The Astrophysical Journal, 711:1044–1050, 2010 March 10 doi:10.1088/0004-637X/711/2/1044

DEMONSTRATIONS THAT THE SOLAR WIND IS NOT ACCELERATED BY WAVES OR TURBULENCE

D. Aaron Roberts
NASA Goddard Space Flight Center, Heliophysics Science Division, Greenbelt, MD 20771, USA
Received 2009 July 14; accepted 2010 January 21; published 2010 February 22

ABSTRACT

The present work uses observations and theoretical considerations to provide both qualitative and quantitative arguments that hydromagnetic waves, whether turbulent or not, cannot produce the acceleration of the fast solar wind and the related heating of the open solar corona. Waves do exist, and can play a role in the differential heating and acceleration of minor ions, but their amplitudes are not sufficient to power the wind, as demonstrated by extrapolation of magnetic spectra from Helios and *Ulysses* observations. Dissipation mechanisms invoked to circumvent this conclusion cannot be effective for a variety of reasons. In particular, turbulence does not play a strong role in the corona as shown both by observations of coronal striations and other features, and by theoretical considerations of line tying to a nonturbulent photosphere, nonlocality of interactions, and the nature of the kinetic dissipation. We consider possible "ways out" of the arguments presented, and suggest that in the absence of wave or turbulent heating and acceleration, the chromosphere and transition region become the natural source, if yet unproven, of open coronal energization through the production of nonthermal particle distributions.

Key words: magnetic fields – solar wind – Sun: corona

1. INTRODUCTION

Two fundamental problems in the physics of the solar wind and corona have persisted for more than 50 years (Marsch 2006; McComas et al. 2007). Why is the open corona orders of magnitude hotter than the photosphere that it comes from? How does the corona give rise to the high-speed streams that originate from coronal holes? Parker's original model of the solar wind (Parker 1958) showed that a wind could be generated by the thermal pressure of the hot corona, but much subsequent effort revealed that high-speed streams required, at least in the context of conventional fluid models, some additional deposition of energy and momentum in the extended corona, above the sonic critical point where the flow and sound speeds were equal (Leer & Holzer 1980). A natural way of transporting energy from the base of the corona outward is to presume that waves are generated low down, in the photosphere or chromosphere, and that these waves carry their flux to be dissipated higher up. The waves act in two ways: through heating by dissipation, and by a ponderomotive pressure proportional to the square of the averaged wave amplitude. Overall, conservation of energy demands that the wave flux be sufficient to supply the kinetic energy flux of the flow and the gravitational potential energy gained as the wind climbs out of the gravitational well of the Sun. The belief that waves produce fast solar wind streams is highly popular, despite the lack of a complete, consistent theory that demonstrates its validity.

The discovery of Alfvén waves in the solar wind (Unti & Neugebauer 1968; Belcher & Davis 1971) gave the above wave picture further credibility, and many authors produced solar wind models based on this (see Hollweg 2008 for a recent review). It was very quickly realized, however, that the observed flux of waves in the solar wind, extrapolated back to the Sun using the short wavelength, WKB, approximation (see discussions by Whang 1973 and Roberts 1989), produced a wave energy flux near the Sun that was roughly 15% of that required. This meant that dissipation was required not simply for heating, but also to resolve this inconsistency: the solar wind wave flux had to be a weakened remnant of the original (or perhaps completely unrelated to the coronal flux) for the scenario to work. The problem was made more vexing by the difficulty in damping linear Alfvén waves (Barnes 1966). The strong damping of the acoustic mode made it a poor candidate for coronal heating, and the fast mode seemed likely to refract out of the fast wind regions (see, however, Barnes et al. 1971, whose model may not be completely ruled out if a source can be found within the streams). This left Alfvén waves (the only other linear mode) as both the natural candidate and what was observed in the solar wind. However, without a damping mechanism, this promising avenue was not viable. The focus turned to various possible nonlinear dissipation mechanisms, usually augmented by kinetic dissipation at small scales. While none of these approaches has claimed to develop a fully self-consistent model, the wave-acceleration idea still appeals due to its qualitative explanation of various observations, as discussed in the reviews cited above.

This paper presents many arguments that, both individually and together, argue strongly against wave and turbulence models of fast solar wind acceleration and heating. The basic conclusions are that WKB estimates are not far from the correct extrapolation of the wave power to the Sun; that the corona does not strongly exhibit signs of turbulence; and that current wave and turbulence models of the acceleration of high-speed streams make a number of untenable assumptions or approximations. (Note that the quantitative arguments apply directly only to the problem of acceleration.) The following section deals with constraints based on measurements by spacecraft in the solar wind and coronal imaging. These constraints are very stringent, and no model has fully taken these into account. Section 3 treats arguments concerning modeling issues such as the dissipation and reflection of fluctuations, and discusses some "ways out" of the presented objections. Section 4 briefly discusses alternative scenarios, and the final section poses a set of questions that must be answered by any solar wind model, based on this work. In the end we suggest, as proposed by the authors discussed in Section 4 and others, that the driver of the solar wind and corona is most likely to be where the energy and dynamic processes are: in the chromosphere and transition

