The strawman telecommunication system consists of two 20 watt transponders which allow two-way coherent S- and X-band data. The S-band link is for carrier (Doppler) tracking and ranging only. The X-band link also carries about 1000 bits per second of telemetry data at perihelion. About 10^7 bits of data can be stored to buffer high-rate data or to provide some backup in case of loss of real-time data. With this dual-frequency system,

FIGURE 1. SOLAR PROBE STRAWMAN DESIGN

- CONFIGURATION - EVOLVING
  - THERMAL SHIELDS CASTING SHADOW CONE
  - SELENIDE ISOTOPE GENERATORS (SIG'S) AND RADIATORS
  - DRAG COMPENSATION SENSOR ISOLATION
    - SELF-GRAVITY PROBLEM
  - HIGH GAIN ANTENNA ARTICULATION
  - SCIENCE INSTRUMENT DEPLOYMENT AND ARTICULATION
- STRAWMAN SCIENCE PAYLOAD
- TELECOMMUNICATIONS
  - DUAL FREQUENCY (X/S) UPLINK AND DOWNLINK
  - 20 WATT TRANSMITTERS, 1.5 HIGH GAIN ANTENNA (VIKING)
  - 1000 BPS @ PERIHELION ($4R_S$), 10^7 BITS STORAGE
  - 0.1 MM/SEC DOPPLER ERROR
- ATTITUDE CONTROL PROPULSION
  - UNIFIED SYSTEM CONCEPT
  - BI-STABLE ATTITUDE
  - DRAG COMPENSATION SYSTEM, 10^{-10} G DESIGN GOAL
  - MOMENTUM WHEELS FOR ATTITUDE CONTROL

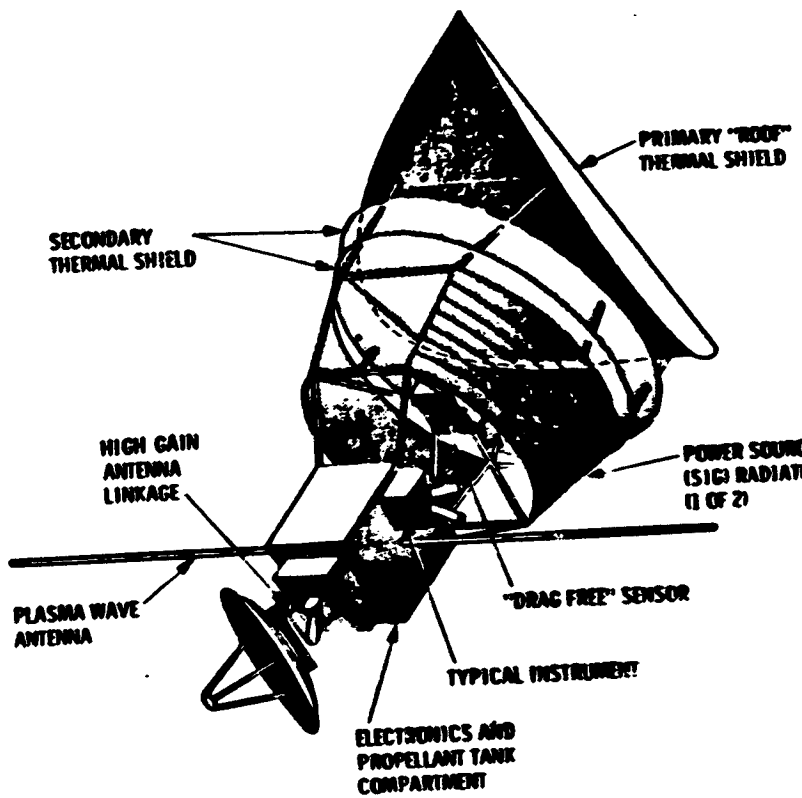


FIGURE 2: SOLAR PROBE INTERNAL BUS TEMPERATURE ESTIMATE

ROOF SHIELD

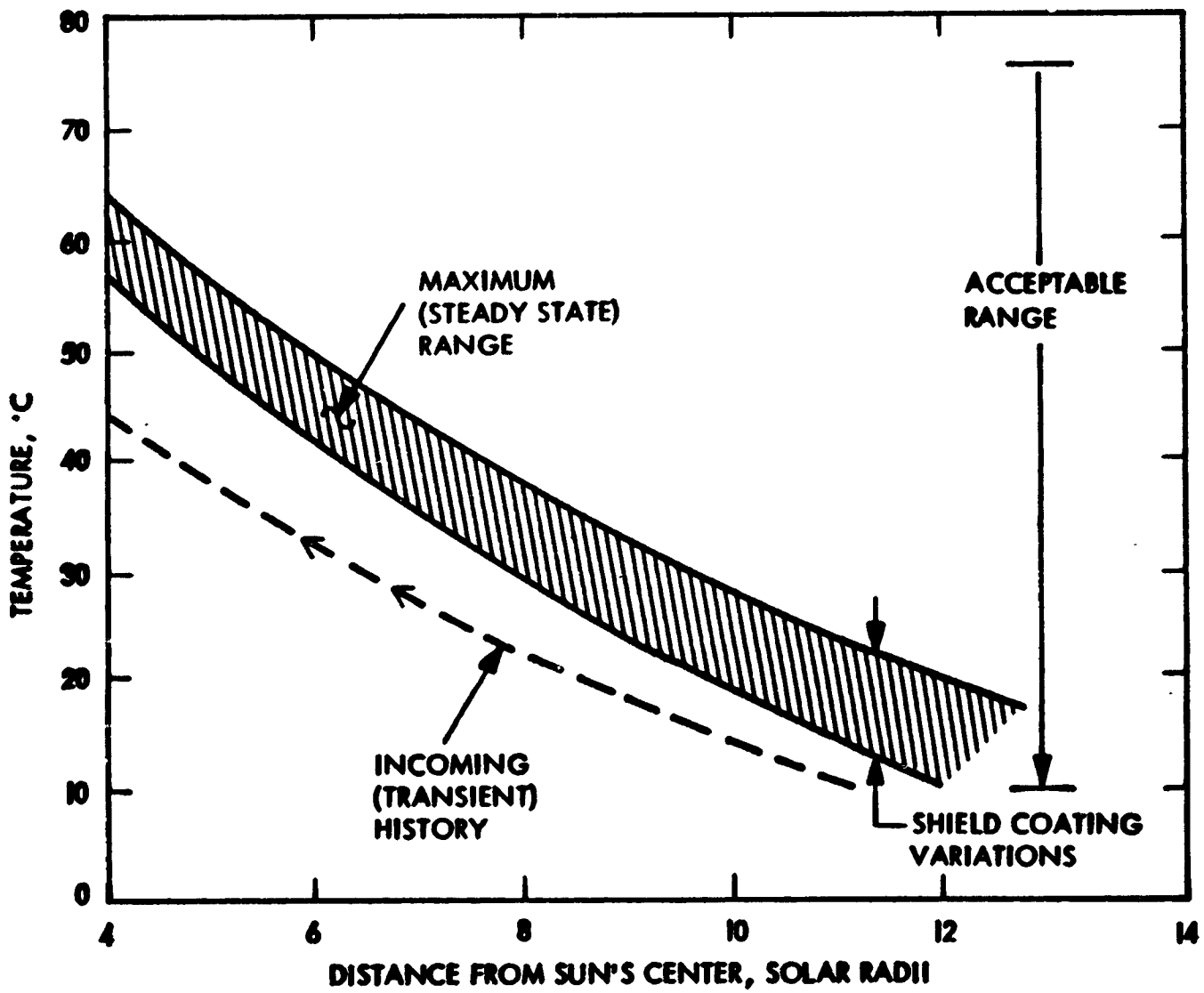
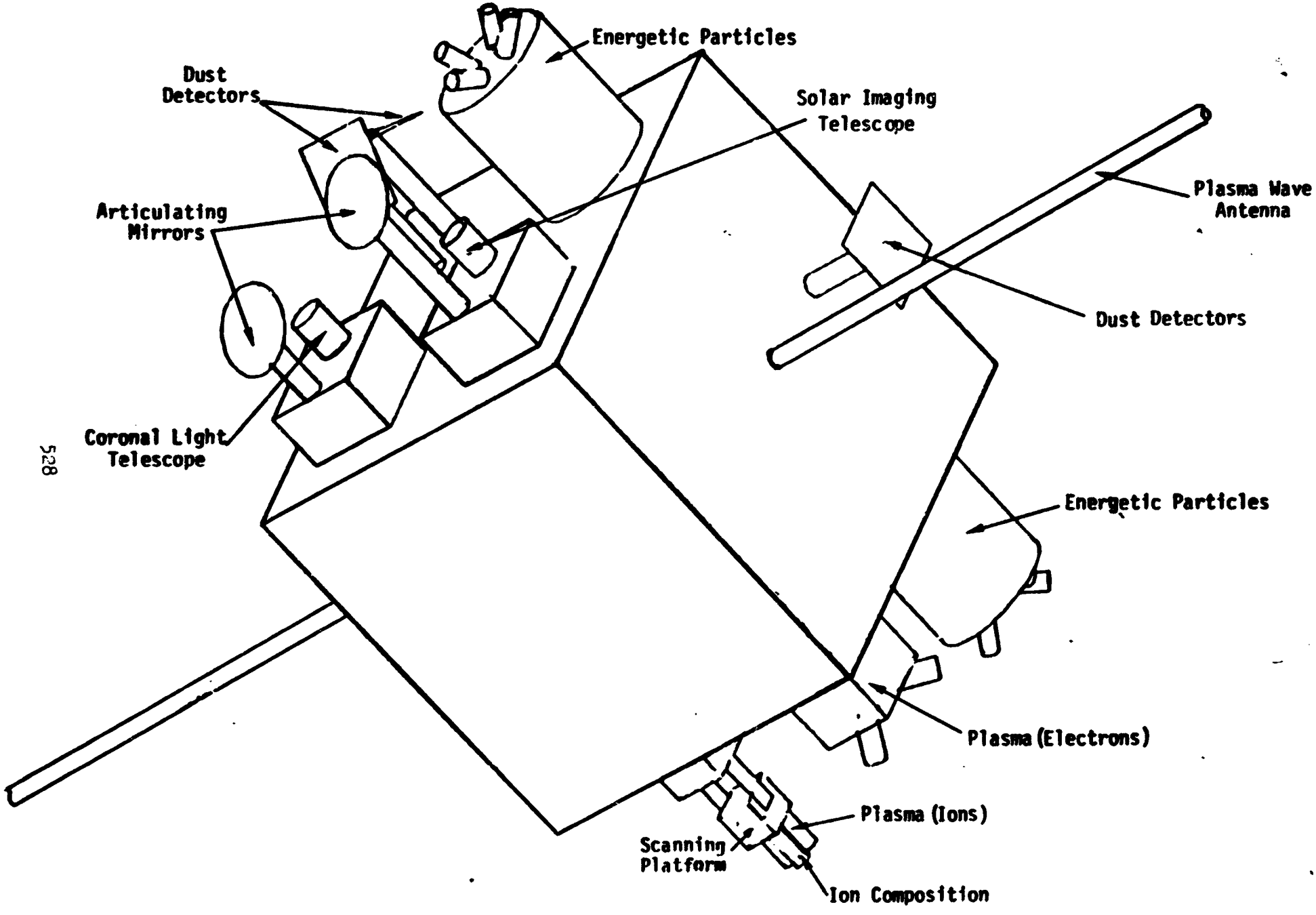


FIGURE 3. SOLAR PROBE STRAWMAN PAYLOAD CONFIGURATION



the Doppler tracking error is estimated to be in the range 0.1 to 1.0 mm/sec near perihelion; see Table 3. The significance of this error estimate is discussed in later papers.

Attitude Control

The strawman attitude control system is consistent with the requirements for drag compensation, attitude stabilization, and midcourse maneuvers. When the spacecraft is outside the Earth's orbit, the high-gain antenna cannot view the Earth unless the heat shield is on the side of the spacecraft away from the Sun. Thus the spacecraft attitude is bistable, with the heat shield pointed either at the Sun or 180° away.

A cartoon illustrating the drag-compensation system is shown in Figure 4. The fundamental concept is to keep a proof mass centered in a cavity. The isolation of the sensor, discussed earlier, is expected to yield drag compensation to an accuracy of 10^{-10} g; the significance of this accuracy is also discussed in later papers.

