SOLAR FLARES AND MAGNETIC TOPOLOGY

by

Peter A. Sturrock

CSSA-ASTRO-87-15
October 1987

(to be published in Solar Physics)
SOLAR FLARES AND MAGNETIC TOPOLOGY

P.A. STURROCK
Center for Space Science & Astrophysics
Stanford University, Stanford, California 94305

ABSTRACT. This article is a very brief review and comparison of the observational properties of flares and theoretical concepts of models of flares, especially the concepts of magnetic topology and its evolution. We examine the "environmental" aspects of flare behavior. Some of these aspects must be consequences of unknown processes occurring below the photosphere. Other aspects involve structures--such as filaments--that are closely related to flares. We then examine properties of flares to try to distinguish the different phases of energy release that can occur in the course of a flare. Finally we offer a schematic scenario and attempt to interpret these phases in terms of this scenario.

1. INTRODUCTION

It is normal practice for a conference organizer or an editor to categorize a talk or an article as either "observational" or "theoretical." This may be an indication--if not a reason--that we spend too little time looking at the interface between observation and theory. In addressing a very complex problem, such as that of solar flares, the relationship between observation and theory needs at least as much attention as pure observational studies or pure theoretical studies. This article represents only a quick look at this relationship.

It is interesting to see how the nature of the solar flare problem has changed in response to observational information and theoretical concepts. When Carrington (1859) first observed a flare in 1859, he described it in meteorological terms as comprising "clouds" high above the surface of the sun. It is notable that Carrington did not assert a cause and effect relationship between the flare and the remarkable geomagnetic and auroral disturbances that followed 26 hours later.

From the theoretical side, understanding of the flare process was strongly influenced by Kelvin's theorem (Kelvin, 1892) that the sun could not possibly be responsible for geomagnetic disturbances. His mathematics was impeccable, but of course his model was incorrect since he had no way of knowing about the solar wind! Kelvin's influence was so great that, even in 1931, Hale (1931) himself still referred to "solar eruptions and their apparent terrestrial effects." One is reminded of three famous statements (especially the second) attributed to Arthur C. Clark:
"If a famous scientist says something is possible, he is probably correct.

If a famous scientist says something is impossible, he is probably wrong.

Science of the next century would be regarded as magic if it were seen today."

The introduction of the magnetograph showed the clear association of flares with magnetic field. As early as 1933, Swann (1933) proposed that flares may involve particle acceleration by the betatron process. Probably the first theoretical analysis that takes account of the high conductivity of the solar atmosphere was that of Giovanelli (1946), who proposed that flares are to be interpreted as electrical discharges due to changing magnetic fields. At this time, flares were referred to as "chromospheric flares," since the only data available were from visible light (primarily H-alpha), that was believed to be produced entirely at the chromospheric level. It was implicitly assumed that the energy released in the flare was also confined to the chromospheric layer, and this led to high estimates of the required magnetic field.

In 1953, Dungey (1953) showed that at an X-type neutral point, the electric field due to a change in magnetic field leads to a current that tends to enhance the change in the magnetic field. That is, Dungey introduced the two concepts of instability and reconnection. These concepts have dominated theoretical studies of solar flares ever since.

It was not until the late 1950s that rocket-borne experiments made it possible to detect X-rays from solar flares (Chubb et al., 1961). X-ray images obtained from the OSO series of spacecraft and other spacecraft have shown that flares are primarily coronal phenomena; the chromospheric response is now regarded as a "secondary" aspect of the flare process.

In the late 1950s and 1960s, a variety of models of flares were proposed (Sweet, 1958; Gold and Hoyle, 1960; Carmichael, 1964; Sturrock, 1968). Each proposed a magnetic field configuration that was supposed--implicitly or explicitly--to be metastable but subject to an instability that would suddenly release free energy associated with coronal currents. This energy release was identified with the impulsive phase of solar flares. Hence, in most cases, the comparison between observation and theory comprised just two numbers: the energy released in a large solar flare (about $10^{32}$ erg), and the time scale for the impulse phase (about $10^2$ sec).

2. SELECTED ITEMS RELEVANT TO THE FLARE PROBLEM

It was implied, in the previous section, that early flare models addressed only a restricted aspect of the flare
problem, namely, the nature of energy release that occurs during the impulsive phase. Subsequent research indicates that there are more phases of energy release that need to be considered. Furthermore, there are other aspects of the flare problem that constitute a challenge to theorists. Some of these are listed below.