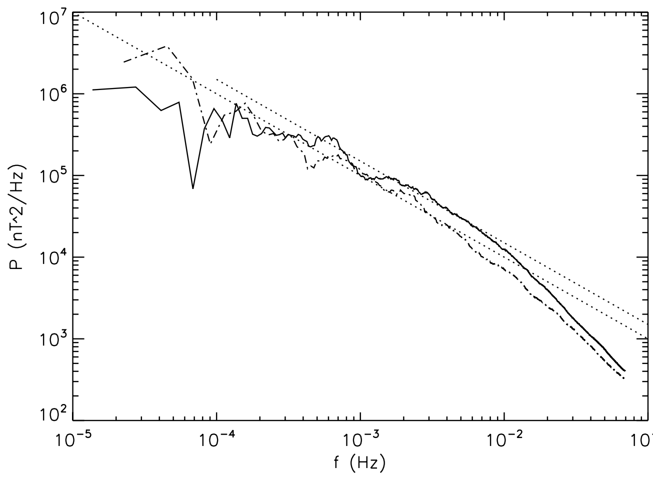

Figure 1. Sum of the component spectra for the magnetic field from Helios spacecraft for a slow and a fast wind interval. The solid line is for the high-speed stream observed by *Helios 2* on day 106 of 1976, and the dot-dashed line is for the slow wind observed by *Helios 1* on day 124 of 1978. The dotted lines are for an ideal f^{-1} power law.

region, fueled by photospheric and chromospheric motions and the magnetic complexity of the lower solar atmosphere that lead to nonthermal particle distributions.

2. OBSERVATIONAL EVIDENCE FROM SOLAR WIND SPECTRA AND CORONAL IMAGES

Consider first the magnetic field spectra shown in Figure 1. Both lines represent logarithmically smoothed (in ±20% width bins) spectra of 6 s data points from a day of data at 0.3 AU from *Helios 2* (solid) and *Helios 1* (dashed) spacecraft. The former spectrum is from day 106 of 1976 and the latter from day 124 of 1978. Both intervals are highly Alfvénic with correlations between the magnetic and velocity fluctuations ∼0.9 and a very nearly constant (and nearly equal) magnetic field magnitudes. The very high Alfvénicity and outward sense of propagation indicates the origin of the fluctuations is almost certainly below the Alfvénic critical point where the flow and Alfvén speeds are equal. The primary difference between the intervals is that the *Helios 2* spectrum is from a well-studied fast wind interval with a roughly constant speed of 700 km s^{-1}, whereas the other wind sample has a roughly constant speed of 350 km s^{-1}. The fast wind has a kinetic energy flux $\propto \rho V_{sw}^3$ about 4 times that of the slow wind (half the density and twice the speed). Thus, these spectra pose the fundamental problem of how waves of nearly identical amplitude can, in one case, produce a high-speed wind, and in another produce no acceleration at all (presuming the slow wind to be thermally driven). To the extent that the two spectra differ, it is in the sense that the slow wind spectrum is slightly lower at high frequencies, consistent with greater dissipation, but therefore directly in contrast to what is required in wave-acceleration models. Although it has been known for a long time (Roberts et al. 1987; Marsch et al. 1981), few appreciate that large amplitude Alfvén waves are not unique to high-speed flows. More generally, the primary controlling factor for the amplitude of waves in the solar wind is the magnitude of the field, with $|\delta\mathbf{B}|/B$ essentially independent of any bulk parameter such as speed or density (Roberts et al. 1990).

A second feature of these spectra is that they are very flat. The typical magnetic spectrum observed below the “turbulence” or $f^{-5/3}$ regime has a slope of ~ -1, as seen here. In this case, this slope extends from very low frequencies up to 10^{-2} Hz (100 s; see Bavassano et al. 1982 for similar high-speed wind observations). There are shorter intervals in the Helios data, most clearly in the high-speed wind, in which the spectrum does not steepen until 0.05 Hz (20 s). This shallow slope presents a second fundamental difficulty for wave acceleration theories: dissipation is required of the waves, but no known dissipation mechanism will produce such a flat spectrum. Thus, the ∼100–500 s waves found in Hinode observations (De Pontieu et al. 2007) would likely be damped very little in the region where dissipation is needed for heating and acceleration. Note also that ∼20 km s^{-1} amplitude of the Hinode waves is consistent with the extrapolations of Roberts (1989) and others, and is not adequate to accelerate the wind, contrary to the conclusions of De Pontieu et al. (2007) who also provide energy flux estimates that are insufficient by a factor of 5 or so. For example, in a recent paper that simulates the effects of nonlinear, non-WKB, waves of any magnetohydrodynamics (MHD) mode, the fluctuation amplitudes required are as high as 75 km s^{-1} for the fastest streams (Ong et al. 1997). The region from 10^{-3}–10^{-2} Hz is that invoked in recent turbulence models (e.g., Dmitruk et al. 2002; Cranmer & van Ballegooijen 2005), and a turbulent cascade in this frequency range would not result in the observed flat spectrum. (See also the direct simulation of the spectral steepening by Ofman & Viñas 2007.) Below we discuss further the issue of whether “quasi-two-dimensional” (Q2D) turbulence avoids this problem. It is important to note that the observational constraint on the inadequacy of the wave flux near the Sun is independent of the fluctuation velocity at the coronal base, being a direct constraint on the energy flux, and thus is quantitatively valid if the dissipation of the waves is low.