Mass and Power Summaries

Tables4 and 5 list the mass and power by subsystem for the strawman design. From Table 4, it is apparent that the thermal control system accounts for about one quarter of the entire spacecraft mass. The dominant power consumer is the telecommunications subsystem; Table 5 shows that nearly one third of the total power is needed for the high-power, dual frequency communications at perihelion.

Research Required

The thermal shield design is only in the conceptual stage at this time. A significant amount of research in materials and manufacturing processes is needed before a final design is possible. An understanding of the extent of "thermal" radiation beyond the visible solar disk is also crucial to the final design of the heat shield.

If the required tracking accuracy is about 0.1 mm/sec, development of a dual-frequency, two-way telecommunication system is clearly essential. The effect of the solar corona on the telecommunications link must be modeled to provide confidence in the design.

Support has also been requested for research into the theory and design of a drag-compensation system for a complex spacecraft with articulating parts. The dynamic range of the drag-compensation system can be reduced if a reasonably accurate model of the expected drag can be developed.

Table 3. SOLAR PROBE DOPPLER TRACKING ACCURACY ESTIMATES AT 4 SOLAR RADII

RADIO LINK CONFIGURATION	ACCURACY (mm / sec)	
• S-BAND UPLINK / S-BAND DOWNLINK	25.0	
• S-BAND UPLINK / S & X-BAND DOWNLINK	5.0	EXTRAPOLATED FROM S-BAND DATA
• X-BAND UPLINK / X-BAND DOWNLINK	2.5	
• X-BAND UPLINK / X & S-BAND DOWNLINK	0.5	
• X & S-BAND UPLINK / X & S-BAND DOWNLINK	0.1	

FIGURE 4: THE DRAG COMPENSATION CONCEPT

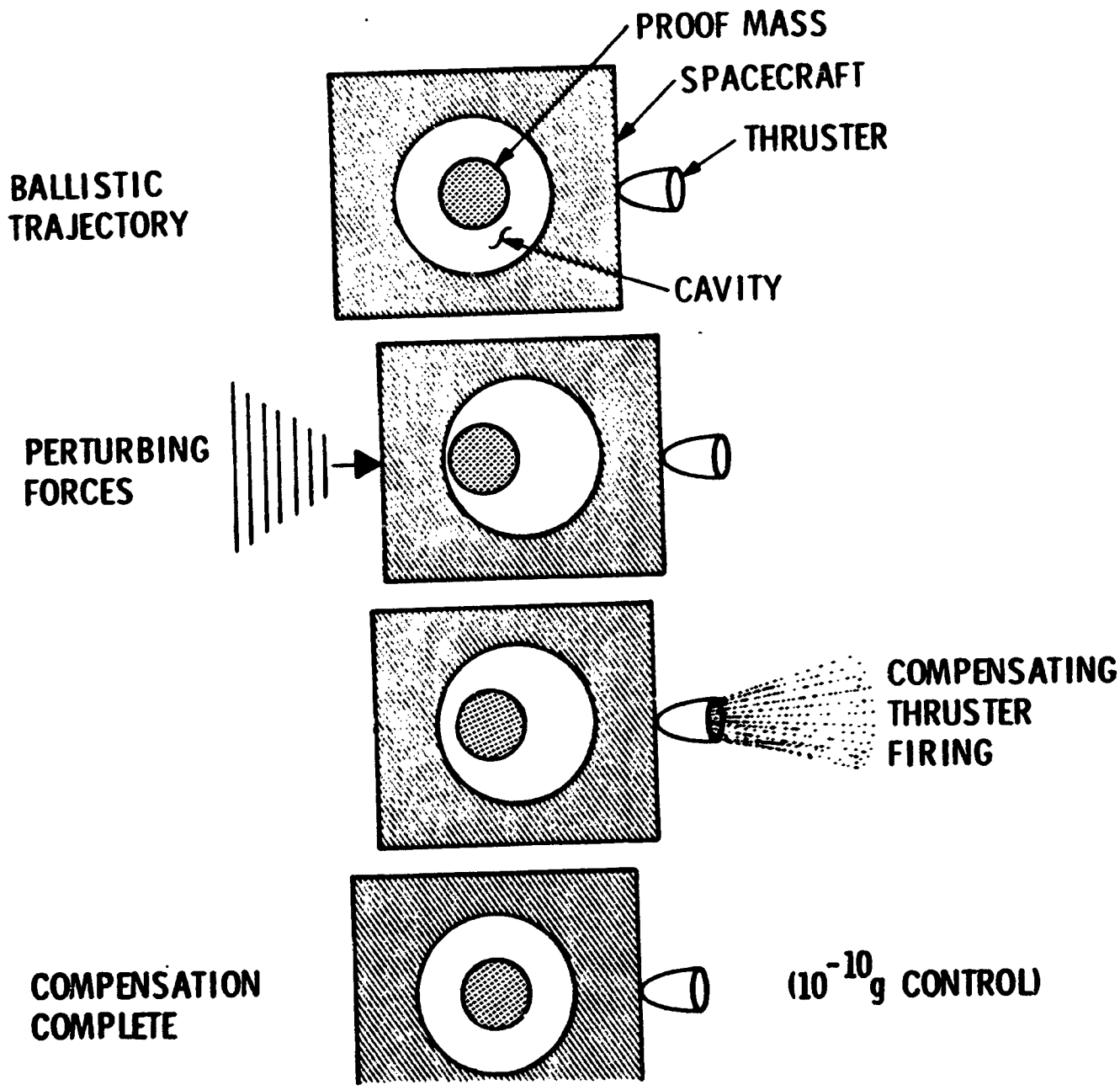


TABLE 4. SOLAR PROBE MASS SUMMARY

SOLAR PROBE MASS SUMMARY
3 AXIS STABILIZED CONFIGURATION
(PRELIMINARY ESTIMATES)

SUBSYSTEM	MASS (KG)	PERCENT OF TOTAL	HERITAGE
SCIENCE	60	7	SPM*/GALILEO/NEW
STRUCTURE	80	9	NEW/CID**/SPM
DEVICES	15	2	NEW/SPM
COMMAND AND DATA HANDLING	40	5	SPM
ATTITUDE CONTROL/PROPULSION SYSTEM:	(205)	(24)	
CONTROL HARDWARE AND ELECTRONICS (INCL. WHEELS)	85	10	NASA STD/GALILEO
DRAG FREE SENSOR AND ELECTRONICS	5	1	TRIAD
PROPULSION HARDWARE AND ELECTRONICS	15	2	VOYAGER
PROPULSION TANK	15	2	VOYAGER
AACS PROPELLANT/ORBIT	5	1	NEW
DRAG FREE PROPELLANT/ORBIT	35	4	NEW
ΔV PROPELLANT	45	5	SPM
PYROTECHNICS	5	1	CID/SPM
TEMPERATURE CONTROL	200	24	NEW
CABLING	35	4	CID/SPM
POWER (450 WATT SIG's)	75	9	NEW SIG's (7.7 w/kg)/SPM
TELECOMMUNICATIONS	40	5	VOYAGER/SPM
IUS ADAPTER	35	4	NEW/SPM
CONTINGENCY	50		
TOTAL	840		

* SPM = Solar Polar Mission
** CID = Comet Ion Drive

TABLE 5. SOLAR PROBE POWER SUMMARY

3 AXIS STABILIZED CONFIGURATION

(PRELIMINARY ESTIMATES)

SUBSYSTEM	SOLAR PERIHELION POWER (w)	PERCENT OF TOTAL	HERITAGE
SCIENCE	45	10	SPM/GALILEO/NEW
TELECOMMUNICATIONS	146	32	VOYAGER/SPM
POWER CONDITIONING	9	2	SPM
COMMAND AND DATA	21	5	SPM
DATA STORAGE (BUBBLE MEMORY)	11	2	SPM
ATTITUDE CONTROL/PROPULSION	60	13	GALILEO/SPM
MOMENTUM WHEELS (4)	85	18	NASA/STD
DRAG FREE	13	3	TRIAD
	390		
DEGRADATION AND CONTINGENCY	60	15	
	450		

The question has also arisen whether or not the perihelion distance might be limited by radiation damage by energetic particles. This question was addressed by a Solar Probe Environmental Workshop held at JPL on January 19 and 20, 1978. The concensus was that, for perihelion at 4 R_s, the flux and fluence of energetic particles would probably be less than that encountered during the Jupiter flyby. However, the Report of the Environmental Workshop (available from JPL) indicates that there are many unknown factors which remain to be determined.

REFERENCES

1. Close Approach Solar Probe, AVCO Technical Report RAD-TR-63-35, 23 September 1963

2. Colombo, G., Lautman, D., Pettingill, G., "An Alternative Option to the Dual-Probe Out-of-Eliptic Mission via Jupiter Swingby, NASA TM-X-71097, p. 37-47, March 1976

3. Anderson, J., Colombo, G., Freidman, L., Lau, E., "An Arrow to the Sun" presented at the International Symposium on Experimental Gravitation, Pavia, Italy, Septmber 1976

4. Bender, D.F., Neugebauer, M.M., Randolph, J.E., Yen, C.L., "Preliminary Mission Design Studies for the Close Solar Probe Mission", AIAA paper number 78-1440, 7 August 1978

D41

BALLISTIC TRAJECTORIES

D. F. Bender

Jet Propulsion Laboratory

N78-33005

The only ballistic trajectory mode feasible for a close solar probe or for an orbit inclined approximately 90° to the ecliptic is the Jupiter gravity assist mode. Figure 1 shows a comparison of the trajectories of the Solar Polar and the Solar Probe Missions for 1983 launches. For any other launch year, the figure would be almost identical if the Jupiter encounter is placed at the edge of the figure as shown. The ticks are at 100 day intervals.

The inner part of the Solar Probe trajectory is plotted in Figure 2. To the accuracy required for preliminary discussions, the geometry of the solar encounter phase is practically the same for the 4.3 year orbit achieved by a Jupiter gravity assist and for a one year orbit (which might be attained with either a drag chute or a retro-rocket maneuver following a Jupiter gravity assist).

Tables 1 and 2 list data describing the geometry of an orbit with perihelion at 4 solar radii and aphelion at Jupiter.

Figure 3 shows the range of apparent directions of the solar wind if it is flowing radially outward from the Sun with a speed of either 150 or 300 km/sec.

The earth-node angle η is defined as the solar longitude of the Earth at the time of the solar encounter minus the longitude of the node of the trajectory (assumed to be in the ecliptic). Figure 4 shows the minimum Sun-Earth-probe angle during the solar encounter as a function of the Earth-node angle and the orbital inclination. If the inclination is 60° or more, the minimum SEP angle is not greatly different from the 90° value.