Concerning the "environmental" situation in which flares occur, it has long been recognized that they occur in active regions, and the case is overwhelming that the connection is due to magnetic field. However, there are some puzzles concerning the relationship between flares and active regions. For instance, in any solar cycle there is a tendency for flares to occur in certain bands of longitude, sometimes called "active longitudes" (Svestka, 1976). Even taking the longitude factor into account, it is remarkable that some active regions produce far more than their "fair share" of flares; these are sometimes termed "super-active regions" (Bai, 1987). It was recognized long ago by Ellison and his collaborators (Ellison et al., 1960) that flares sometimes occur in sequences, with the remarkable property that the flares of a given sequence are very similar to each other. These are called "homologous" sequences of flares. As Ellison once remarked, it is as if the magnetic field that is destroyed by one flare is rebuilt in its original form in time for the second flare.

More recently, analysis of gamma-ray flares observed by instruments on board the Solar Maximum Mission spacecraft led to the remarkable discovery that these flares tend to recur at an interval of about 150 days (Rieger et al., 1984). This periodicity is found to be present also in microwave data and appears to be present in other related data as well (Bogart and Bai, 1985). Bai and Sturrock (1987) have shown that the simplest interpretation, that the periodicity is due to the interplay of two different factors rotating at different rates, is invalid. The periodicity appears instead to be a global solar phenomenon.

Another important "environmental fact" concerning solar flares is that large two-ribbon flares are invariably associated with filaments. This association was stressed some years ago by Kiepenheuer (1963) who wrote:

"Those who have seen in an accelerated movie the brightening of a flare out of a dark filament, and the almost chaotic interaction of bright and dark structures, will not doubt the existence of a causal relation between the activation of a dark filament and the formation of a flare."

Further important evidence of the association between filaments and flares comes from a study of precursors to solar flares. It was pointed out some years ago by Smith and Ramsey (1964) that filaments are typically "activated" before a flare. This activation shows up as movements that can be detected in dopplergrams. Kahler et al. (1987) have
pointed out that the commencement of the flash phase of a solar flare, as observed in H-alpha, occurs at the same time as rapid motion of the associated filament.

Another form of pre-flare "activation" was discovered during the Skylab Mission. It was found that X-ray brightenings typically occur in an active region shortly before a flare occurs. This process was referred to as "pre-flare heating" (Van Hoven et al., 1980). It would indeed be interesting to know what relationship, if any, exists between filament activation and pre-flare heating.

3. PHASES OF SOLAR FLARES

The simplest flares involve a sudden brightening in H-alpha on a time scale of minutes. This is usually referred to as the "impulsive phase." Earlier studies, based entirely on H-alpha data, use the terms "flash phase" and "expansion phase." This phase has traditionally been the main focus of attention of theorists, as was pointed out in the introduction. However, it is not clear that the impulsive phase itself provides the secret of the entire flare phenomenon. We have already noted that there is often activity in an active region just before the flare occurs. Furthermore, there is evidence for independent energy release following the impulsive phase. Different sequences of energy-release phases seem to take place in different classes of flares. In what follows, we attempt to identify the minimum number of phases of energy release.

A. Activation Phase. This term subsumes filament activation and pre-flare heating described in the previous section. These processes clearly represent the sudden conversion of energy from one form to another. It is likely that the basic form of energy that is being tapped is magnetic free energy, that is the energy associated with currents in the solar atmosphere.

B. Impulsive Phase. This is probably the most dramatic phase of a solar flare. In H-alpha, it can be visible as the sudden appearance of an H-alpha brightening that rapidly extends along the length of a filament. It may be accompanied by a rapid hard X-ray burst and by continuum gamma-ray emission and line gamma-ray emission. When there is hard X-ray emission, there is usually also microwave radiation. It appears that there is a sudden conversion of energy, probably beginning with magnetic free energy, into the heating and acceleration of electrons and ions. The various forms of radiation may then be interpreted as secondary processes due to this heating and acceleration. Beginning with the impulsive phase, but extending much later, there may be a soft X-ray burst that can be interpreted as bremsstrahlung from a hot dense coronal plasma. The current view is that this "flare plasma" is
produced by "chromospheric evaporation," in response to the sudden flux of energy from the corona to the chromosphere.

C. Gradual Phase. Bai (1986) has summarized data showing that, in some flares, there is a second more gradual period of particle acceleration after the impulsive phase. It may come five or ten minutes after the impulsive phase and last for five or ten minutes or more. When this phase occurs, it may be more effective at accelerating particles to very high energies than was the impulsive phase. Hence such flares may be "microwave rich," and the inferred electron energy distribution may have a harder spectrum than that produced during the impulsive phase. Flares that exhibit a gradual phase of particle acceleration appear to be associated with coronal mass ejections.