Further evidence that the waves in the f^{-1} region are not damped comes from at least three independent means of determining the radial evolution of the fluctuations. Roberts (1989) used both variances of fluctuations and filters in particular bands to show that the WKB approximation holds very well for the f^{-1} frequency region of the high-speed stream referred to in Figure 1 here. Horbury & Balogh (2001) used other means to isolate frequency bands and found, based on a large statistical sample, that for both Helios data inside 1 AU and *Ulysses* data outside 1 AU, there is a strong correlation between the WKB scaling with distance of the fluctuation amplitudes and the f^{-1} spectrum. Hollweg (1974; see also Verma & Roberts 1993) showed that the WKB scaling is a simple consequence of the dissipationless propagation of fluctuations that have an Alfvénic equipartition of magnetic and kinetic energy; they do not even have to be waves, although they almost certainly are given the Alfvénic correlations and the “surfing” of the Helium at the Alfvén speed ahead of the protons (Goldstein et al. 1995, and references therein). The required equipartition is more prevalent in spacecraft data closer to the Sun, and in particular in the same high-speed wind as studied here at 0.3 AU it holds to within ∼10%. Thus, there is very strong evidence that the fluctuations leading to Figure 1 have not been dissipated except at the smallest scales where the spectra eventually show evidence of turbulence. Roberts (1989) has shown that the dissipation possibly implicit in the higher frequency evolution leading to the break in Figure 1 will not change the conclusion that the waves are not strong enough to accelerate the wind, even when ignoring gravity and including a spectrum of waves that covers an unphysically large range in frequency. We have a good understanding of how turbulence appears and acts in the region

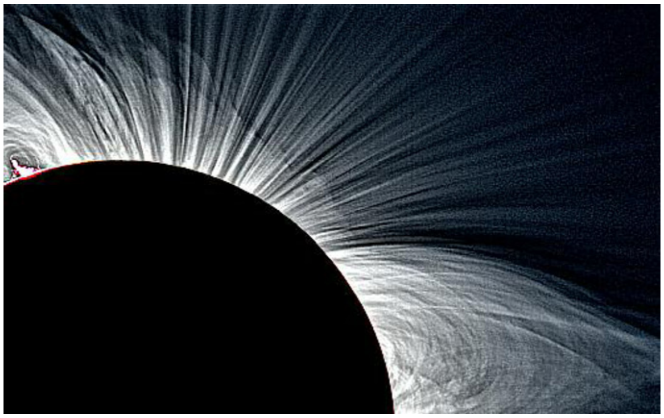

Figure 2. Edge-enhanced subset of a coronal image from the 2006 eclipse of the Sun observed in Libya. See, e.g., Wang et al. (2007). Image obtained from http://www.zam.fme.vutbr.cz/~druck/Eclipse/Ecl2006l/0-info.htm and processed by the present author using Photoshop.

where spacecraft have measured it: the spectrum steepens to typical turbulence slopes starting at high frequencies and working to lower as sufficient time passes for mixing at the larger scales. Note that at lower *f*, below the nominally f^{-1} region, the radial evolution is even slower, approaching the "static limit" in which the fluctuations decay in the same way as the transverse Parker field ($\propto 1/r$). Extrapolation using this radial evolution makes the wave power at the Sun yet lower by about an order of magnitude.

The WKB scaling of the wave amplitudes in the energy containing range of the fluctuations sets the scale for extrapolation of the wave amplitudes to below the Alfvénic critical point at $\sim 15 R_S$. The fluctuation level, $|\delta \mathbf{B}|/B$, decreases with decreasing distance from the Sun, so nonlinear effects, not apparent in the f^{-1} region at 0.3 AU, should be even less likely closer in. Thus, the WKB approximation should provide a very robust means for extrapolating spectral levels. This paper does not deal with the actual origin of the observed interplanetary fluctuations, but the considerations presented here point in the direction of a source above the transition region but below the Alfvénic critical point, most likely involving the differences in the wind speed as a function of the direction transverse to the radial as the dominant free energy in the fluctuations.

Visual inspection of high-resolution, ground-based, and other coronal images does not reveal the patterns expected of a turbulent medium. Figure 2 is a portion of a composite eclipse photograph that shows very straight "rays" or "striations" for implied open magnetic fields. The transverse scale of possible coronal turbulence is generally assumed (with weak justification) to be 30 Mm or so at the coronal base, and to stretch with the expansion of the flow. In the figure, this corresponds to the distance across a few of the striations. One might expect to observe some activity at that scale, but, apart from a few crossings that may be line-of-sight effects, the rays remain parallel and look very much like the flow lines in a laminar fluid. There is no two-dimensional mixing that would lead, at larger scales, to a braiding of the striations and at smaller scales to a diffusion of the striations with both time and space. Neither of these effects are observed, and movies show that the structures mostly just rotate with the Sun. Rays in the coronal images are straight past $15 R_\odot$, where typical predicted amplitudes (without the dissipation we have ruled out above; see, e.g., Cranmer & van Ballegooijen 2005) imply $\sim 30°$ bends in the field. The best direct observations we have of the motion of the solar wind due to waves is that given by white-light images of comet ion tails, which we take to be tracers of the wind motion (see Vourlidas et al. 2007, who are primarily interested in the tail disconnection). These images show deflections of less than 15° in the inner heliosphere, consistent with Helios observations and precisely as predicted by the model of Whang (1973) based on the WKB extrapolation. This is, again, consistent with a wave flux that can supply at most 25% (due to half the amplitude) of the energy flux needed. Although the cited comet tail case may involve a relatively slow wind, there is no general correlation between fluctuation amplitude and wind speed (Roberts et al. 1990).

