Bull. Astr. Soc. India (2013) **00**, 1–**??**

The Solar Corona: What Are The Remaining Fundamental Physical Questions?

Petrus C. Martens*

Montana State University, Bozeman, MT 59715, USA

Received — ; accepted —

Abstract. The two key unresolved physical questions in our knowledge of the solar corona are: (1) How is the corona heated to a temperature of several MK, and, directly related to that, why is the coronal emission structured in nearly constant cross-section loops? And, (2) what is the mechanism that determines the onset of solar flares and eruptions, and, again directly related, can flares be predicted? I will introduce these questions, discuss some proposed solutions that are not complete, and my view on getting to the full solutions.

Keywords : Solar Corona – Coronal Heating – Solar Flare Prediction

1. Introduction

The solar corona has been observed in detail in soft X-rays and EUV for over half a century (Golub & Pasachoff 1997). Ground-based eclipse observations of the corona in visible light have been available for even much longer. There is much we understand now in terms of a physical description about the solar corona. The coronal plasma is mostly very hot, of the order of mega-kelvins, and very tenuous – so tenuous that coronal plasma is almost exclusively optical thin. The magnetic field in the corona typically has a much higher energy density than the plasma – often expressed by saying the plasma-beta is low – and therefore the coronal magnetic field is largely force-free, meaning that the gas pressure gradient and gravity play a negligible role in the force balance. Coronal plasma can move unhindered along the magnetic field lines, but it is almost impossible for it to cross the magnetic field. Hence the myriad of beautiful plasma columns our instruments observe in EUV and soft X-rays outline the coronal magnetic field to great precision. The major exception to the hot state of the corona is found in filaments and prominences, and it turns out that these structures often form a key ingredient of flares and eruptions. Filaments are

---

*email: `martens@physics.montana.edu`

formed in peculiar manner and they and their surroundings often contain large electric current systems that provide the energy for flares when dissipated and the Lorentz force that drives eruptions. The force-free magnetic field in the corona evolves driven by the motions of the footpoints of the magnetic field lines in the photosphere, and in response to the emergence and cancellation of magnetic field at the photospheric boundary.

What is that we do not very well understand about the corona? First and foremost, because the corona is much hotter than the solar atmosphere below it and also than the solar wind above it, there has to be a continuous supply of energy to it to compensate for the energy loss through EUV and X-ray radiation. The physical mechanism for this “coronal heating" is still far from understood even after decades of intensive study (Reale 2010). The obvious candidate for supplying that energy is the magnetic field, since we know there is a lot of free magnetic energy – in the form of electric currents – present in the corona. There is certainly no shortage of proposed theoretical physical mechanisms for converting magnetic energy into plasma thermal energy for the corona, with “nano-flares" (Parker 1988) the most commonly preferred at this point. However, very few of the proposed mechanisms address the major observed features of the corona: why is the emission concentrated in thin coronal loops? Why are the loops in the cores of solar active regions hotter and more stable than those in the periphery? Why can’t we still simply predict the coronal emission from high-resolution magnetic simulations that seem to map out the magnetic field reasonably well? Why do the observed coronal loops appear to have a constant cross-section, while magnetic field extrapolations produce bundles of field lines that are almost always expanding with height?

The second major unresolved problem regarding the solar corona is the prediction of the timing and magnitude of solar flares and filament eruptions – the location of major solar flares is reasonably well foreseeable from observations of the solar magnetic field (see Fig 6). This issue has major practical implications because solar flares and Coronal Mass Ejections (CMEs) that usually accompany filament eruptions, can seriously disrupt terrestrial radio communications, reduce the accuracy of the GPS system, pose radiation hazards to astronauts and people on polar jet flights, can lead to serious disruptions in high latitude power grids, etc, etc. In fact, NASA’s Living With a Star (LWS) program (*http://lwstrt.gsfc.nasa.gov/*) owes its very existence to the need to study and ultimately predict threats of solar origin to our modern technology dependent society.

