THE OPTICAL FLARE

H. Zirin
Big Bear Solar Observatory
California Institute of Technology

ABSTRACT. Optical observationd now present considerable information on the flare process. It is always associated with filaments and with simplification of existing magnetic connections, and it arises from the emergence and expansion of new flux.

The optical flare divides into impulsive phase, with multiple flashes along the neutral line, and thermal phase, with two-ribbon expansion. The former bears some resemblance to tearing mode phenomena.

The appearance of loops in emission requires very high densities in those phenomena.

The ratios of the hydrogen lines, the excitation of HeII 4686, and the relation of vertical to horizontal structure all remain to be explained.

1. INTRODUCTION

The great Russian humorists Ilf and Petrov pointed out that "pedestrians must be respected." Pedestrians discovered the law of gravitation, the differential calculus, even the automobile. The optical observers are thus the pedestrians of flare research. Most of what we know about flares today is the result of optical observations. Although it would be most difficult to interpret these without the remarkable radio and X-ray observations reported at this conference, the optical data are still our most powerful tool in understanding what governs the flare. Why? What question do we answer by optical observations? After all, most of us agree (but it is not proven) that the initial flare energy release occurs in very high temperature material invisible in the optical range.

Yet optical observations have the following advantages:

Solar Physics **86** (1983) 173–184. 0038–0938/83/0861–0173 $ 01.80.

1) An immense number of photons are produced, permitting high space and time resolution. A flare occupies hundreds of pixels in optical data - a few in X-ray or radio pictures.

2) High resolution telescopes for optical work are easy and cheap to build.

3) Magnetic fields are easily measured.

4) The flare results from preflare conditions that are invisible to X-rays or radio waves, even with low resolution. The origin of the flare must thus be studied in the optical.

5) Almost all high energy flare effects are reflected in optical effects. With improving understanding, we can pick out different elements to deduce time behavior of HXR, SXR and other radiations.

6) It is quite conceivable that the filament that is part of almost every flare is in fact the seat of the energy release, hence it is possible we already see the flare energy release process in the visible but don't understand it.

2. INFORMATION ON LOCATION AND CIRCUMSTANCES OF FLARES

It must be recalled that almost no data on flare location and morphology existed until the introduction of the magnetograph at Mt. Wilson and high quality Hα cinematography by Ramsey and Moreton at Lockheed. Severny (1958) was the first to worry about the location of flares occurring on neutral lines in the longitudinal fields. The Lockheed observers early recognized the activation of filaments before the Hα flare. So, since 1958-1960 we have known that the flare takes place along the longitudinal neutral line. This is one of the key facts in the flare problem.

The fact that the flare took place at a point where the field was not zero bothered many people and made them dubious of this result. But modern day high resolution magnetograms and Hα grams permit us easily to resolve the problem.

In general, two kinds of longitudinal neutral lines are seen in active regions. Both are marked in Hα by dark features. In one case these dark features run directly from one polarity to the other, and for this reason I called them field transition arches (FTA). They are always perpendicular to the neutral line. In the other kind, the field lines run parallel to the neutral line; these we know as filaments, and this is where flares occur. The reason of course is obvious: perpendicular fibrils involve shorter field lines and much less energy than filaments.

How do optical observations demonstrate this fact?

1) The pre-flare fibrils are seen to run parallel to the neutral line.

2) Measurement of fields in filaments show them to be parallel to the axis.

3) Flares along parallel neutral lines show bright elements in either polarity but displaced along the neutral line.

4) After flares we see loops perpendicular to the neutral line and still later, the parallel fibrils are seen to have been replaced by perpendicular fibrils.

Optical observations have established that, except for occasional small events, flares do not occur away from the general neighborhood of filaments. The parallel boundary is established by sunspot motions, usually connected with newly erupting flux. Because reconnection cannot take place quickly, the expanding flux pushes the old flux aside and a sheared boundary parallel fibril results. It is well established that flares occur as the result of sunspot motions. The prime generator of this is flux emergence in the middle of existing fields. In other cases spot motion occurs for mysterious reasons. For example, the spots of August 1972 were inverted and apparently the p spot was pulled forwards by the unknown Hale-Nicholson force.

A similar event occurred in June 1982. A pair of spots appeared close together in NS alignment on the east limb. Between June 4 and June 6 the following spot (or, more likely, preceding if the group belonged to the other hemisphere) moved rapidly westward, covering 30,000 km in 2 days and culminating in a huge flare on June 6 (Figure 1). The rapid east-west motion produced the sharp east-west neutral line that was the site of the big flare, but many flares also took place on the advancing edge of the moving spot. The flare is the way the magnetic field finally adjusts to the changes produced by spot motion; changes which cannot be accommodated by diffusion but only by instabilities. The process seems quite similar to the process of earthquakes which are also of common interest in California and Japan.