Figure 3. Transition to turbulence in a rising hot air plume. From the Iowa Institute of Hydraulic Research (Rouse 1950). While this will not be exactly the expected behavior in the solar wind, the evidence presented in the text implies that turbulence in the solar wind is similar in becoming important at some distance from the source region.

We also know what happens in ordinary fluid flow when it becomes turbulent. A commonly observed case involves a laminar flow that becomes chaotic when the streamwise Reynolds number becomes sufficiently high. The chaotic state is dramatically and suddenly different from the laminar state. Figure 3 (from Rouse 1950) shows an example of such a transition to turbulence for a rising plume of hot air traced by smoke. The observed coronal striations do not look like the turbulent fluid case, and while we would not expect identical behavior for an ordinary fluid and an MHD fluid with a strong mean field, there should be some evidence of the turbulent state. The most likely observational consequence of a truly turbulent corona would be the complete elimination of the striations within a few solar radii. If the turbulence were confined to the region of the striations, it should cause a diffusion that would smear out the linear features, but many of the striations are very fine and well-columnated for long distances.

It is natural to ask where the striations are in three dimensions. Wang et al. (2007) find evidence that coronal rays come from interchange reconnection regions over bipoles in coronal holes and near the edges of the holes. Hole edges may be regions of strong shear, where the wind speed is either changing from fast to slow, or they may be the origin of highly structured slow winds that are seen in spacecraft observations to be regions of more developed turbulence. If the rays come directly from sites of reconnection and/or large shear, then they are more likely to be in regions of strong turbulent effects than if they were in the center of a relatively uniform hole flow. Thus, we see what appears to be laminar flows where they are least expected if the corona is a region of significant turbulence.

There are many other ways to examine coronal images for evidence of motions at or near the foot points of coronal features. For example, any of the LASCO EUV Imaging Telescope (EIT)/195 movies (http://lasco-www.nrl.navy.mil/daily_mpg/, e.g., 2009 March 2) show bright points in coronal holes and elsewhere that clearly rotate with the Sun, with a maximum rotational surface speed of 2 km s^{-1} at the equator, and thus the motions about the mean motion are much slower than this. Although these points do not correspond to foot points of open field lines, their motion is probably typical of other regions of the corona, and certainly not 50 or more times slower as would be required by many wave acceleration theories. These speeds are below those implied by WKB extrapolations of interplanetary wave power back to the Sun, and this would be consistent with a source for the interplanetary fluctuations above the coronal base but below the Alfvénic critical point. Between two EIT frames, taken 12 minutes apart, there is a clear rotational motion of nearly all features with little change in any. The typical transverse correlation scale assumed in turbulence models is more than 10 times as large as the displacement of the points between frames. There is no sense of much evolution either along or transverse to the features in these movies that has anything to do with the assumed correlation scale. While it may be possible that MHD turbulence is so different from fluid turbulence that the above observations are not important, this has not been demonstrated for any theory.

3. INADEQUACY OF TURBULENT AND OTHER FLUCTUATION-BASED MODELS

A common scenario involving a turbulent cascade of fluctuation energy for wind acceleration and heating involves the feeding of the small scales by the cascade, explicitly or implicitly with a significant cascade with wave vectors along the mean magnetic field direction, followed by perpendicular heating of the ions due to cyclotron resonance (see the review of Hollweg & Isenberg 2002). This model requires waves with ∼kHz frequencies near the Sun. This is 5 or 6 orders of magnitude removed from the typical "stirring scales" in the models. Observed solar wind turbulence shows a removal of existing energy in the spectrum as the primary source of turbulent heating, not a cascade from a steady, large-scale driver. The spectrum in Figure 1 retains the WKB evolution at the large ("f^{-1}") scales, while the smaller scales relax to a turbulent spectrum that declines in level as the spectral break point moves to lower frequencies. The cascade rate through the inertial range is likely governed by a generalized Kolmogoroff phenomenology, but the energy source is the non-equilibrium spectrum at wave numbers in what becomes the inertial range rather than a continuous feeding from a fixed scale below.

The inadequacy of the WKB extrapolated fluctuation power thus implies that any high-frequency source of fluctuations near the Sun cannot be provided by a parallel cascade from low frequencies, but must be provided by some other source of high-frequency fluctuations. Such a source must contain at least the 500 W m^{-2} minimally required for acceleration, and probably much more, since the waves will be only one energy release channel; if microflares are taken as the source of the waves, these are much more likely to release energy in jets (as in Hinode observations) or direct heating. Axford et al. (1999, and references therein) suggested that nanoflares could provide such a source (without turbulence) through small-scale reconnection in the magnetic carpet near the Sun. It has been argued that an unobserved level of compressive fluctuations will be generated by this process unless the waves generated are purely parallel propagating, which seems unlikely (Hollweg & Isenberg 2002). In addition, the timescale of waves should be given by the timescale of the reconnection processes, which would imply a very large number of millisecond reconnections, with a distribution of slower processes consistent with the required heating function (see the discussion of Hollweg 2008). There is no reason to believe such fast processes occur in the required manner. Moreover, reconnection as observed in magnetospheric and solar wind plasmas, as well as in closed regions in the corona, generates heating and jetting far more than waves. Such processes, at a much wider range of timescales, could be partially responsible for the production of nonthermal particle distributions that could power the wind in a different manner (see below). We also note that recent Hinode observations show that the coronal holes are not covered in bipolar regions that would be the sources of nanoflares, but, instead, strong unipolar regions are the dominant polar field sources (Tsuneta et al. 2008).