D. Late Phase. It has already been pointed out that the soft X-ray emission that begins during or before the impulsive phase typically extends considerably after the impulsive phase. It is likely that, in some cases, this continued X-ray emission is due simply to bremsstrahlung from the hot flare plasma that was evaporated from the chromosphere during the impulsive phase. However, from detailed analysis of the flare of 1973 September 5, Moore et al. (1980) showed that the energy content of the flare plasma increased after the impulsive phase, showing conclusively that there was continued energy release after the impulsive phase (see also Wu et al., 1986). The time scale of this release was of the order of hours. It is this process that is referred to as the "late phase" of energy release.

H-alpha observations alone give evidence for such a late phase of energy release. In the case of large two-ribbon flares, one finds that the ribbons slowly separate, with speeds of order 1 km/s, and this drift can continue for many hours. It is clear that, in such cases, new regions of the chromosphere are receiving a flux of energy from the corona. This phenomenon also points to a slow energy release high in the corona, and we assume that this is another manifestation of the "late phase" of energy release.

Observations reveal other forms of energy associated with solar flares. For instance, the bulk of the energy released during a flare may be associated with the kinetic energy of a coronal transient. Coronal transients are closely related to filament eruptions. Hence we see once more that it is difficult to disentangle the so-called flare process from the process of filament eruption.

During most gradual flares all the above-described phases occur. However in some gradual flares, the hard X-ray emission time profiles do not show impulsive behavior but show gradual behavior from the beginning (e.g., 1981 May 13 gradual flare). In impulsive flares, the gradual hard X-ray emitting phase does not develop, and the late phase also
is lacking. The gradual soft X-ray emission observed after the impulsive phase is thermal radiation by plasma heated during the impulsive phase.

4. A SCHEMATIC SCENARIO FOR SOLAR FLARES

Part of the flare problem is to understand the "environmental" factors that were discussed in Section 2. This means that part of the flare problem requires an understanding of subphotospheric processes. The existence of active longitudes and the development of superactive regions may be a manifestation of long-lived giant convection cells. The interpretation of the 150-day periodicity remains a puzzle.

Figure 1. Schematic model of evolution of a twisted subphotospheric flux tube after it erupts through the photosphere. (a) A "side view" ignoring the effect of the sub-photospheric twist on the coronal configuration. (b) An "overhead view" showing how the coronal field will be modified by the sub-photospheric twist.
Some of the environmental factors can be understood on the assumption that an active region develops from the eruption of a twisted subphotospheric flux tube. After the tube penetrates the photosphere to extend into the corona, it is energetically favorable for the subphotospheric portion to unwind and transmit the subphotospheric twist through the photosphere into the corona. This has the effect of producing rotations at the photosphere, as indicated in Figure 1, including in particular a high shear along the neutral line. Such shear is a notable property of flare-producing regions. Furthermore, such shear is indicated by the fine-scale structure of filaments.

The concept of untwisting of the subphotospheric field provides a natural interpretation of homologous flare sequences. As the twist proceeds, it builds up a stressed magnetic field that ultimately reconnects to return the coronal region to an unstressed current-free form. However, continued unwinding reconstructs the stressed field so that the flare recurs, and the sequence can continue until the sub-photospheric tube is relieved of its helical stress.

Figure 2. Schematic representation of possible magnetic field configuration of a filament (Sturrock et al., 1984).

A crucial question for understanding the flare process is that of the structure of the magnetic field associated with a filament. The field is certainly highly sheared, and it probably also contains a great deal of fine structure, since it is known that most of the magnetic flux at the photosphere is confined in small "magnetic knots," with field strengths of order 1,000 to 1,500 gauss and diameters of order 300 kilometers (see, for instance, Tarbell and Title, 1977). This implies that the magnetic field of a filament probably resembles a rope with many strands, as indicated in a highly simplified form in Figure 2. Each magnetic strand is rooted in the photosphere, one root on
either side of the neutral line. Where two feet, on opposite sides of the neutral line, are in close proximity, there is the possibility that reconnection will occur. Reconnection has two consequences. One is that energy is released that may lead to heating of the atmosphere, resulting in X-ray emission from the coronal level and H-alpha brightening at the chromospheric level. The other effect is that the balance of forces is changed, so that part of the filament can now move upwards, and more stress is placed on the remaining strands of the filament. This may lead to further reconnection of adjacent strands, and this process may at some point develop into a catastrophic process in which all connections along the length of the filament tear (rather like tearing a piece of cloth), except that the field must remain rooted at the two ends of the filament. The effect of this process is then the formation of a large twisted flux tube, rooted at its end points, as shown in Figure 3.

Figure 3. Schematic representation of the development of an extended current sheet beneath an erupting filament (Sturrock et al., 1984).