In the remainder of this paper I will demonstrate that we understand the global physics of solar flares reasonably well – the build-up of free magnetic energy, the formation of flux ropes (filaments) that ultimately erupt and drive CME’s as well as the flare energy release observable as coronal radiation (Benz 2008). From the understanding of the global physics follows the knowledge of where on the Sun the flares are likely to occur. I will argue that the prediction of the timing of flares may be a problem that is analogous to forecasting the collapse of a house of cards; doable in principle, but seriously complicated, and dependent on very detailed observations in practice.

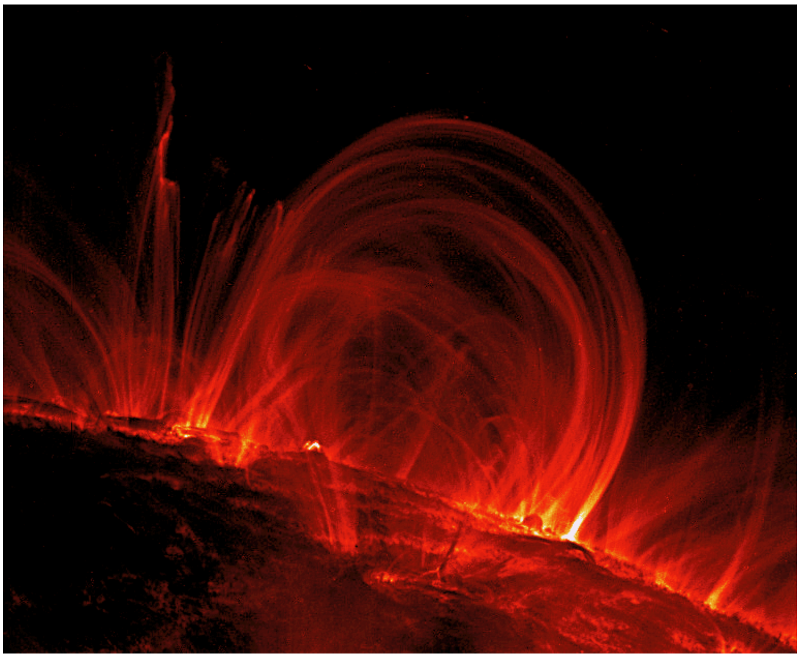

Figure 1. A TRACE 171 Å image of coronal loops on the limb., taken on November 6 1999 at 2:30:44 UT. Note how clearly the loops are outlined and their constant cross-section.

2. Heating of Coronal Loops

Fig. 1 has become the iconic image for coronal loops observed by TRACE. Loops observed in the 171 and 193 Å passbands that are responsive to cooler coronal plasma of 1.0 to 1.5 MK tend to have very clear borders, much more so than the hotter loops in other passbands, such as AIA 94 and 131 Å, and loops observed with broadband filters, such as Hinode's XRT. One of the most remarkable features of these loops is that their cross-sections appear to be constant, while potential field extrapolations produce loops that are broadest at their tops and taper off near their footpoints (López Fuentes, Klimchuk & Démoulin 2006). The very simple example of a dipole field is given in Fig. 2, clearly demonstrating the tapering off of loops at their footprints.

This tapering phenomenon is preserved in more complex fields, with one distinct exception: around the separators in the field topology, the cross-sections of flux bundles, i.e. loops, have been shown to be constant in simulations (Plowman, Kankelborg & Longcope 2009). Hence a straightforward hypothesis is that the coronal loops we observe coincide with separators. Longcope and collaborators (Longcope 1996) have demonstrated that coronal currents are preferably generated and maintained along separators. That could then explain why these loops are heated more strongly than their surroundings – since there is more electrical current to dissipate – and hence are more bright.