There are a number of more technical concerns about the current models for turbulent dissipation in the corona. For example, the theories for the dissipation that assume a parallel cascade do not properly treat interactions due to reflection, which will at most lead to an inefficient nonlocal (in wave number) coupling. Velli (1993) shows the reflections occur efficiently for $f < 10^{-3}$ Hz, and this is consistent with others' results. These are the only inward waves available to provide a cascade. The heating due to the cascade is usually modeled as $\propto (z^{+})^2 z^{-}/L$, where the $z+$ and z^{-} represent the amplitudes of outward and inward waves and L is the correlation length of the fluctuations. This formula assumes a local interaction in wave numbers; the nonlinear interactions produce effects at sums and differences of wave vectors, and these produce substantial changes in wave vectors only when the interactions are between fluctuations of comparable scale. Such a local cascade is implicit in Kolmogoroff and other turbulence phenomenologies, but is typically not taken into account in parallel cascade theories that require that the long-wavelength drivers interact with smaller scale fluctuations to produce a cascade.

Many turbulence models do not treat the dissipation mechanism, but simply assume that the turbulent cascade produces heating according to the formula in the previous paragraph or one like it. In some cases the model is inconsistent in that a turbulent cascade is assumed, but the spectrum itself is not evolved self-consistently. Thus, in the Cranmer & van Ballegooijen (2005) model, the spectrum at 1 AU is a direct reflection of that used as input, linearly propagated, with no change of shape due to a cascade. No power is shown to be generated at the scales where dissipation would occur, and thus there is no opportunity for perpendicular or any other kind of heating. Models that just use a formula for the cascade rate effectively introduce a parameterized heating function, but no mechanism for dissipation. This approach was used in an early model by Hollweg & Johnson (1988), and while it is possible to generate fast wind solutions in this way, it is necessary to introduce variable dissipation lengths and to ignore the constraints posed by Figure 1 and the related discussion.

Models that invoke quasi-two-dimensional turbulence (Q2D) have a variety of fundamental problems. In their favor, these models avoid the objection of nonlocality in wave-vector space by giving the driver a strong transverse component and having it drive a cascade in wave vectors largely perpendicular to the field. In fact, the only reason to have any time dependence for the boundary condition in this case is to generate, by reflections off the Alfén speed gradient, the required fluctuations with "inward propagation" (the appropriate sign of the correlation between velocity and magnetic field fluctuations) that are needed for nonlinear couplings. (Roberts et al. 1999 used a stationary boundary condition to drive a two-dimensional cascade in a three-dimensional expanding flow in which both signs of correlation were convected from the inner boundary.) The current Q2D coronal heating models also have the advantage of dissipating the fluctuation power in modes that do not propagate along the mean magnetic field, and thus their power would not be observed in the solar wind. However, these models do not self-consistently include the wind flow in the mechanism; this will convect the Q2D fluctuations outward (Roberts et al. 1999), with the magnetic field decaying more slowly than the WKB prediction (apart from dissipation) leading to potential conflicts with observations depending on the details of the model.

Typically, it is at least implicitly assumed that the parallel fluctuations are evolving fairly independently from the Q2D modes, and thus may have a different signature. The slowly decaying Q2D magnetic fluctuations should be evident in solar wind spectra unless they are all dissipated. If the latter occurs, then all the fluctuations in the solar wind are due to another process that occurs in the same region as the Q2D evolution but is independent of it despite boundary conditions that vary in time as well as space across the solar surface and that produce flows that are highly inhomogeneous (Markovskii et al. 2006). Moreover, solar wind observations support the idea that the "wave-like" (parallel) and Q2D turbulence components of the fluctuations are strongly coupled (Milano et al. 2004). Shear in the radial velocity, which is highly likely, will turn wave vectors from the parallel direction, leading to further mixing (e.g., Roberts & Ghosh 1999, and references therein). Thus, it is not plausible that the parallel component of the fluctuations could retain an "unevolved" (nonturbulent) spectrum while the perpendicular component evolves and strongly dissipates, and the flatness of the spectra in Figure 1 should constrain Q2D models as well as others.

Some of the problems with the Q2D scenario are explicitly recognized by the modelers themselves. One particularly significant point is that Q2D turbulence does not provide, at least directly, the parallel wave power that is needed for the perpendicular ion–cyclotron heating that is usually assumed to be the ultimate end of the cascade process and that would explain observed anisotropic ion temperatures. Hollweg (2008) discusses this issue at length, without providing a solution to the problem. One such attempted solution turns the perpendicular power into parallel power by a succession of interactions in a nonuniform corona (Markovskii et al. 2006); while this helps with the cyclotron heating, it revives the objections raised in the connection with Figure 1. Another attempt to address this issue (Cranmer & van Ballegooijen 2003) does not produce enough power to heat the heavy ions, much less the protons.