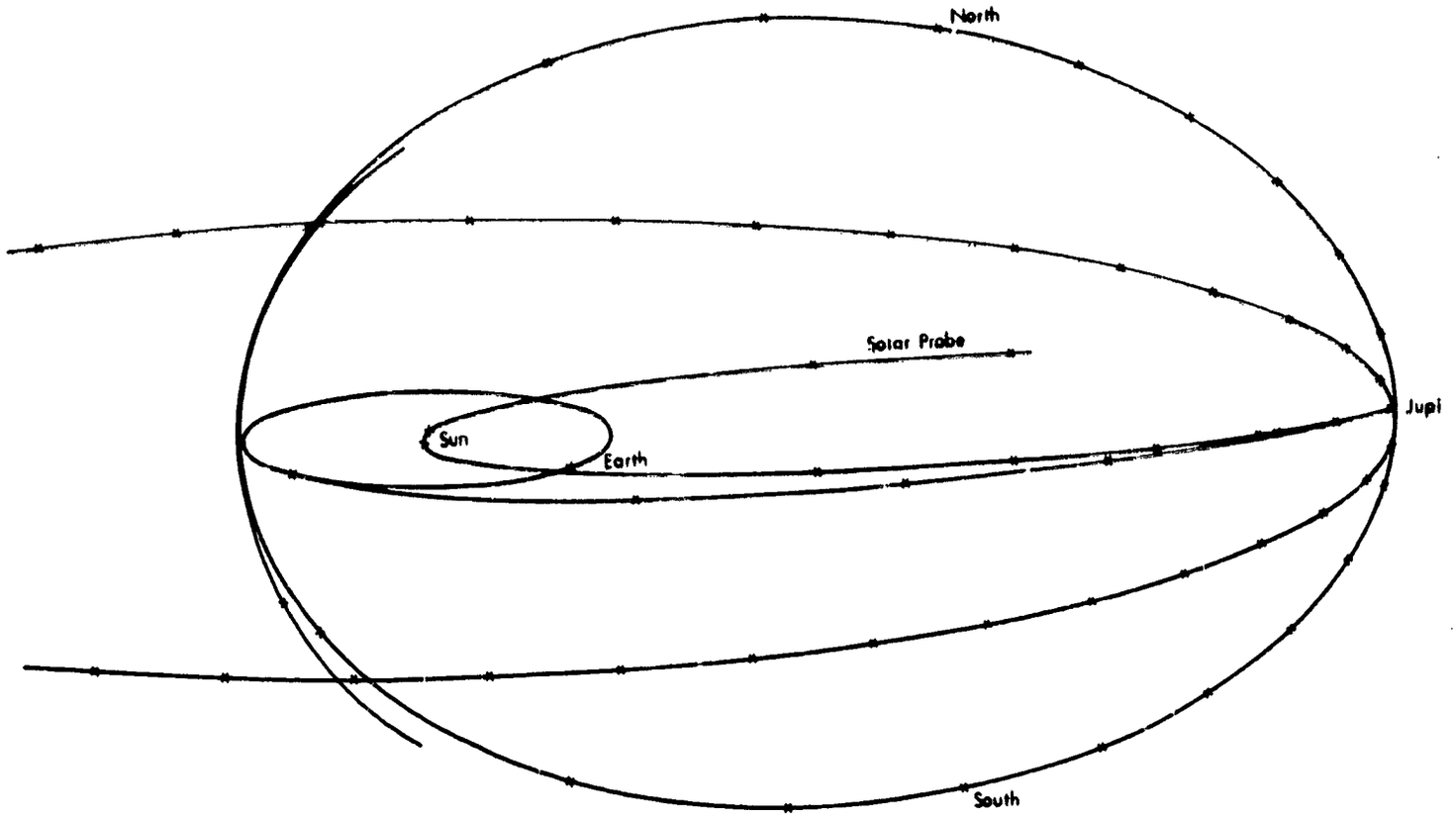


Figure 1. Comparison of Solar Probe and Solar Polar Mission trajectories.

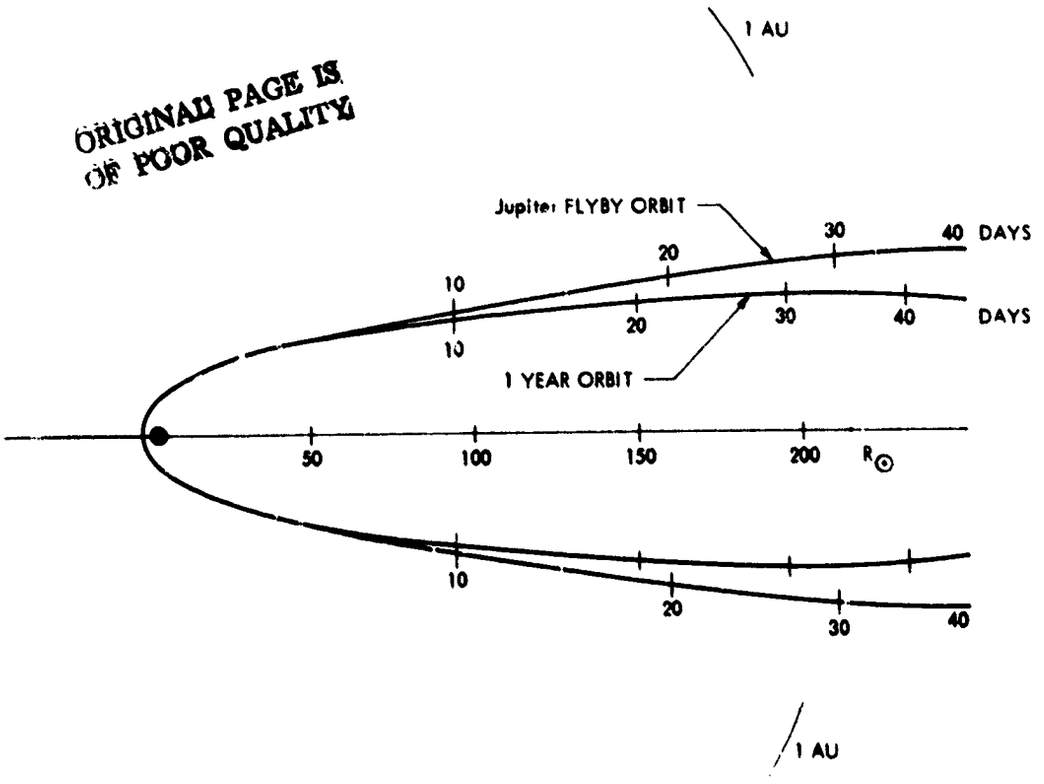


Figure 2. Comparison of trajectories for 1-year and 4.2-year orbits.

Table 1. Near Sun Trajectory Data 4 R_S Orbit

True Anomaly* (deg)	Radius+ (Solar Radii)	Time from Perhelion (hours)	Circumferential** Velocity (km/s)	Radial Velocity (km/s)	Angular Rate (°/min)	Shadow Cone Angle
0	4.00	0	308.2	0	.381	14.5
10	4.03	.44	305.9	26.6	.375	14.4
20	4.12	.99	299.0	51.5	.368	14.0
30	4.29	1.38	287.6	76.1	.332	13.5
40	4.53	1.91	272.3	98.7	.297	12.8
50	4.87	2.51	253.4	117.6	.257	11.8
60	5.32	3.22	231.4	133.0	.215	10.8
70	5.95	4.08	207.2	144.3	.172	9.7
80	6.80	5.19	181.3	151.2	.132	8.5
90	7.97	6.67	154.7	153.6	.096	7.2
100	9.63	8.77	128.0	151.2	.066	6.0
110	12.07	12.0	102.1	144.3	.042	4.8
120	15.83	17.2	77.9	133.0	.024	3.6
130	22.03	26.8	56.0	117.6	.012	2.6
140	33.30	47.1	30.0	98.7	.006	1.7
145	42.70	66.7	28.9	88.1	.003	1.3
150	56.88	100.1	21.7	76.1	.002	1.0
155	79.56	162.7	15.5	64.9	.001	.7

* Also latitude (or 180-lat) if polar and perihelion on Equator

\+ 1 AU = 215 Solar Rad.

** Component perpendicular to Radius

Table 2. Nominal Solar Probe Trajectory Orbital Elements

Earth to Jupiter

Launch: April 15, 1985, C_3 = 115.7 km^2/s^2, DLA = -27.1°
Time of Flight: 462 days to July 21, 1986
Elements:* a = 6.4323 AU = 962.2645 Gm
e = .844024 ω = 179.8126
i = 1.7328° M = .0023° (at launch)
Ω = 24.5646° V_{hp} = 13.8 km/s

Jupiter Centered Elements* July 21, 1986

North Going	South Going
B·T = -875600 km	B·T = -868700 km
B·R = 87000 km	B·R = -64580 km
a = -662098 km	a = -662098 km
e = 1.66342	e = 1.65246
i = 174.3198	i = 175.7389
Ω = -37.8486	Ω = 135.4436
ω = -50.1598	ω = 123.3914
q = 6.15 R_j	q = 6.05 R_j

Jupiter to Sun Elements* July 21, 1986 to April 1988

North (Des Node at Sun)	South (Asc Node at Sun)
a = 2.52729 AU	a = 2.52927 AU
= 378.0772 Gm	= 378.3734 Gm
e = .99261	e = .99263
i = 89.661	i = 90.282
Ω = -17.389	Ω = 162.610
ω = 178.117	ω = .414
M = -156.735 (at Jupiter)	M = -158.7838
n = 46	n = 43
TF = 638.6 days	TF = 635.8 days

* All angles in degrees, reference system Earth Ecliptic (1950)

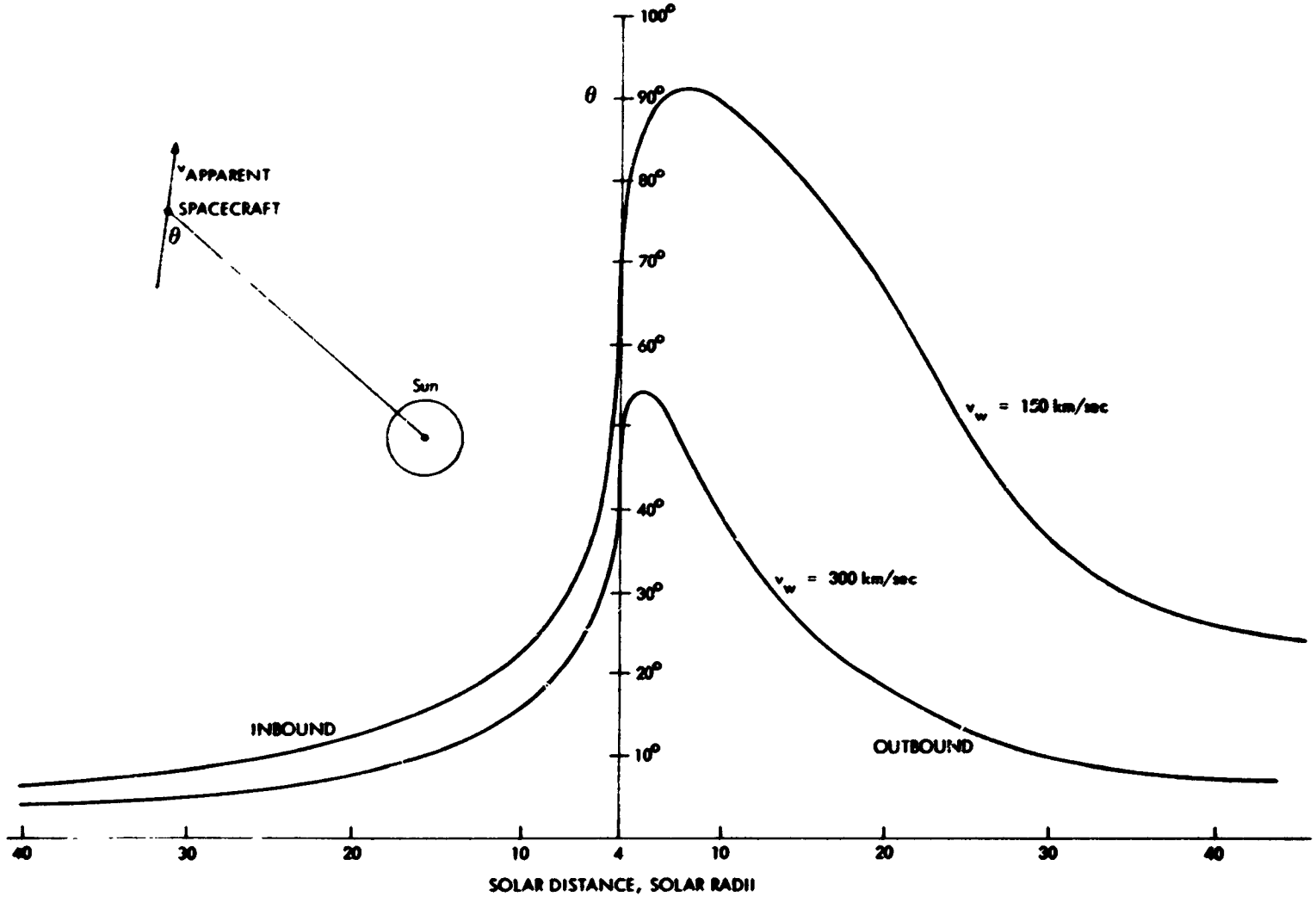


Figure 3. Apparent direction of the solar wind if the velocity is 150 or 300 km/sec.