The early spasmodic reconnection is interpreted as the cause of both filament activation and preflare heating that comprise the "activation phase" of a flare. The rapid runaway tearing that disconnects the filament almost entirely from the photosphere is identified in this model with the impulsive phase of the flare. As we shall see, this particular model leads to a two-ribbon flare. After the filament is disconnected from the photosphere, magnetic stresses will cause it to rise.
However, in so doing, it is attempting to move through a region that already contains plasma permeated by magnetic field. Hence this stage of the process involves the "brushing" of one set of field lines past another. This stage seems to correspond to the gradual phase of acceleration. If the relative velocity of the "filamentary" field lines with respect to the "coronal" field lines is sufficiently large, then the Kelvin-Helmholtz instability should occur (Dobrowolny, 1972). This will lead to MHD turbulence in the vicinity of the interface, that could then lead to acceleration of electrons and ions. It is possible that acceleration during the impulsive phase also is due to such a mechanism since it has proved difficult to explain particle acceleration as a direct result of magnetic reconnection.

As a result of the eruption of a filament, the overlying coronal field lines will form a current sheet below the erupting filament, as indicated in Figure 3. It has been proposed (Sturrock et al., 1984; Cliver et al., 1986) that the slow reconnection of this current sheet is responsible for the "late phase" of a solar flare, resulting in soft X-ray emission and in the drifting two-ribbon topology of large flares.

Figure 4. Schematic representation of a toroidal magnetic flux tube encircling an erupted prominence, as a result of the reconnection indicated in Fig. 3. The toroid would be detectable as a stationary type IV radio burst (Sturrock, 1986).
As a result of this reconnection, the coronal field below the filament can return to a current-free form, being visible as a post-flare loop-prominence system (Figure 4). However part of the coronal magnetic field is likely to form a toroidal configuration wrapping around the filament. Some high-resolution coronograph photographs and images of coronal transients show evidence that the magnetic field of a transient is comprised of two different flux systems as would be expected on the basis of the present model.

Figure 5. Schematic representation of situation that arises when a filament, encircled by a toroidal flux tube, is completely ejected from the sun. The toroid would be detectable as a moving type IV radio burst. The shock wave would give rise to a type II radio burst (Sturrock, 1986).

It may be that, in some cases, the filament erupts but does not expand out into interplanetary space. In other cases, the filament may be so highly stressed that it expands outward, attempting to form an open field
configuration. If this occurs, one must expect a shock wave to develop ahead of the expanding configuration, as indicated in Figure 5.

5. DISCUSSION

The flare problem, at its current stage of evolution, now comprises a closely interlocking set of sub-problems involving the acquisition and interpretation of data, the analysis (by a combination of analytical and numerical techniques) of theoretical concepts, and the difficult task of assembling these concepts into a comprehensive model that is theoretically self-consistent and in accord with observational facts. After this brief review of the interface between observation and theory, we are clearly faced with the need for more detailed and specific observational data. We need observations with higher spatial resolution, and sufficiently high temporal resolution, to help us determine the spatial and temporal relationship of various forms of radiation from flares, and hopefully to obtain a more detailed picture of the magnetic field structure that precedes a flare and how that structure changes as the result of the flare. The continuous acquisition of high quality vector magnetograph data will be most valuable, but they should be supported by high-resolution X-ray images.

These observational challenges are matched by comparable theoretical challenges. We are faced with the general problems of trying to model possible pre-flare plasma-magnetic-field configurations, and then following the evolution of that system, realizing that a plasma process in one part of the system is likely to influence a neighboring region and stimulate a similar--or a different--process in that region. To take a specific example: It is usually assumed that reconnection plays a key role in the flare process; the theory that is invoked follows the pattern of the Furth, Killeen and Rosenbluth (1963) tearing-mode analysis. However, in such analyses, it is assumed that conditions over a large current sheet evolve only as a function of time. A more likely evolution in a solar flare is that reconnection commences in one part of a large sheet and then propagates to other parts of the sheet. How does this occur? Can it occur sufficiently rapidly to be consistent with the proposal that such reconnection is the explanation of the impulsive phase of a flare?

However, after a general picture has emerged, has been tested, and appears to be satisfactory, observers and theorists will face a more difficult challenge--that of understanding the development of a specific solar flare from detailed observations of the pre-flare state, and detailed theoretical modeling based on those observations. Perhaps we will be ready for this challenge at the beginning of the 21st Century.
ACKNOWLEDGEMENTS

This work was supported in part by Office of Naval Research Contract N00014-85-K-0111, by NASA Grant NGL05-020-272, and as part of the Solar-A collaboration under NASA Contract NAS8-37334 with Lockheed Palo Alto Research Laboratories.

REFERENCES


END
DATE
FILMD 3-88
PTIC