A fundamental issue is that of the boundary conditions in the Q2D scenarios. Turbulent cascades are more efficient perpendicular to the mean magnetic field because the two-dimensional interactions shuffle flux tubes around rather than performing the extra work needed to bend the tubes. If we consider a two-dimensional fluctuation initiated near the surface of the Sun, it will have to violate the line-tying condition on the magnetic fields in the highly conductive solar surface: any attempt to shuffle flux tubes that is not consistent with surface motions will result in a kink that will relax the original stress by emitting an Alfvén wave outward from the Sun. This problem was dealt with by Dmitruk et al. (2001), who acknowledge the requirement that "zero-frequency modes" must be possible. Their solution is to impose a condition on the derivatives of the transverse wave fields at the lower and upper boundaries with respect to the vertical coordinate, s. Thus, any solution in the box for the transverse motions that is independent of s, i.e., the zero-frequency modes, meets the boundary conditions, and thus the flow inside the box can be completely unrelated to the flow in the boundary, which can be any other function of the transverse variables. Therefore, the derivative conditions sacrifice line tying, and this is what allows the two-dimensional interactions to occur. There are simulations in which line-tied boundaries produce strong nonlinear interactions (e.g., Rappazzo et al. 2008), but in these cases the two ends of the flux tubes are driven differently, simulating closed loops. All observational solar movies showing features arising from the solar surface show line tying, and this is how many models drive (finite-frequency) waves. The line-tying condition is made more significant by simulations of the solar convective zone (Brummell et al. 1996) that show that the solar photosphere is largely laminar, masking the turbulence below. Thus, the coronal lines tied to the photosphere may be well driven by motions that are nonturbulent, making it all the more difficult to produce turbulence in the corona above.

Compressive effects to damp the waves near the Sun, such as nonlinear Landau damping based on parametric decay, will not be effective due to the small amplitude of the magnetic fluctuations. Parametric decay rates are proportional to the square of the magnetic fluctuation amplitude normalized by the field magnitude (e.g., Wong & Goldstein 1986). In the solar wind in regions such as Figure 1, where the relative fluctuation amplitude is $\sim$0.35 and the flow is highly Alfvénic, it is impossible to find cases of the purely backward propagating waves (toward the Sun) implied by parametric decay, since such inward waves either do not exist or are completely masked by large fluxes of forward propagating waves at all scales. As pointed out above, these regions do not show dissipation where the relative fluctuation amplitudes are large, so there is no reason to believe they would at amplitudes nearly an order of magnitude smaller compared to the mean field, as expected near the Sun. Similar concerns limit the effectiveness of a compressive parallel cascade; this would involve steepening that depends on the relative fluctuation amplitude, and thus this would be a slow process near the Sun.

Could the interplanetary spectrum be completely unrelated to the one that comes from the Sun? This is one way to render the observations in the solar wind irrelevant to the acceleration problem, and it is invoked, implicitly, by the Q2D and high-frequency models of the process, discussed above. In another approach, it has been suggested (Borovsky 2008; Cranmer & van Ballegooijen 2005) that the larger scale fluctuations in the solar wind might be "uncorrelated flux tubes," in which most of the observed power at scales larger than a few hr is produced by the convection of flux bundles from different footpoints on the Sun. Borovsky (2008) makes a strong case for the existence of the flux tubes, but these tubes themselves interact as part of a larger scale turbulence; this is discussed to some extent in the paper. The uncorrelated-tube viewpoint does not account for the observed space filling, highly Alfvénic fluctuations with transverse velocities at all scales that must make the tubes interact (see, e.g., Goldstein et al. 1995). There is a continuous spectrum between the "bundle power" and "wave power," and all scales are observed to participate in a turbulent evolution as the wind moves outward. Thus, there are many reasons to believe that the power observed in the solar wind, certainly at the highest two decades in Figure 1, which are within the flux tube boundary scales of Borovsky (2008), represent propagating, interacting fluctuations generated below the Alfvénic critical point. The wave power in the solar wind has to be a fossil spectrum from the Sun; this is why it is outward propagating, since in situ generation destroys the Alfvénic correlation (Roberts et al. 1987). It is not likely or perhaps even meaningful to have a damped and an undamped population coexist below the Alfvénic critical point at the same frequencies, nor is it plausible that the waves are damped, producing a steeper spectrum, and then are regenerated with a nonturbulent, flatter spectrum, all within the region below the Alfvénic critical point.

Three other mechanisms should be mentioned here. Microinstability generated waves are currently being considered as possible sources of solar wind heating and acceleration (see, e.g., Hollweg 2008). Such instabilities provide one means to generate the sunward wave flux required by cascade models. However, these waves will be on kinetic scales, and thus not relevant to models that start with large-scale waves. Also, there is a required source of free energy, such as beams in the distribution functions; this pushes the problem back one stage to how such distributions were generated, and raises the question of whether such mechanisms more efficiently produce other phenomena, such as nonthermal distributions, rather than waves. The possibility of plasmoids being accelerated by magnetic field gradients in expanding tubes (Pneuman 1983), sometimes called "melon seed" acceleration, does not provide an explanation for the unipolar fields in coronal hole flows, and is now invoked in the context of transient ejecta. Finally, magnetic reconnection is often invoked to heat coronal loops (Parker 1990). However, for the open corona, reconnection is generally considered to be a source of waves or nonthermal particle distributions, as in many of the theories cited above, since in the unipolar, uniform flows most relevant here it is only likely to be strongly active near the base of the acceleration region, consistent with the relatively quiescent, noninteracting rays in Figure 2.