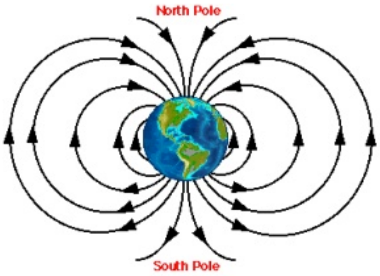

Figure 2. Sketch of the dipole field of the Earth. Note that every flux bundle emanating from the Earth's surface expands in width with height, reaching maximum width at its apex.

Fig. 3 elucidates the concepts of separatrices and separators. Separatrices are 2D surfaces in a 3D magnetic field. Magnetic field lines that are infinitely close together at one photospheric boundary of the field, but on different sides of a separatrix surface, will reenter the photosphere at finite distances from each other. Some field lines escape out to infinity with the solar wind, but those will end up at finite distances as well. Separators are field lines at the intersections of of separatrix surfaces. Separators are exceptional, as mentioned above, in that they are the preferred locations for coronal currents to build up and that the flux bundles around them have constant cross-section.

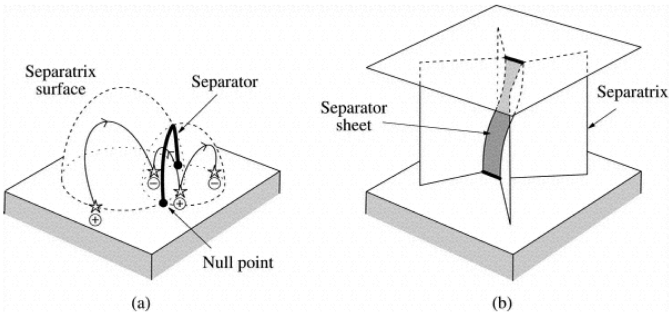

Figure 3. Separatrix surfaces and separators at their intersections in a model solar coronal magnetic field bounded by the photosphere at the bottom. From Priest et al. (2005)

Testing this hypothesis has led to disappointing results. First, an attempt to correlate potential field extrapolations from MDI magnetograms (Schrijver & Title 2002) with loops observed by TRACE has been unsuccessful. Perhaps that can be ascribed to the limited resolution of the MDI magnetograms, which would cause the extrapolation to miss many separators. Even more contrary evidence has recently been found by Scott, Martens & Tarr (2013). They found dark lanes with very low emission in the AIA passbands studied at the boundaries of Active Regions. Further analysis revealed that these dark lanes were at the intersection of separator surfaces with the photospheric boundary, exactly the opposite of what one would expect from the hypothesis above.

So, clearly, for now we will have to shelve the hypothesis that the observed coronal loops coincide with magnetic separators. While there are many proposals for coronal heating mechanisms, none of them, except for the separator hypothesis, explains at the same time why the observed loops have nearly constant cross-section. A promising suggestion, independent from the heating mechanism, was made by DeForest (2007). Deforest notes that the observed emission at the tops of EUV loops observed by TRACE is larger than what one would expect in comparing with the emission from the footpoints. Hydrostatic equilibrium holds along the axes of loops since there is no magnetic force parallel to the field. Hence the decrease in pressure with height is simply calculated from the observed loop temperature profile, and the ratio between apex and footpoint pressure follows directly. The observed emission shows that the apex density and pressure are higher than those that follow from hydrostatic equilibrium. How can that be accounted for?

First, it is possible of course that the temperature profile inside the TRACE loops is much more complex than one would find from simple passband ratios (Martens, Cirtain & Schmelz 2002), and a higher average temperature along the loop would resolve the puzzle. The solution that Deforest proposes is that the the loops we observe are not sufficiently resolved by even the best telescopes at that time, and he backs that up by comparing observations by EIT on SoHO with simultaneous TRACE images at five times higher resolution. While the EIT loops appear to have constant cross-section, the TRACE images reveal fraying near the footpoints and a smaller filling factor there. If confirmed, this leads to a new puzzle: why would the small bright subloops (strands) inside a constant cross-section container loop expand in cross-section with height, while the non-emitting remaining strands in the loop container apparently shrink with height in cross-section?