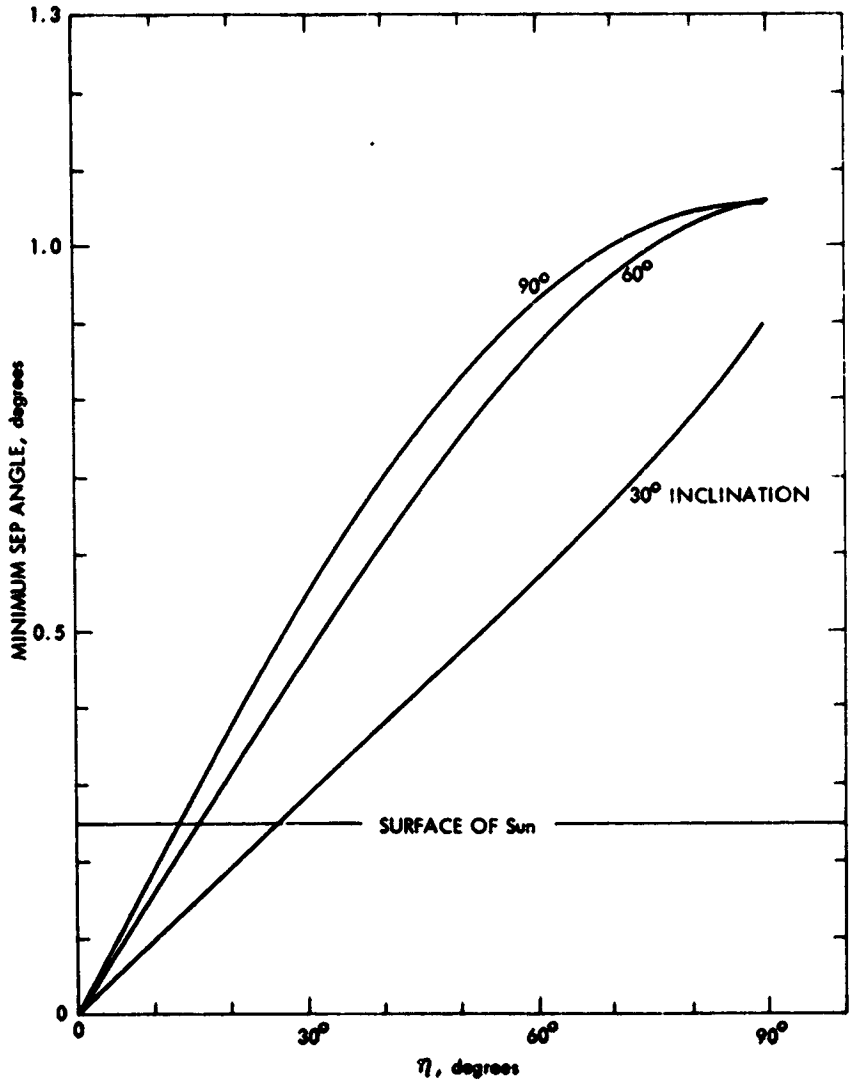


Figure 4. Minimum Sun-Earth-probe angle as a function of Earth-node angle and inclination.

The fact that any value of Earth-node angle and any inclination can be attained by means of a Jupiter gravity assist is illustrated in Figure 5, which is a plot of launch C_3 versus arrival time for a launch on April 15, 1985. The present estimate is that 830 kg can be launched on this trajectory for $C_3 = 115$ km^2/s^2. Unfortunately, the C_3 requirements increase for several years after 1985 and will not be lower again until 1992 for these trajectories.

More mass can be delivered ballistically if the direct launch from Earth is replaced by a ΔV-EGA trajectory. Such a trajectory is shown in Figure 6. The spacecraft is launched into a loop to about 2.5 AU, where a propulsion maneuver gives a large velocity change, ΔV. The spacecraft then performs an Earth gravity assist and flies on to Jupiter. The Earth flyby could be performed either a few months before or a few months after two or three years after launch (these are called 2^-, 2^+, 3^-, and 3^+ trajectories). For the 2^+ trajectory, shown in Figure 6, the C_3 is 32 km^2/s^2 and the velocity increment required is 2.62 km/s; with such a trajectory, the final spacecraft mass would be about 1500 kg and solar encounter would occur about 5 years after launch. A still greater payload is possible using a 3^+ trajectory which has a launch C_3 of 52 km^2/s^2 and a ΔV of 0.68 km/s; this system could deliver over 1900 kg to 4 solar radii seven years after launch.

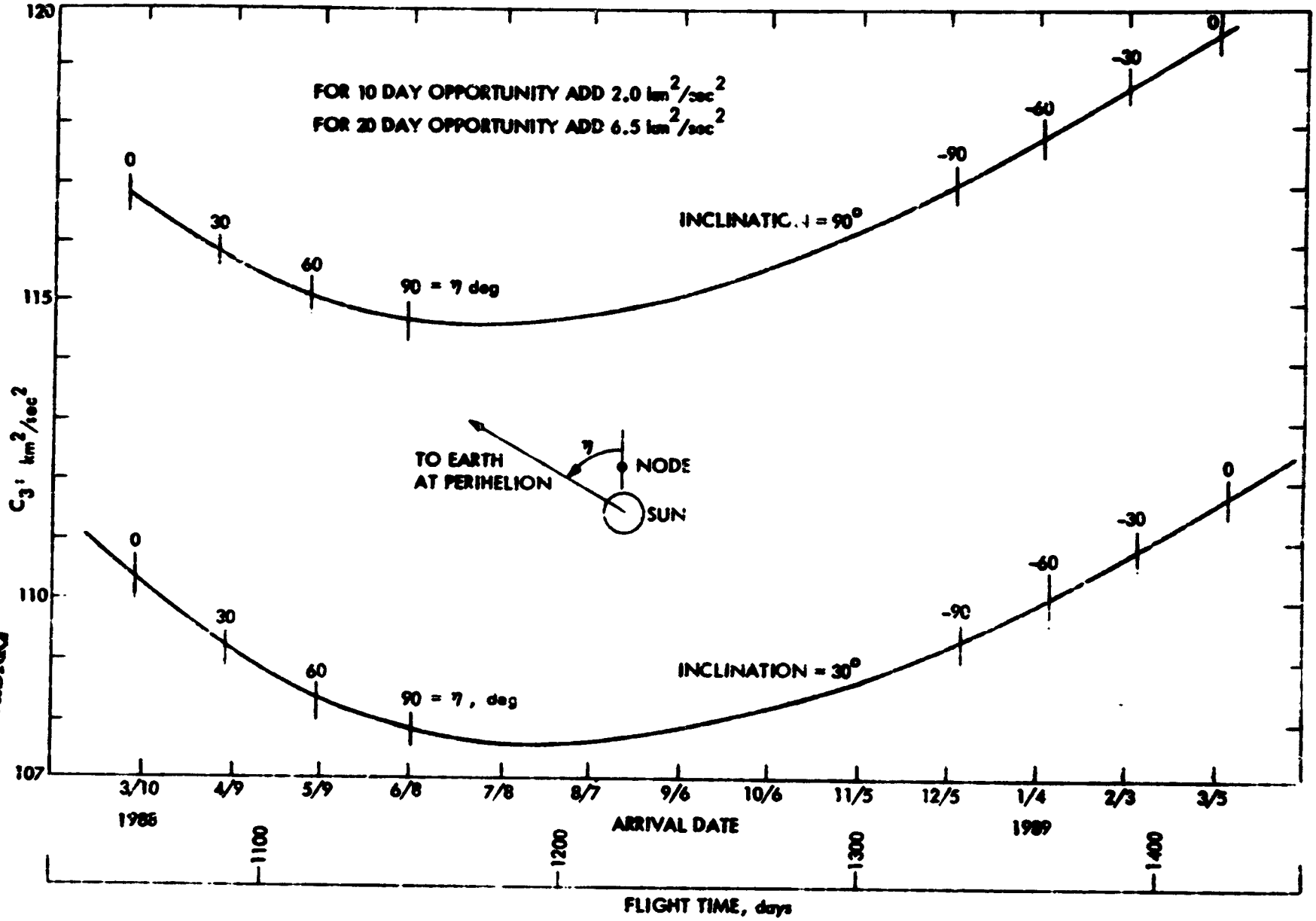


Figure 5. Launch C_3 versus arrival date for 1985 ballistic solar probe opportunity.

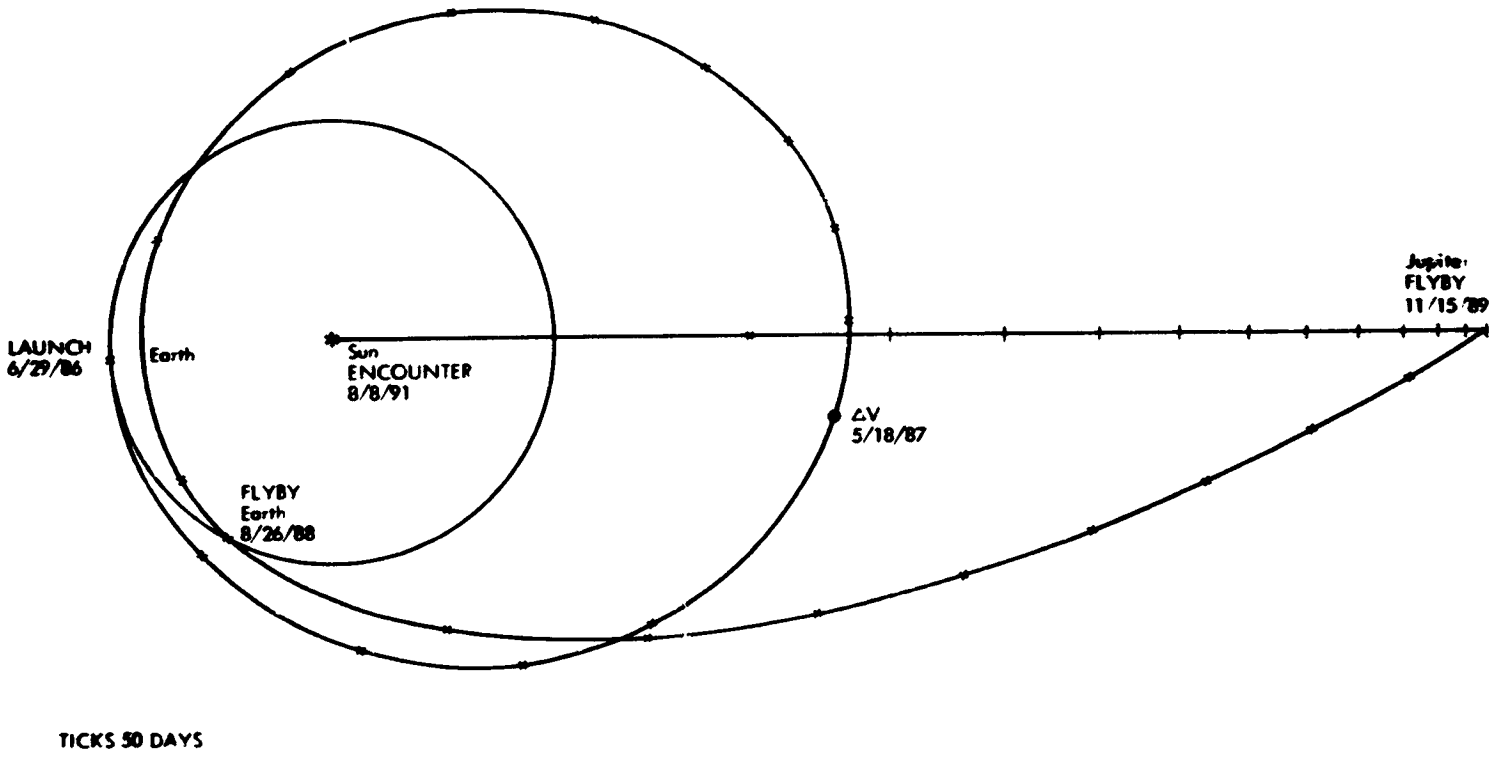


Figure 6. A 2^{+} ΔV-EGA solar probe trajectory.

D42

N78-33006

ION DRIVE PERFORMANCE AND TRAJECTORIES

C. L. Yen

Jet Propulsion Laboratory

This paper is concerned with the use of Solar Electric Propulsion (SEP) for the Solar Probe Mission. The dependence of the payload mass on the size of the SEP and the flight time is described on the basis of preliminary data.

First, for those with little familiarity with SEP, we summarize the range of hardware performance expected to be available in the 1980's. The power source is a solar array with concentrators for extra sunlight collection. The concentration factor is expected to be in the range of 2 to 4 and the total power between 20 and 60 kw. This power is used to accelerate mercury ions to 30-45 km/sec, which corresponds to a specific impulse I_{sp} = 3000 - 4500 sec, or perhaps even higher.