4. IF NOT WAVES, THEN WHAT?

The above discussion leaves few alternatives for the production of the fast solar wind. The early studies mentioned in the introduction (see Leer & Holzer 1980) show that some nonthermal process must be at work. If it is not in the waves, the natural alternative is that it is in the particles (see also Cranmer 2009 and Marsch 2006 for extensive references and related discussion). Options that do not invoke nonthermal distributions have been suggested, but they are less plausible, as discussed above. There are two types of nonthermal features that are currently invoked in models: anisotropy (high perpendicular temperature, e.g., Hollweg 2000) and nonMaxwellian tails to the distributions (e.g., Scudder 1992; Maksimovic et al. 1997). There are no complete theories of how these features could be generated, although the above and other papers offer possible suggestions. Once such a distribution is created, the authors mentioned above and others provide models that will produce high-speed winds based on them, and these can explain many observed features of the winds.

A distribution of protons with high perpendicular temperature low in the corona leads to a mirror force that, when added to the other forces involved, provides a high-speed wind. Typically, ion–cyclotron waves are invoked to produce the anisotropies; this is a very different version of a "wave" scenario from that considered above, in that the waves are not involved in the acceleration, but the energy flux requirements and constraints on possible turbulent cascade origin of the required waves are the same or similar. Models based on non-Maxwellian tails invoke some version of the "velocity filtration" approach. In these "kinetic exospheric" models, an apparent heating occurs due to the gravitational elimination of the core of the particle distribution that leaves a broader and hence "hotter" distribution as the main particle distribution higher up in the corona (e.g., Scudder 1992; Maksimovic et al. 1997). Scudder (1992) and Pierrard et al. (2004) have also shown that preferential heating and acceleration of minor ions in the corona and solar wind are consistent with kinetic exospheric models, given sufficiently strong nonthermal tails for the distributions at the base, although wave fluxes are, in this case, large enough to provide the small energies involved.

While a turbulent cascade of the sort typically invoked does not seem plausible as a source for nonthermal distributions low in the corona, for the various reasons given above, nonthermal distributions in the chromosphere and transition region seem likely, as there is a great deal of energy present there, and nonthermal processes—reconnection, the shocks often invoked for chromospheric heating, converging fields, other induced electric fields, etc.—are highly likely. There are still issues of the effects of collisions, although these do not preclude the counterintuitive flux of energy up a temperature gradient (Dorelli & Scudder 2003). (The latter paper contains the intriguing suggestion that if temperature increases with height, even a hot Maxwellian must have come from a non-Maxwellian distribution.) If the above arguments are correct, the production of nonthermal distributions quite low in the solar atmosphere is the problem that most urgently needs to be solved to understand the hot open corona and the acceleration of the solar wind.

5. CONCLUSIONS

The conclusions of this work may be phrased as a set of questions that must be answered by any complete model of the acceleration of the fast solar wind, in addition to those posed by the usual constraints on density, speed, and temperature.

1. Can the model produce nearly equal amplitudes of waves in fast and slow streams beyond the Alfvénic critical point?

2. Does it produce an interplanetary f^{-1} spectrum up to at least 10^{-2} Hz at 0.3 AU?
3. Does it produce the observed evolution of the spectrum of fluctuations at 0.3 AU and beyond, and in particular is it consistent with the seemingly dissipationless evolution of the fluctuations at frequencies at or below $10^{-2} Hz$ in uniform streams?
4. Does the model apply the proper (line-tied) boundary conditions at the photosphere?
5. Does it treat any required damping and reflection mechanisms self-consistently?
6. If required, is there a mechanism for creating fluctuations at kHz frequencies?
7. Is the model consistent with high-resolution coronal images that, to the current limits of resolution, show straight, unstructured striations out to many solar radii?
8. Is the model consistent with convection zone models and time-dependent imaging of the photosphere and other near solar regions?

Each of these tests eliminates or strongly constrains various wave acceleration models of the fast solar wind, and no current model survives all the tests (or even, e.g., just test 2). The changes required to the models are not matters of varying parameters, but are fundamental matters of principle or of the violation of empirical constraints. Thus, the main point of this paper is to emphasize that all possible constraints on solar wind models must be taken into account, not just general conditions on thermodynamic quantities and wave variances. These arguments are, in many cases, quantitative. For example, the inadequacy of the predicted wave flux near the Sun only depends on the lack of dissipation that is well established by the observed spectral shape and evolution of the fluctuation levels, as well as by theoretical considerations based on the relative fluctuation level and the high Alfvénicity of the flows. Likewise, the visual amplitude of the fluctuations in coronal rays is a strong quantitative constraint on parallel propagating waves, and the coherence of the rays is likely to be a similar constraint on Q2D fluctuations (as yet not explicitly demonstrated). Parallel-cascade turbulence is too inefficient and slow to work, and Q2D models depend on boundary conditions that are almost certainly nonphysical. The direct production of high-frequency waves and their dissipation within the corona, while the waves observed in the heliosphere are generated in some other way, is not ruled out by the quantitative aspects of the arguments presented here, but there is no known, plausible source of such fluctuations. Thus, the arguments presented here pose very strong constraints on models of the solar wind.

The author thanks the Helios magnetometer teams for producing high-quality data, and NASA's Space Physics Data Facility for providing the files needed here. The provision of high-resolution eclipse photographs by M. Druckmüller on his excellent Web site is also deeply appreciated. The active discussion of these ideas by many participants in the Solar Wind 12 conference and in other contexts is also gratefully acknowledged, as are helpful interchanges with J. Dorelli and the referee.