Regardless of the resolution of this puzzle, the work of Deforest indicates the need for higher resolution observations: if the five-fold improvement in resolution from EIT to TRACE and AIA led to important new insights is it not to expected that a further increase in resolution might teach us even more? This question has been answered with a very strong affirmative by the Hi-C rocket flight of the summer of 2012 (Cirtain et al. 2013), which carried an EUV telescope with a resolution of 0.2 arcsec, an increase of another factor five over TRACE and AIA. A comparison between the EUV 193 Å images of that flight and simultaneous AIA images is shown in Fig. 4. These images are consistent with several of the claims made above: the coronal loop has constant cross-section and it is frayed at the end. What is new in the Hi-C images is that they reveal strands inside the AIA loop that twist around the axis and cause heating events where they touch each other, as in the case indicated by the arrow. The corresponding AIA 94 Å image shows that the heating event generates hot 7 MK plasma.

Note that if the individual strands represent current concentrations then there must be separatrices and separators inside the loop, which may explain the constant cross-section according to Plowman et al. (2009). Also note that the fraying at the ends is consistent with Deforest's original observations and can explain excess brightness at the loop tops.

Questions that remain are the following. Hi-C has observed relatively cool (≈ 1.5 MK) loops, and Active Region loops tend to be hotter and denser. In fact, the "moss" (Berger et al. 1999)

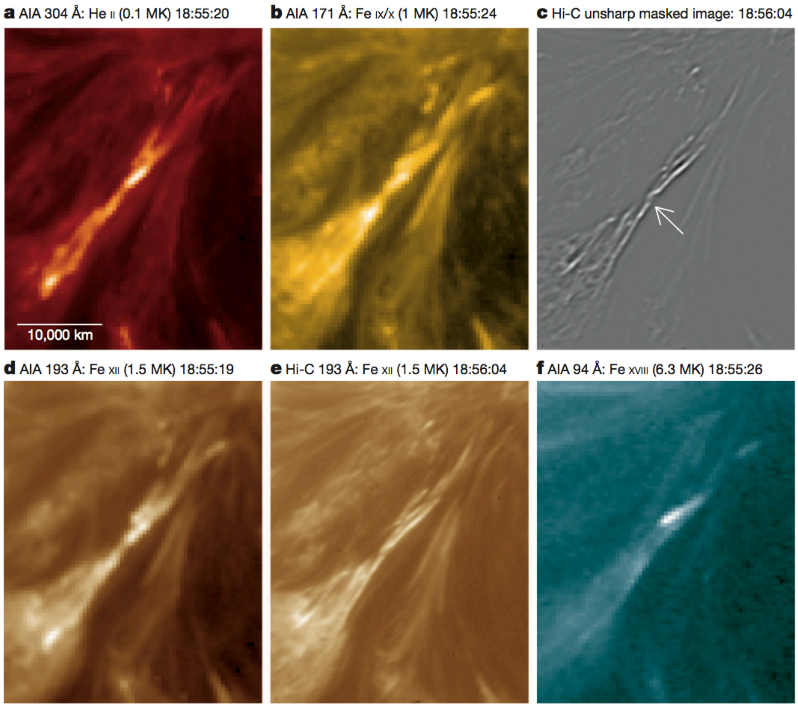

Figure 4. A coronal loop seen at several different coronal temperatures by AIA and Hi-C. Labels indicate the passbands, the dominant emitting ion, and the peak formation temperature of the ion. The unsharp mask image enhances the shapes of fine-scale structures relative to the image background. All panels show a narrow magnetic arch that Hi-C resolves to be twisted along its length. The twisted loops converge and appear to intersect (arrow in c). From Cirtain et al. (2013).

associated with the footpoints of these hotter loops is clearly visible in the Hi-C images. Does the loop internal structure observed by Hi-C also occur in the hotter active region loops? Do events like this release enough energy to explain active region loop emission? Or is there a need for additional steady heating? What drives the currents in the strands within the loop? As for the latter, there have been interesting results from Wedemeyer-Böhm, Scullion, Steiner, Rouppe van der Voort, de La Cruz Rodriguez, Fedun & Erdélyi (2012) that show that there are strong small-scale rotational motions in the photosphere and chromosphere that may generate the braiding of strands that is observed. What remains to be established is that these vortex motions coincide with the footpoints of braided coronal loops.