The final orbit is assumed to be in the ecliptic with a perihelion distance of 0.02 AU and a period of 1 year. Less ambitious, longer period orbits were not considered to be worth the advanced nature and greater cost of SEP compared to conventional chemical propulsion systems. High inclination orbits are also feasible, but have not been studied yet. The velocity increment, ΔV, needed to achieve the assumed final orbit is 20-25 km/sec. If the shuttle is used to escape from the Earth with C_3 = 0 and SEP is used to supply all the required ΔV, then a crude upper limit to the SEP performance can be obtained from the simple rocket equation

$$M_N = M_o e^{-\Delta V/I_{sp}g} - M_{ipm}$$

where M_N is the payload mass, M_o is the mass delivered to Earth escape by the shuttle, and M_{ipm} is the dry mass of the ion propulsion module. For M_o = 6200 kg, and I_{sp} = 4300 sec, M_N = (3400-3800) kg - M_{ipm}. The magnitude of M_{ipm} depends on the SEP size which, in turn, depends on the flight time of the mission.

There are several classes of optimal low-thrust trajectories for the Solar Probe Mission. This classification is based on the number of orbital revolutions (NREV) the SEP spacecraft makes before it completes the transfer to the final orbit. Trajectory types A, B, C correspond to one, two, and three orbital revolutions, respectively.

The typical flight time, SEP size, and other performance factors associated with each trajectory type are summarized in Table 1. Performance is expected to improve as NREV is increased. Nonetheless, the duration of the mission and the anticipated thermal problems make it impractical to consider trajectory types with NREV greater than 2. A type C trajectory (NREV = 3) is feasible only if the thermal problems can be alleviated. One possibility is to use the asteroid Icarus as a heat shield. With a period of about 1 year and a perihelion distance of 0.18 AU, Icarus appears to be a suitable object for this purpose. Another method of enhancing performance is to incorporate a Venus swingby in the SEP flight. Since the analysis for this hybrid mission mode is more difficult, little data have been generated, but the trend in improved performance can be clearly shown.

Figures 1 through 4 show plots of transfer trajectories corresponding to type-A, type-A with a Venus Swingby, type B, and type-C with an Icarus Rendezvous. In each of these examples, the flight time and the SEP size were selected to give a delivered mass ranging from 700 to 1000 kg, the range of payload deemed reasonable by the study team. These examples do not represent upper limits of SEP performance nor the only available geometry and one should not infer the SEP capability from these few data points. The results of a preliminary parametric investigation of the performance as a function of SEP size and flight time are shown in Figures 5 through 7. These data are too crude to form the basis for drawing final conclusions, but a clear indication of the SEP performance trend has emerged.

A summary of the SEP performance is given in Table 2. The following concluding remarks can be made:

(1) To do the mission in a two year flight time, one must fly a type-A trajectory. The required SEP size is very large; a

Table 1. TRAJECTORY CHARACTERISTICS

Type	No. of Rev.	Flight Time, Days	SEP-Size	Characteristics
A	1	650 - 800	Very Large	Simple Path
			∿60 kw/(4:1)	Insufficient Payload
			∿2700 kg	Venus Swingby Can Help
B	2	1200 - 1450	Moderate	Thermal Protection needed for SEP
			∿35 kw/(2:1)	
			∿1700 kg	Sufficient Payload
C	3	1700	Moderate	Thermal Protection needed for SEP
			∿35 kw/(2:1)	
			∿1700 kg	Sufficient Payload

Table 2. PERFORMANCE SUMMARY

Type	Flight Time (Days)	Power (kw)	Concentration Factor	Payload (kg)	Comments
A	650 - 800	55-60	4	700	Launch vehicle capability limited, no growth potential.
A/VGA	650 - 800	55-60	4	1000	Higher performance expected if swingby conditions and the trajectories are optimized, but probably not by more than 200 kg. Launch window constraint exists.
B	1300 - 1450	30	2	900	R_p = 0.256 AU R_p = 0.3 AU thermal problem?
	1200	48	2	1200	Performance growth expected for larger power.
B/VGA	NO DATA AVAILABLE YET				Payload gain over type B is not expected to be large.
C/Icarus	1720	35	2	1070	R_p = 0.47 AU Performance growth expected at higher power. Launch window constraint exists. Too long a flight?

R_p = intermediate perihelion distance

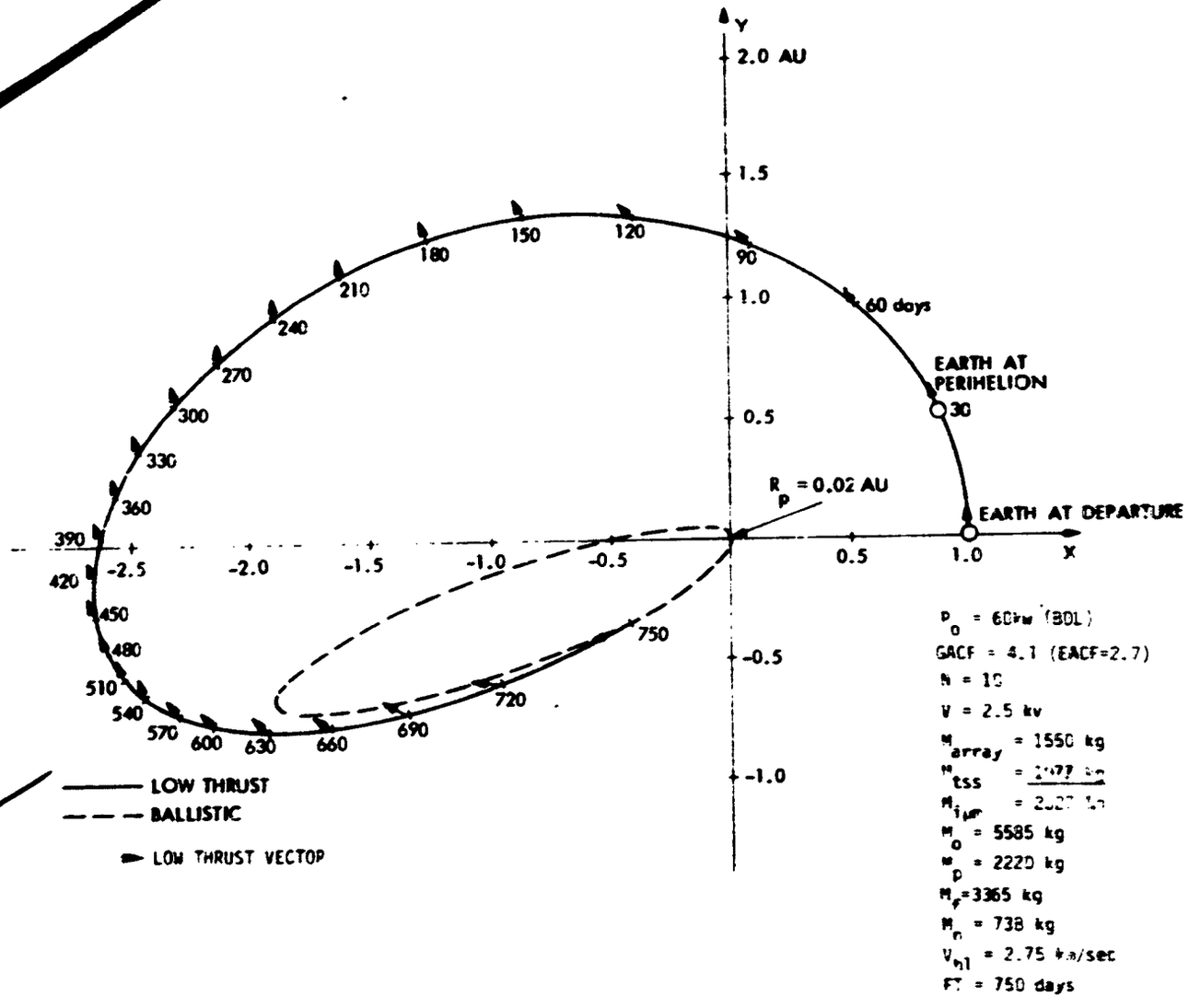


Figure 1. Ion drive Solar Probe Type-A trajectory.

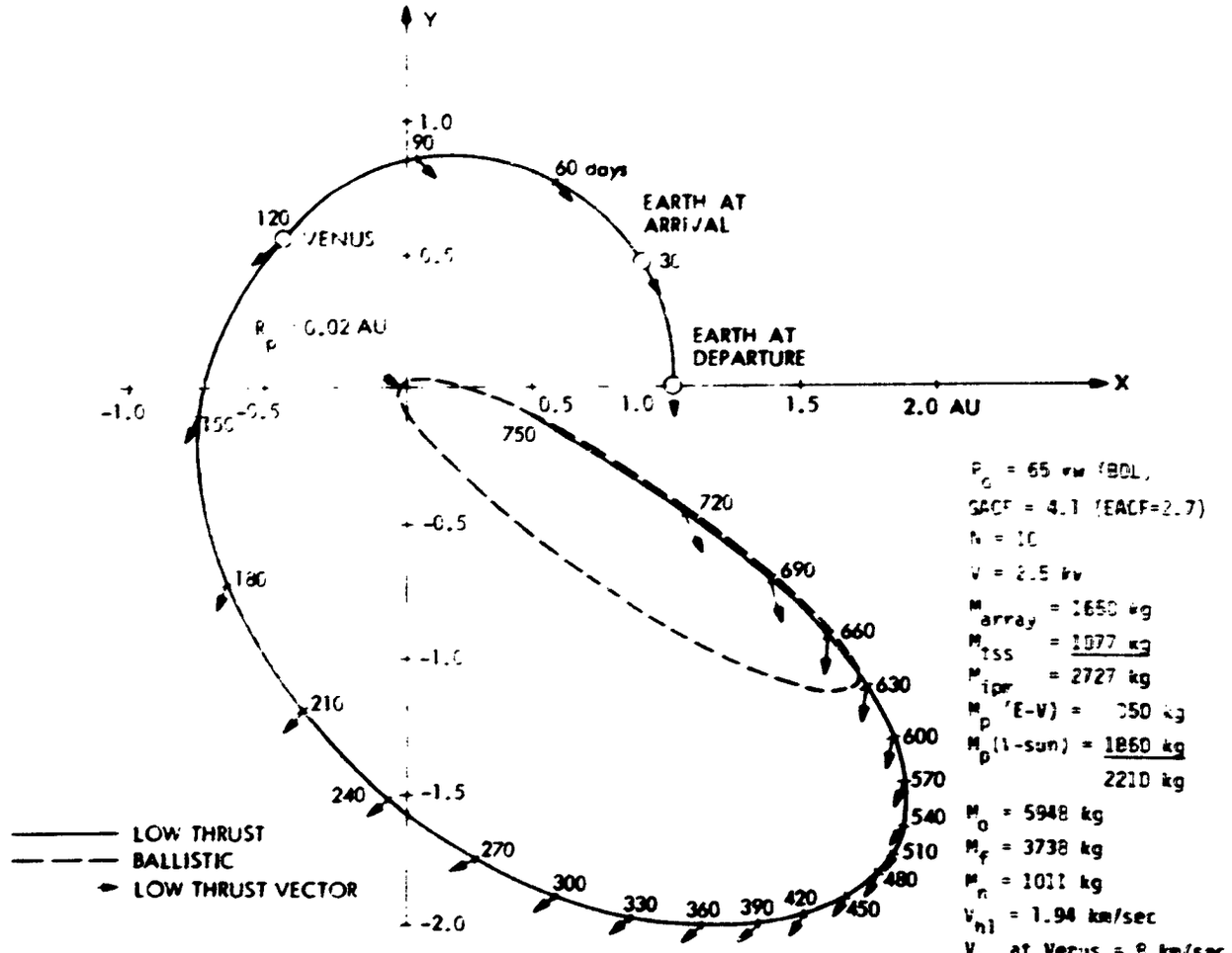


Figure 2. Ion drive Solar Probe Type-A trajectory with Venus swingby.