REFERENCES

Axford, W. I., McKenzie, J. F., Sukhorukova, G. V., Banaszkiewicz, M., Czechowski, A., & Ratkiewicz, R. 1999, Space Sci. Rev., 87, 25
Barnes, A. 1966, Phys. Fluids, 9, 1483
Barnes, A., Hartle, R. E., & Bredekamp, J. H. 1971, ApJ, 166, L53
Bavassano, B., Dobrowolny, M., Mariani, F., & Ness, N. F. 1982, J. Geophys. Res., 87, 3617
Belcher, J. W., & Davis, L., Jr. 1971, J. Geophys. Res., 76, 3534
Borovsky, J. E. 2008, J. Geophys. Res., 113, 8110
Brummell, N. H., Hurlburt, N. E., & Toomre, J. 1996, ApJ, 473, 494
Cranmer, S. R. 2009, Liv. Rev. Sol. Phys., 6, 3
Cranmer, S. R., & van Ballegooijen, A. A. 2003, ApJ, 594, 573
Cranmer, S. R., & van Ballegooijen, A. A. 2005, ApJS, 156, 265
De Pontieu, B., et al. 2007, Science, 318, 1574
Dmitruk, P., Matthaeus, W. H., Milano, L. J., & Oughton, S. 2001, Phys. Plasmas, 8, 2377
Dmitruk, P., Matthaeus, W. H., Milano, L. J., Oughton, S., Zank, G. P., & Mullan, D. J. 2002, ApJ, 575, 571
Dorelli, J. C., & Scudder, J. D. 2003, J. Geophys. Res., 108, 1294
Goldstein, B. E., Neugebauer, M., & Smith, E. J. 1995, Geophys. Res. Lett., 22, 3389
Goldstein, B. E., Smith, E. J., Balogh, A., Horbury, T. S., Goldstein, M. L., & Roberts, D. A. 1995, Geophys. Res. Lett., 22, 3393
Hollweg, J. V. 1974, J. Geophys. Res., 79, 1539
Hollweg, J. V. 2000, J. Geophys. Res., 105, 15699
Hollweg, J. V. 2008, J. Astrophys. Astron., 29, 217
Hollweg, J. V., & Isenberg, P. A. 2002, J. Geophys. Res., 107, 1147
Hollweg, J. V., & Johnson, W. 1988, J. Geophys. Res., 93, 9547
Horbury, T. S., & Balogh, A. 2001, J. Geophys. Res., 106, 15929
Leer, E., & Holzer, T. E. 1980, J. Geophys. Res., 85, 4681
Maksimovic, M., Pierrard, V., & Lemaire, J. F. 1997, A&A, 324, 725
Markovskii, S. A., Vasquez, B. J., Smith, C. W., & Hollweg, J. V. 2006, ApJ, 639, 1177
Marsch, E. 2006, Liv. Rev. Sol. Phys., 3, 1
Marsch, E., Rosenbauer, H., Schwenn, R., Muehlhaeuser, K.-H., & Denskat, K. U. 1981, J. Geophys. Res., 86, 9199
McComas, D. J., et al. 2007, Rev. Geophys., 45, 1004
Milano, L. J., Dasso, S., Matthaeus, W. H., & Smith, C. W. 2004, Phys. Rev. Lett., 93, 155005
Ofman, L., & Viñas, A. F. 2007, J. Geophys. Res., 112, 6104
Ong, K. K., Musielak, Z. E., Rosner, R., Suess, S. T., & Sulkanen, M. E. 1997, ApJ, 474, L143
Parker, E. N. 1958, ApJ, 128, 664
Parker, E. N. 1990, Geophys. Res. Lett., 17, 2055
Pierrard, V., Lamy, H., & Lemaire, J. 2004, J. Geophys. Res., 109, 2118
Pneuman, G. W. 1983, ApJ, 265, 468
Rappazzo, A. F., Velli, M., Einaudi, G., & Dahlburg, R. B. 2008, ApJ, 677, 1348
Roberts, D. A. 1989, J. Geophys. Res., 94, 6899
Roberts, D. A., & Ghosh, S. 1999, J. Geophys. Res., 104, 22395
Roberts, D. A., Goldstein, M. L., Deane, A. E., & Ghosh, S. 1999, Phys. Rev. Lett., 82, 548
Roberts, D. A., Goldstein, M. L., & Klein, L. W. 1990, J. Geophys. Res., 95, 4203
Roberts, D. A., Goldstein, M. L., Klein, L. W., & Matthaeus, W. H. 1987, J. Geophys. Res., 92, 12023
Rouse, H. 1950, Characteristics of Laminar and Turbulent Flow, 3rd in the "Digitized Hunter Rouse Movies" series, http://www.iihr.uiowa.edu/products/dhrm.html, c. 1950
Scudder, J. D. 1992, ApJ, 398, 319
Tsuneta, S., et al. 2008, ApJ, 688, 1374
Unti, T. W. J., & Neugebauer, M. 1968, Phys. Fluids, 11, 563
Velli, M. 1993, A&A, 270, 304
Verma, M. K., & Roberts, D. A. 1993, J. Geophys. Res., 98, 5625
Vourlidas, A., Davis, C. J., Eyles, C. J., Crothers, S. R., Harrison, R. A., Howard, R. A., Moses, J. D., & Socker, D. G. 2007, ApJ, 668, L79
Wang, Y.-M., Biersteker, J. B., Sheeley, N. R., Jr., Koutchmy, S., Mouette, J., & Druckmüller, M. 2007, ApJ, 660, 882
Whang, Y. C. 1973, J. Geophys. Res., 78, 7221
Wong, H. K., & Goldstein, M. L. 1986, J. Geophys. Res., 91, 5617