Answering the questions above requires much additional observations at Hi-C resolution. The key result of the Hi-C flight is though that we may have resolved the elementary magnetic structures that interact in heating the corona, which, if confirmed, would be a milestone in explaining coronal heating.

3. Predicting Solar Flares

Decades of observations in in all wavelengths from radio to γ-rays have taught us much about solar flares (Benz 2008). One of the reasons to be interested in flares is that they can pose considerable threats to our technological infrastructure, from radiation hazard to astronauts and airline passengers, to power surges in large-scale electrical grids, to damage to satellites electronic systems, to radio communications that depend on reflections of waves from the ionosphere, and to irregularities in GPS accuracy to name just a few. Therefore it is desirable to have the capability to predict the timing and severity of solar flares to take protective measures and mitigate the damage.

We know from observation that solar flares are often accompanied by large eruptions, those of Hα-filaments, and Coronal Mass Eruptions (CME's), the latter best observed with coronographs such as the one on LASCO. CME's represent giant magnetic loops that travel all the way through the heliosphere and that can great great havoc when they collide with the Earth's magnetosphere, leading, for one thing, to the magnificent display of Northern Lights. It can take one or several days for a CME to impinge on the Earth and from satellite observations we can see them coming. Flares on the other hand represent the large burst in radiation that comes with the launch of CME's, and their arrival is much less predictable.

The connection between flares and CME's is reasonably well understood. See Fig. 5 for an early cartoon. Many such schematic drawings have been produced over the years, and it is interesting to inspect their evolution at *http://solarmuri.ssl.berkeley.edu/ hhudson/cartoons/*.

As the figure shows, an electric field is induced in the reconnecting current sheet below the erupting magnetic flux tube that forms the filament. That electric field is thought to accelerate the electrons and ions that impinge on the chromosphere with great speed thereby generating the observed radiation, from EUV, to X-ray and γ-rays. The eruption of the magnetic flux tube is driven by the Lorentz-force, in this case a simple hoop force that can no longer be contained by the downward magnetic pressure of the overlying magnetic field. Clearly the eruption and the flare are closely connected both resulting from a magnetic instability or lack of equilibrium.

In three dimensions the magnetic topology of flaring regions is more complex of course. The structure that emerges is more complex with multiple separatrices and separators, just as in depicted in Fig. 3. Longcope, Des Jardins, Carranza-Fulmer & Qiu (2010) have shown convincingly for a number of flares that they studied that the brightest X-ray loops in flares coincide with separators. Note, as was pointed out above, how that is different for stationary X-ray loops in active regions.

Flares always occur over a polarity inversion line (PIL), for large flares always within an active region where the magnetic field is strong (hundreds of Gauss) and strongly sheared, i.e. there is a strong component of the magnetic field parallel to the PIL. The presence of a filament that later erupts very often accompanies a large flare, and a very distorted PIL is another factor conducive to flares. The largest M- and X-class flares invariably occur in the largest active re-