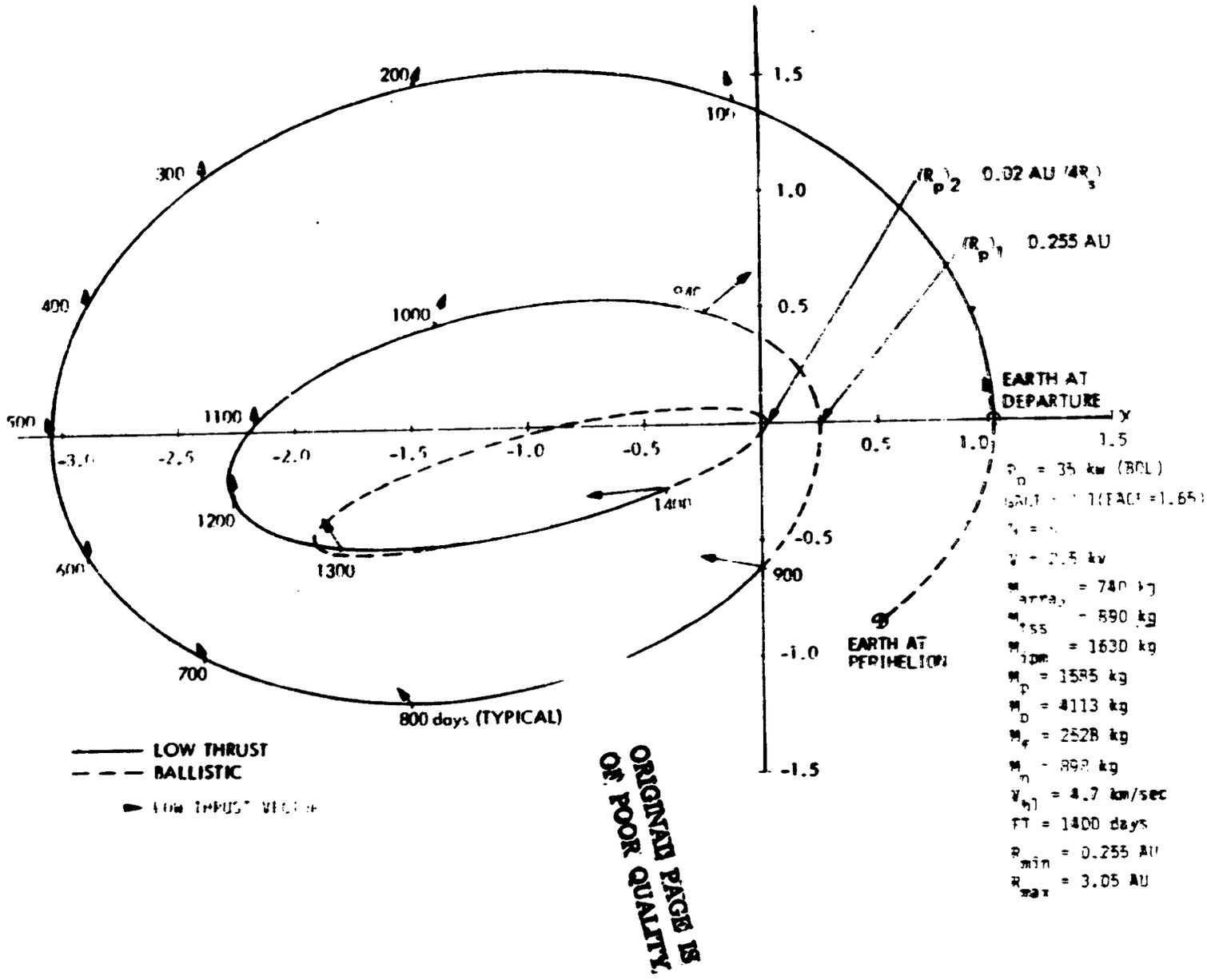


Figure 3. Ion drive Solar Probe Type-B trajectory.

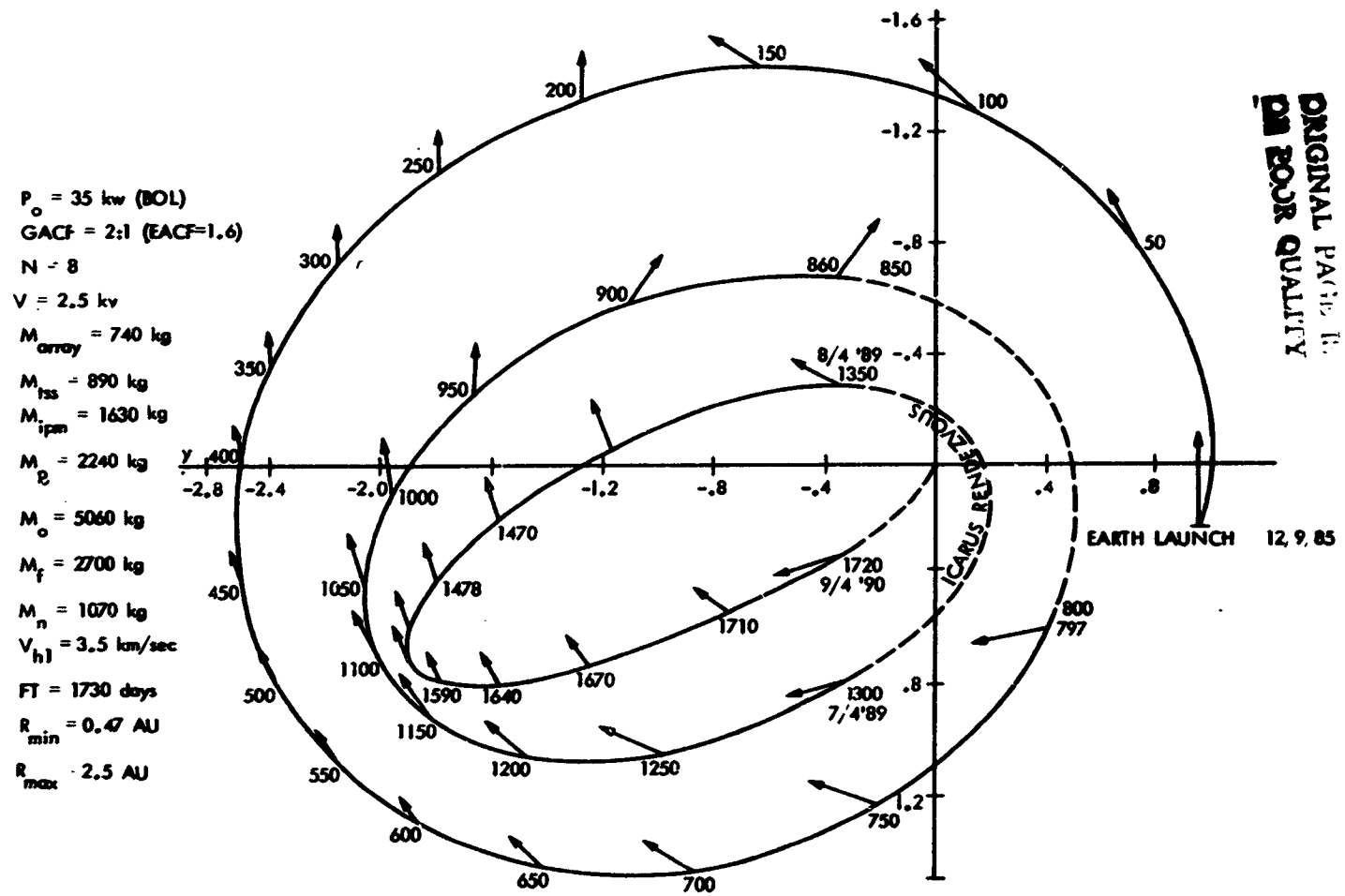
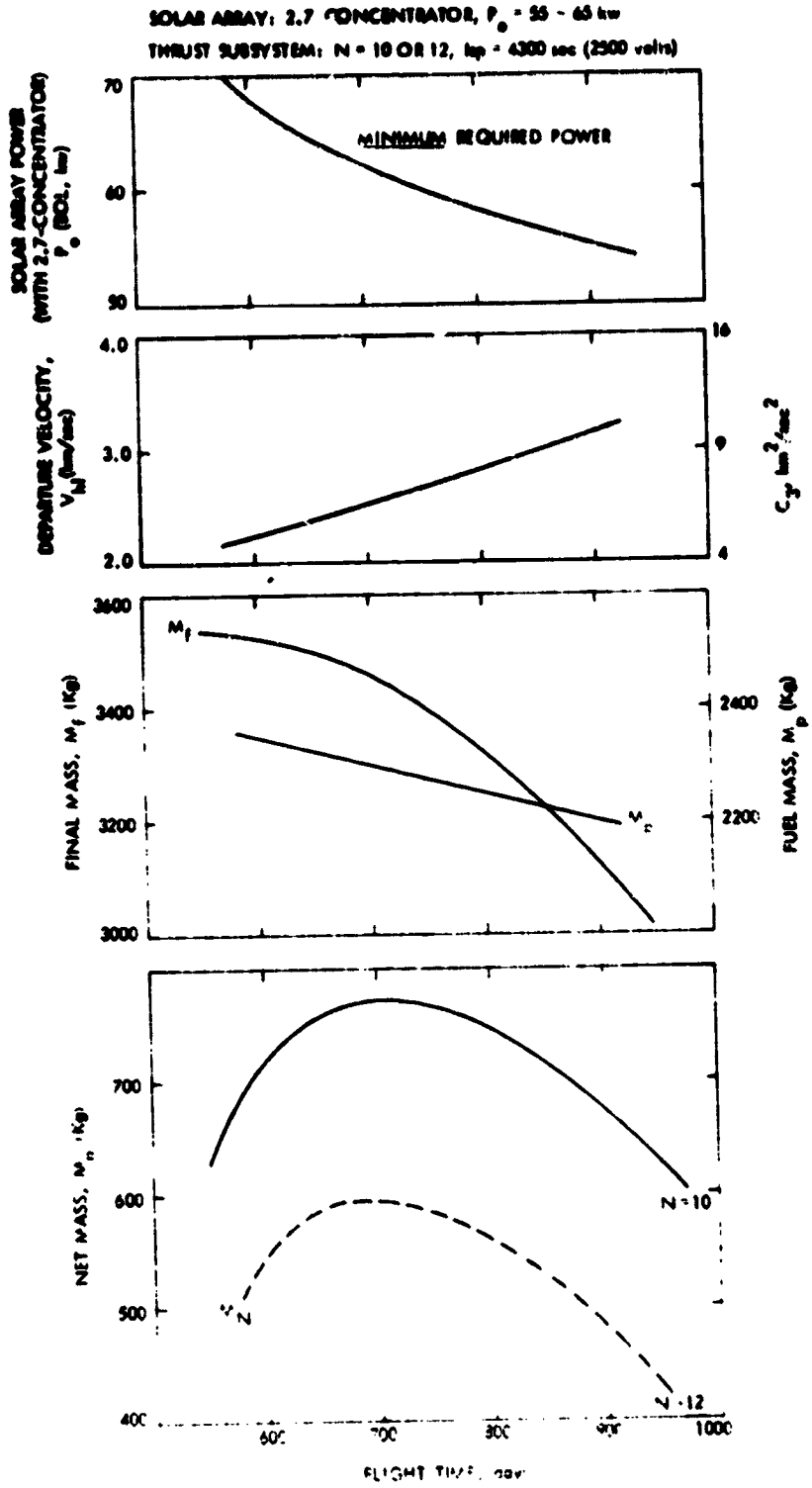


Figure 4. Ion drive Solar Probe Type-C trajectory with an Icarus rendezvous.

C-4



ORIGINAL PAGE IS
OF POOR QUALITY

Figure 5. Performance summary of an ion drive Solar Probe Type-A trajectory.

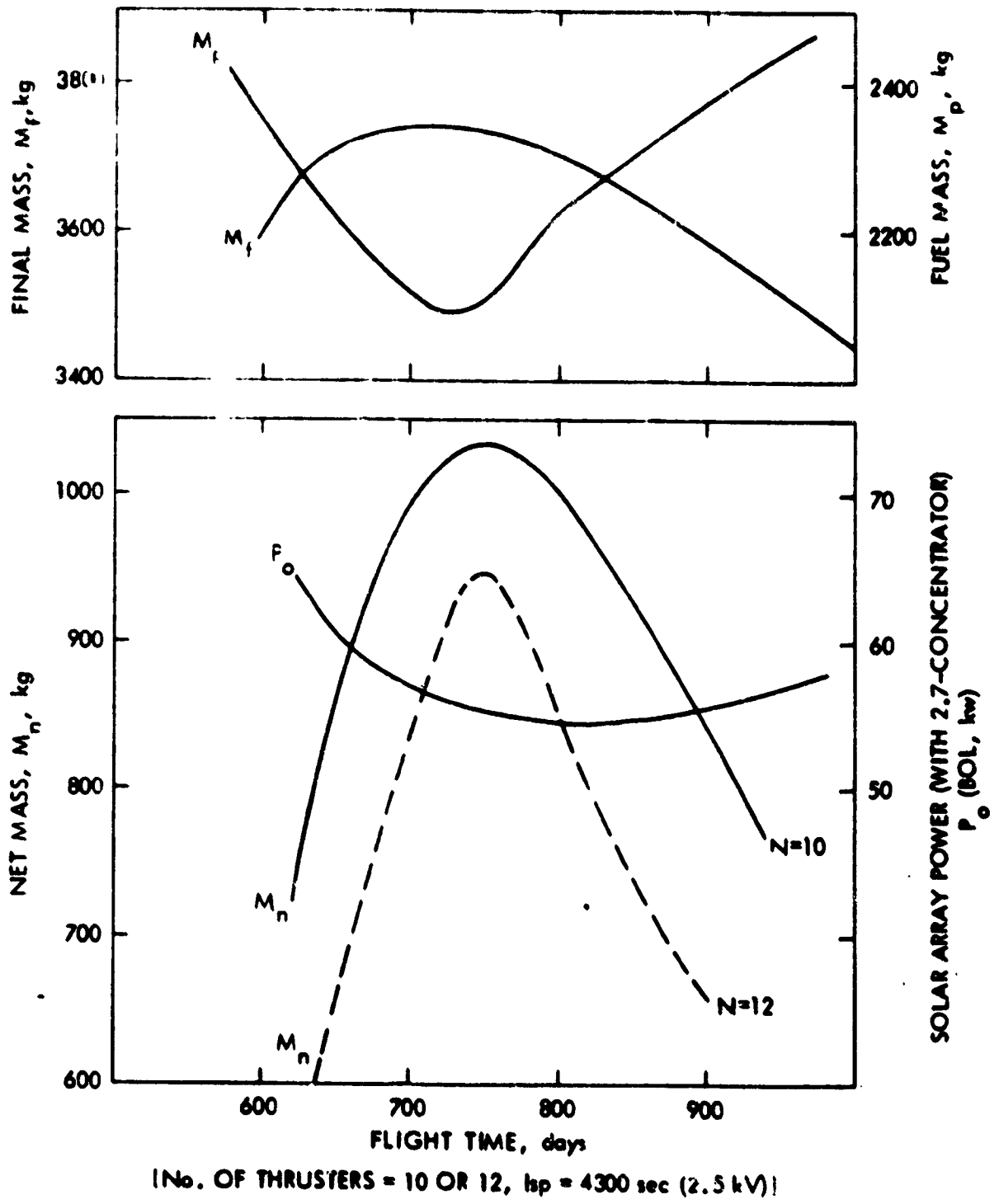


Figure 6. Performance summary of an ion drive Solar Probe Tyne A trajectory with Venus swingby.

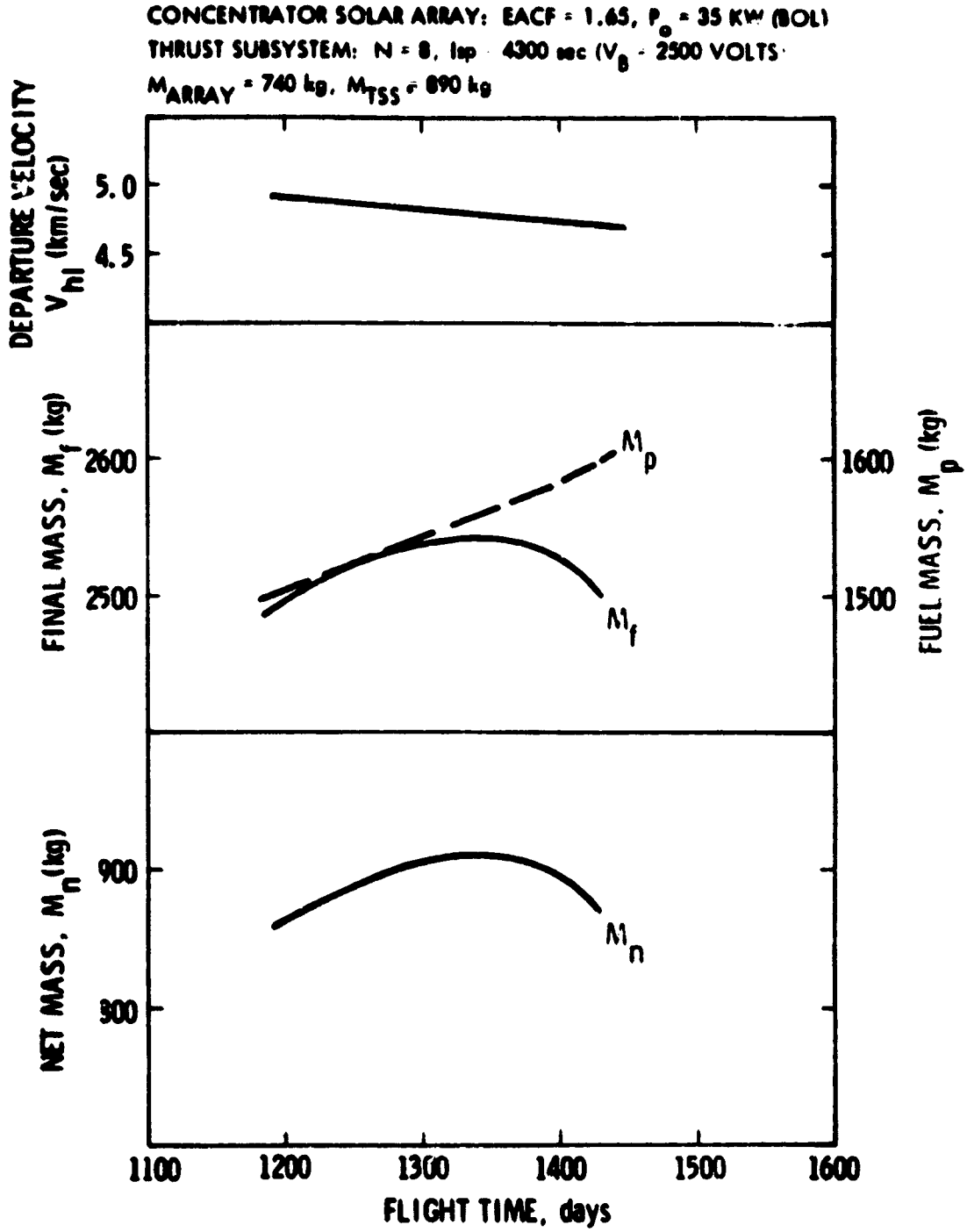


Figure 7. Performance summary of an ion drive Solar Probe Type B trajectory.

50 to 60 kw, 4-fold concentrator array and 10 to 12 thrusters must be used. The extrapolated maximum payload is about 700 kg, but an improvement can be made with a Venus swingby. Even then, a payload of not more than 1000 kg is predicted. The type-A option appears unattractive in view of the expensive propulsion system associated with it.

(2) For a 3.5 to 4 year mission, the type-B trajectory is the mode to fly. The required SEP is moderate in size; a 30 to 50 kw, 2-fold concentrator array and an 8-thruster system is adequate. (A 2-fold concentrator array is a much more state-of-art array than a 4-fold concentrator array.) A payload delivery of up to 1200 kg is feasible and as much as 1400 kg may be possible. The thermal protection of the SEP is a requirement inherent in this type of trajectory and awaits solution.

(3) If one were willing to add an additional year and consider a 4.5 to 5 year mission, a type-C trajectory with an Icarus rendezvous is an attractive option. It requires a moderate sized SEP, presents little SEP thermal-protection problem, and offers an opportunity to have a rendezvous with an Apollo asteroid.

D43

N78-33007

EUROPEAN SPACE AGENCY STUDIES OF THE SOLAR PROBE

Ian W. Roxburgh

Queen Mary College
University of London

Mile End Road
London E1 4NS England

The feasibility and scientific objectives of a solar probe were studied by a Mission Definition Group in 1975 and 1976. The group consisted of A. Balogh, B. Bertotti, G. Columbo, G. Duchossois, G. Haskell, H. Rosenbauer, E. Roth, I. Roxburgh, T. Sanderson and D. Young. The orbit analysis programme was developed by M. Hechler and an extended study of the orbit analysis was done in 1977 by Bertotti, Hechler and Roxburgh. The results of these studies are in the Report of the Mission Definition Study (1976) and an E.S.O.C. report (1978), and the reader is referred to these sources for greater details.

Mission Concept and Objectives The Solar Probe Mission would send a satellite close to the Sun to make in situ measurements of the solar wind plasma, to look at the Sun, to determine the solar quadropole and to test theories of gravity. In the E.S.A. study, no solar surface experiments were considered.

Satellite Design Figure 1 gives a sketch of the E.S.A. configuration. It has two heat shields, the front one of diameter 3 m, the secondary of diameter 2.5 m, located 0.5 m behind the first. The probe and external equipment lies in the shadow cone of the shield. The front shield gets very hot

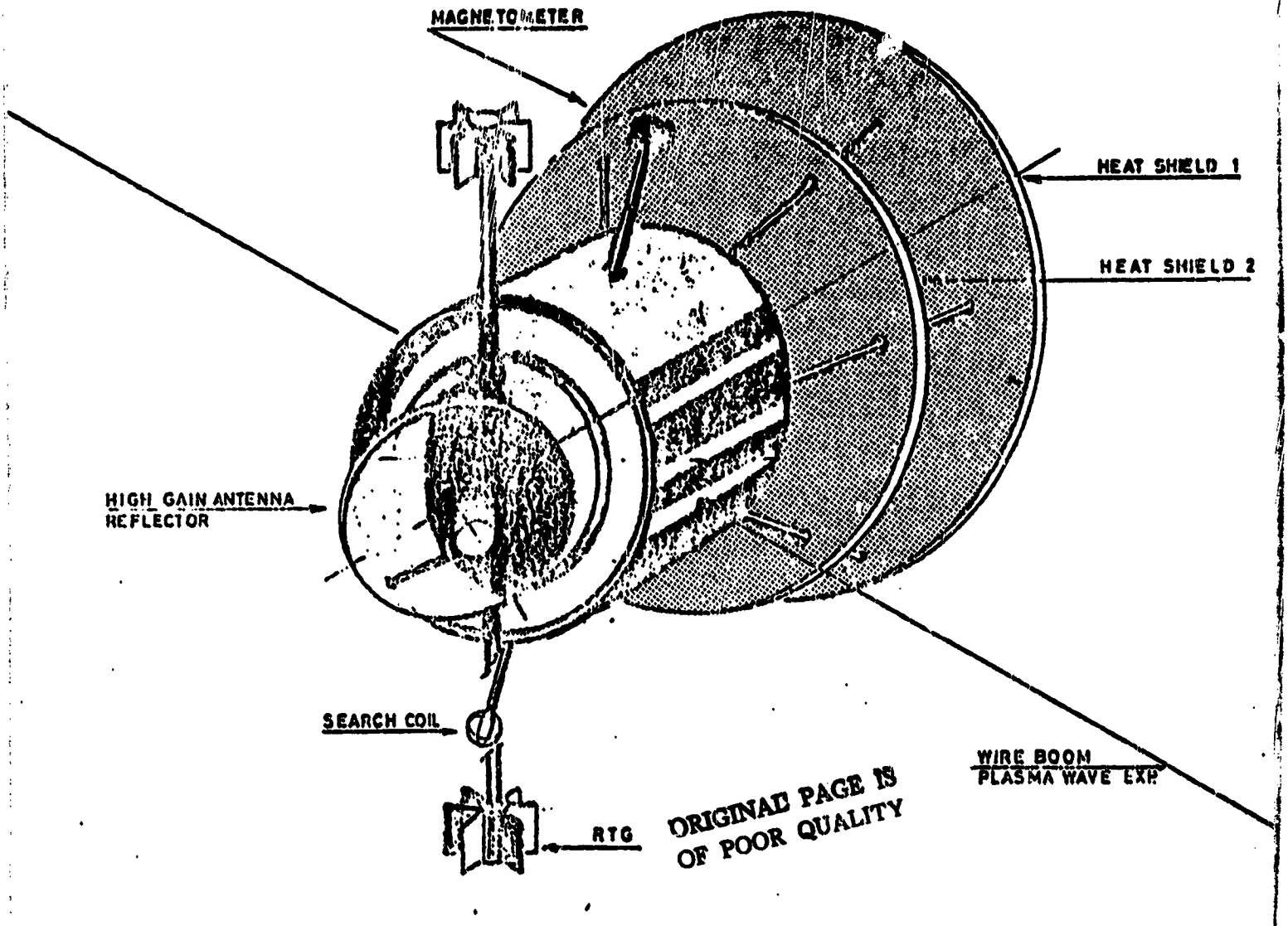


Figure 1. The E.S.A. solar probe concept.