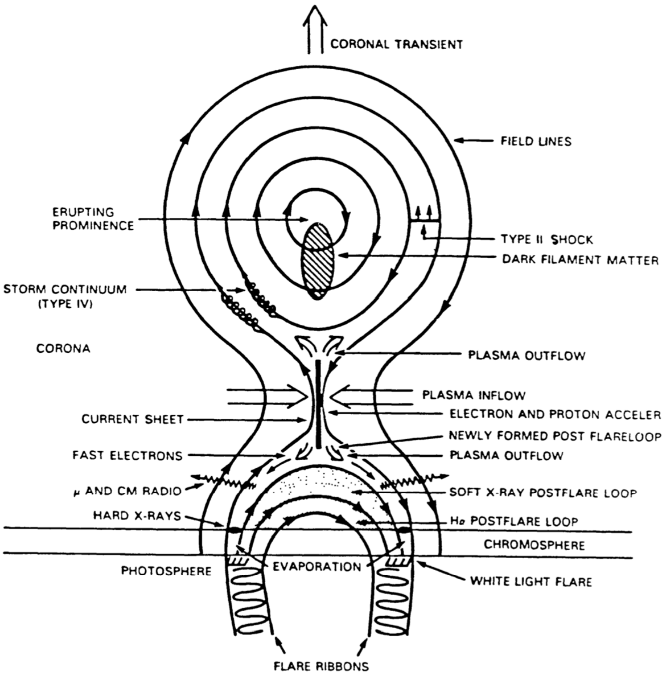

Figure 5. Schematic drawing of the magnetic field topology (drawn lines) and main physical processes (labels) in a solar two-ribbon flare that is accompanied by a filament eruption and a CME. The figure represents a cross-section perpendicular to the magnetic polarity inversion line. From Martens & Kuin 1989.

gions with the strongest magnetic fields with large shears. Fig. 6 shows the magnetic field of an emerging active region with in red the PIL segments where a flare is likely to occur.

The multi-million dollar question is in predicting the exact timing of flares. Clearly one would like to know that to avoid the hazardous X-ray and γ-rays, but a very general warning for a flare potential is not enough; it is imperative to avoid too many false warnings, so that airlines and NASA can take action to protect passengers and astronauts without wasting a lot of resources.

The argument I will make in the rest of this paper is that the prediction of the timing of flares, while it must be a solvable physical problem in principle, may still escape solar physicists for a long time to come. The analogy is with a house of cards (see Fig. 7). We understand very well the physics involved in both building and the collapse of a house of cards. The weight of the cards at the top is distributed along the cards leading to a vertical and a horizontal force at the bottom edge of each of the cards. The horizontal force is destabilizing but it is offset by the static slip resistance force of the edge of the card over the surface of the horizontal card on which it rests. Once the horizontal force resulting from the weight of the card overcomes the static

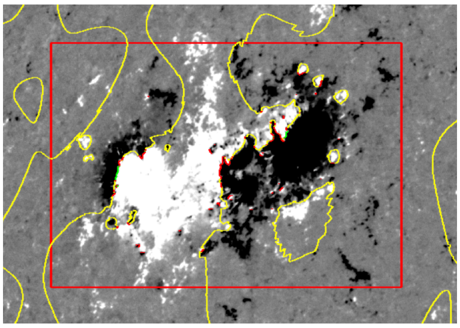

Figure 6. Magnetogram of an emerging active region observed by the Helioseismic and Magnetic Imager (HMI)) on SDO taken on March 6 2011 at 00:23:56 UT. White indicates positive polarity and black negative. The polarity inversion line is indicated in yellow with the red segments indicating the regions where the magnetic shear is larger than a certain threshold, and a flare is likely to occur. Figure courtesy of Alec Engell (MSU).

slip resistance the bottom edge of the card will start moving outward resulting eventually in the collapse of part of or the whole structure.

But now comes the problem of predicting the exact timing of the ultimate collapse of the house of cards. For each card that is not horizontal one has to keep track of the horizontal force at the edge which is determined by the weight resting on the card and the angle with the vertical that the card makes. Clearly doable but requiring a lot of continuous measurements. The gravity force is offset by slip resistance. Slip resistance in turn depends on the roughness of the surface of the card edge and the surface of the card below it (e.g. is it an new or an old deck of cards), the presence of an adjacent edge, the presence of water and/or grease on the cards, etc. All these factors can in principle be accounted for but continuously keeping track of all of them is a difficult, yet required task.