(3000 K) and pyrolitic graphite was considered a suitable material for the shield. Figure 2 gives the steady-state temperature of different parts of the satellite as a function of heliocentric distance.

Both spin stabilised and 3 axis stabilised spacecraft were considered and no preference emerged from the technical viewpoint. However, the in situ experiments preferred spin-stabilised and the orbit experiments preferred 3-axis stabilised for the inclusion of a drag free system, although the possibility of using a CACTUS type microaccelerometer on a slowly spinning craft was considered viable.

Orbit

A variety of possible orbits were considered with C_3 energy parameters ranging from 98 km^2/sec^2 for in ecliptic oribits to 125 km^2/sec^2 for 90° out of ecliptic orbits. For the purposes of the ESA study, a launch date of 1st January 1982 was used. In the orbit analysis, a wide variery of orbits was studied with perihelion distances of 0, 4, and 10 solar radii and inclinations of 0, 45° and 90°. A typical orbit is given in Figure 3.

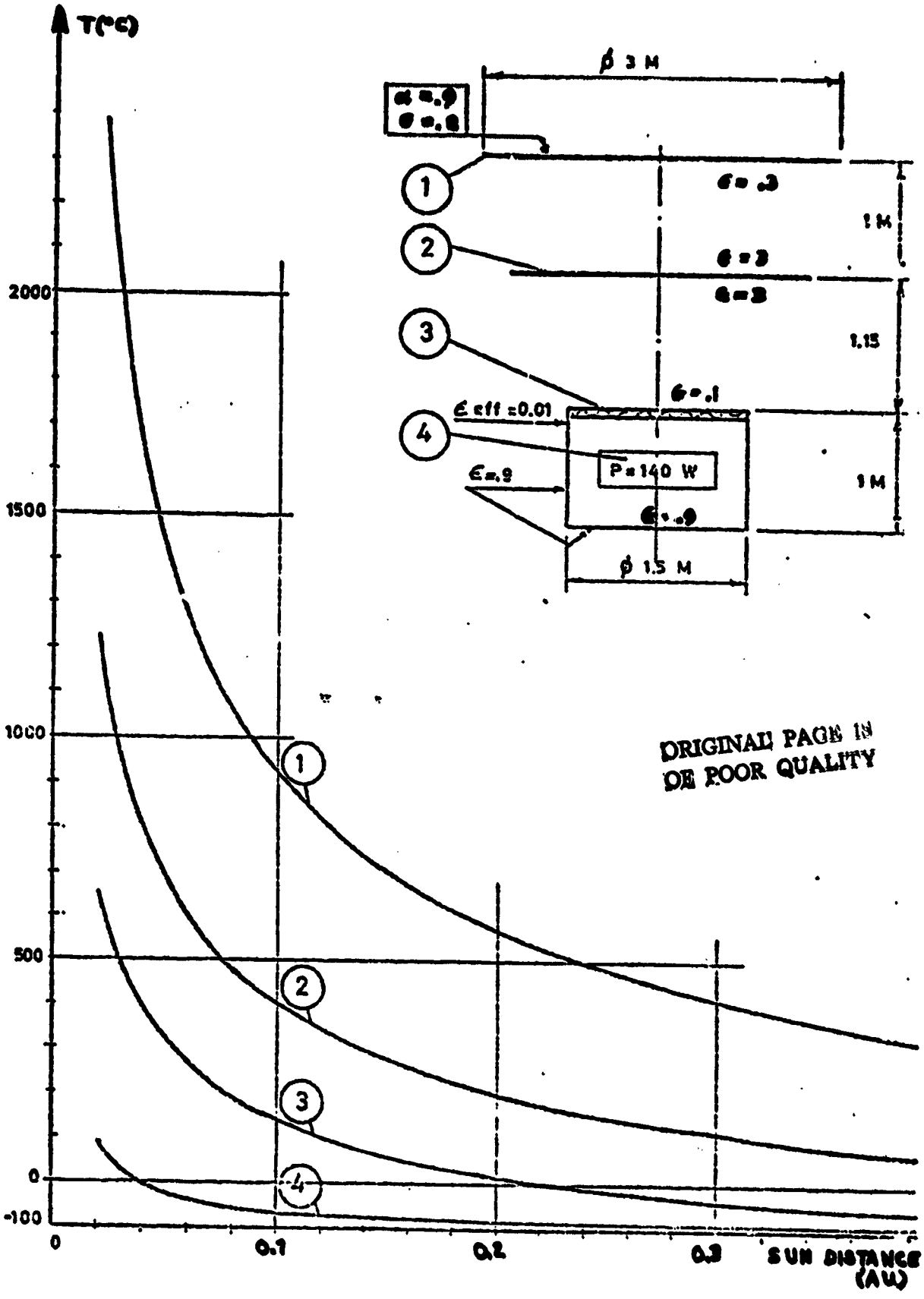


Figure 2. Variation of temperature with heliocentric distance.

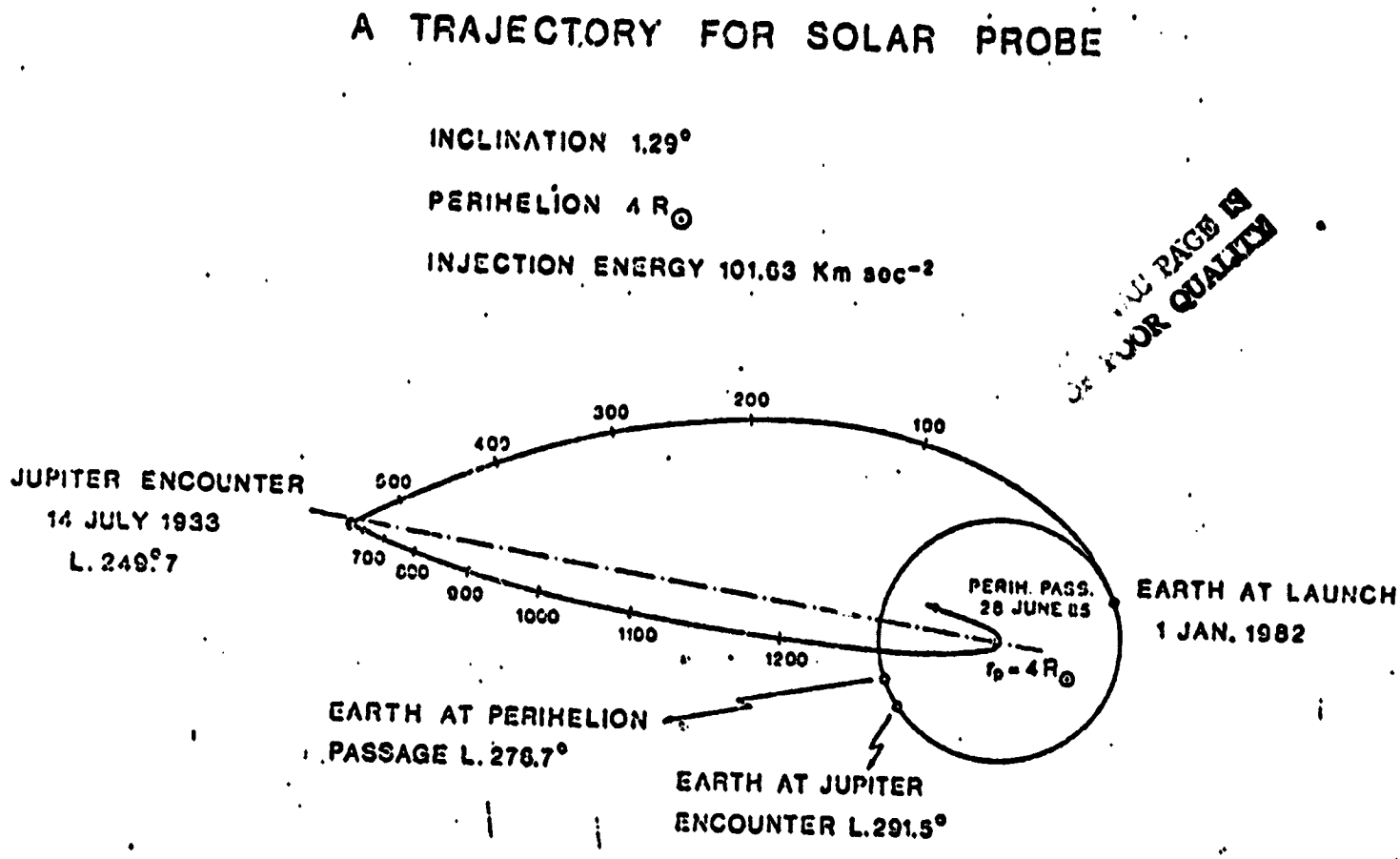


Figure 3. A trajectory for the solar probe.

Model payload for In Situ Experiments The following experiments were considered for a normal payload

a) Plasma Particles: Proton and electron velocity distributions
Protons: 20 eV to 20 keV
Electrons: 1 ev to 2 keV
Electrostatic analysers with open secondary electron multipliers or electrometers.

b) D.C. Magnetic Field: Magnetic fields 10^{-6} to 1 Gauss
2 instruments 10^{-6} to 10^{-2}, 10^{-2} to 1 Gauss
Triaxial flux gate magnetometers.

c) Plasma Waves: Frequency range 10 to 10^{7} Hz
Electric 2 axis dipole
Magnetic 3 axis search coil.

d) Ion Composition: Mass Spectrometers
Mass per charge range 1.5 to 6 AMU
Mass per charge resolution 15
Energy per charge range 0.3 to 10 keV
Energy per charge resolution 0.1

e) Energetic charged particles: Proton range 3 MeV to 30 MeV
Electron range 50 keV to 3 MeV
Nucleon range 300 keV/N to 30 MeV/N
Geiger Muller Tubes
Solid state detectors or photo-multiplier tubes

f) Coronal light detector: Photomultiplier, channel multiplier or silicon photo diode
Visible light

Orbit analysis The aim of the orbit analysis experiments is to determine the solar quadropole moment J_2, and the post-Newtonian relativistic parameters β and γ, and higher order if the accuracy can be obtained. For these experiments it was assumed that the probe will contain either a drag free device to compensate for nongravitational forces, or an on-board accelerometer to measure them. For an orbit going down to a few solar radii perihelion, it was found necessary to develop a third order $(v/c)^3$ error analysis code and this was done by M. Hechler at E.S.O.C. A large number of runs have been performed for several orbits, measurement parameters and system noise. A determination of range and range rate proportional to r^{-2} (to take account of coronal effects) was included and the system noise was taken to be isotropic and equal to

$$\sqrt{\langle (\delta a)^2 \rangle} = 10^{-8} \left(\frac{1 \text{ A.U.}}{r}\right)^2 \text{ cm sec}^{-2}.$$

Range and range rate errors were assessed for a system using an S band uplink and an S and X band downlink giving accuracies of 21 m and 3 cm/sec for a distance to the Sun of 4 solar radii. Some results of the orbit analysis are given in Table 1.

Mass and Power A preliminary mass budget gave a total mass of 425 kg and a power requirement for the experiments of 40 to 50 w.

Conclusions

The E.S.A. study showed that the solar probe is an exciting mission which would make a significant contribution to several branches of science. Further studies are needed on thermal control, drag free/accelerometers, and telecommunications, but this should be one of the most exciting missions to be flown.