Consider now the magnetic structure of a potentially flaring active region. Assume we understand the physics involved, which is not a certainty at all. However, unlike as for the the house of cards we may not be able to resolve each of the individual cards in sufficient detail, or at all – our telescopes have very limited resolution. Hence it is quite possible that we are not able to carry out the measurements required to determine the stability of the active region house of cards. In my personal view the key measurement for flare prediction is that of the electrical currents in active regions; it is after all the currents that carry the free magnetic energy for the flares, and the MHD or plasma instabilities that may trigger flares are related to the current strength.

For example, a standard instability in plasmas is the current instability in which the Ohmic resistance in the plasma shoots up exponentially when the electron current drift velocity approaches the sound velocity of the plasma. As the plasma current slowly increases all of a sudden the release of a lot of energy is triggered and that may destabilize the entire magnetic house of cards. Measuring the drift velocity and the sound velocity in loops prior to a flare is still completely out of the question with our current instrumentation.

The conclusion must be that not knowing the exact physical process that triggers a flare –

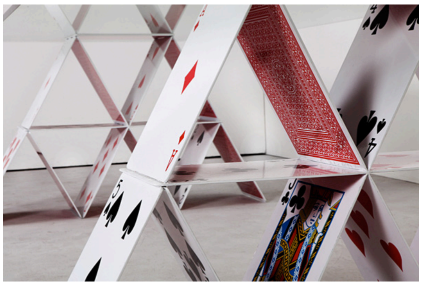

Figure 7. A house of cards. The stability of the structure is determined by a card by card balance between the force from the distributed weight of the cards and the staid slip resistance, as described in detail in the text. The house of cards may be an analog of the magnetic stability of an active region, explaining why the exact prediction of a flare is such a difficult problem.

and there may be more than one – and henceforth not knowing the instrumental capabilities that are required to monitor that physical process, we are likely still far removed from being able to predict the exact timing of the onset of flares.

References

Benz, A. O. 2008, Living Reviews in Solar Physics, 5, 1

Berger, T. E., De Pontieu, B., Fletcher, L., Schrijver, C. J., Tarbell, T. D., & Title, A. M. 1999, Sol. Phys., 190, 409

Cirtain, J. W., Golub, L., Winebarger, A. R., de Pontieu, B., Kobayashi, K., Moore, R. L., Walsh, R. W., Korreck, K. E., Weber, M., McCauley, P., Title, A., Kuzin, S., & Deforest, C. E. 2013, Nature, 493, 501

DeForest, C. E. 2007, ApJ, 661, 532

Golub, L., & Pasachoff, J. M. 1997, The Solar Corona (Cambridge, UK: Cambridge University Press, September 1997,)

Longcope, D. W. 1996, Sol. Phys., 169, 91

Longcope, D. W., Des Jardins, A. C., Carranza-Fulmer, T., & Qiu, J. 2010, Sol. Phys., 267, 107

López Fuentes, M. C., Klimchuk, J. A., & Démoulin, P. 2006, ApJ, 639, 459

Martens, P. C. H., Cirtain, J. W., & Schmelz, J. T. 2002, ApJ Letters, 577, L115

Martens, P. C. H., & Kuin, N. P. M. 1989, Sol. Phys., 122, 263

Parker, E. N. 1988, ApJ, 330, 474

Plowman, J. E., Kankelborg, C. C., & Longcope, D. W. 2009, ApJ, 706, 108

Priest, E. R., Longcope, D. W., & Heyvaerts, J. 2005, ApJ, 624, 1057

Reale, F. 2010, Living Reviews in Solar Physics, 7, 5

Schrijver, C. J., & Title, A. M. 2002, Sol. Phys., 207, 223

Scott, J. T., Martens, P. C. H., & Tarr, L. 2013, ApJ, 765, 82

Wedemeyer-Böhm, S., Scullion, E., Steiner, O., Rouppe van der Voort, L., de La Cruz Rodriguez, J., Fedun, V., & Erdélyi, R. 2012, Nature, 486, 505