3. THE OPTICAL FLARE: APPEARANCE IN Hα

It did not take very good data for early observers to recognize the two-ribbon flare. Virtually all large flares show this property, the strands starting close together at the neutral line, and moving apart at 5-15 km/sec as the thermal phase of the flare continues. As can readily be seen from pictures, such as those of the August 7, 1982 flare (Zirin and Tanaka, 1973), the two strands are connected by bright Hα loops. X-ray pictures of post-flare loops (Svestka, 1976, Figure 20b) clearly show the association between the post-flare coronal condensation and the bright Hα loops. The appearance of bright Hα loops against the disk is a diagnostic of a definite minimum density and temperature; to my knowledge the values have not been

6 JUNE 82

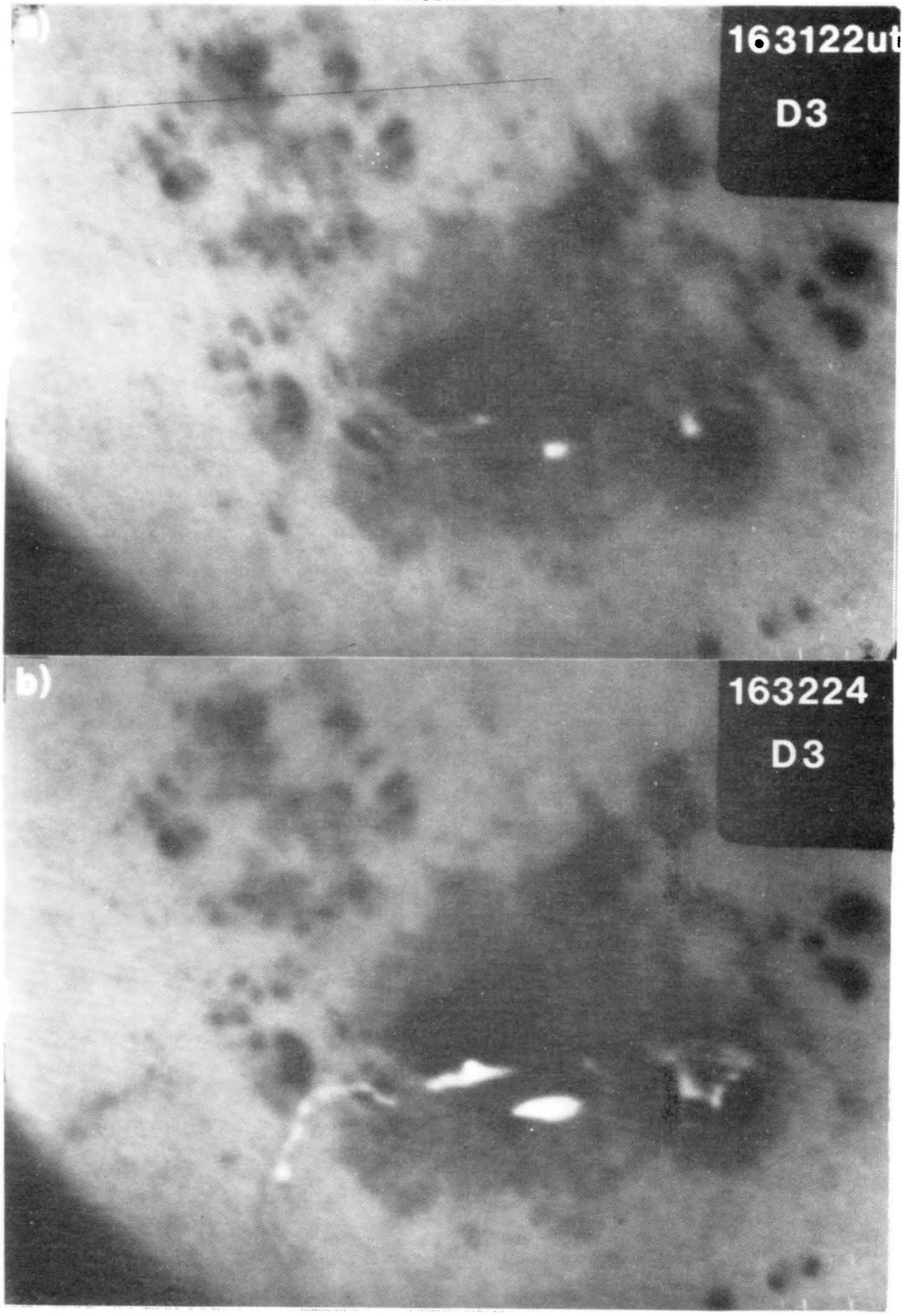

Fig. 1: D3 and Hα images of the June 6, 1982 flare. The large round spot at lower right had been directly S of the large main spot on 6/4; it travelled 30,000 km in two days.
(a) D3 16:31:22 UT -- Impulsive beginning: transient bright kernels appear in the moving spot (p polarity), the large delta spot (p below, f above) and a small f spot to the left.
(b) D3 16:32:24 UT -- Development of the main phase of the flare: although numerous transient flashes still appear, two intense kernels appear in the delta spot where the hot thermal phase is beginning. Field lines that once connected the p spot at right now connect the two intense kernels.

6 JUNE 82

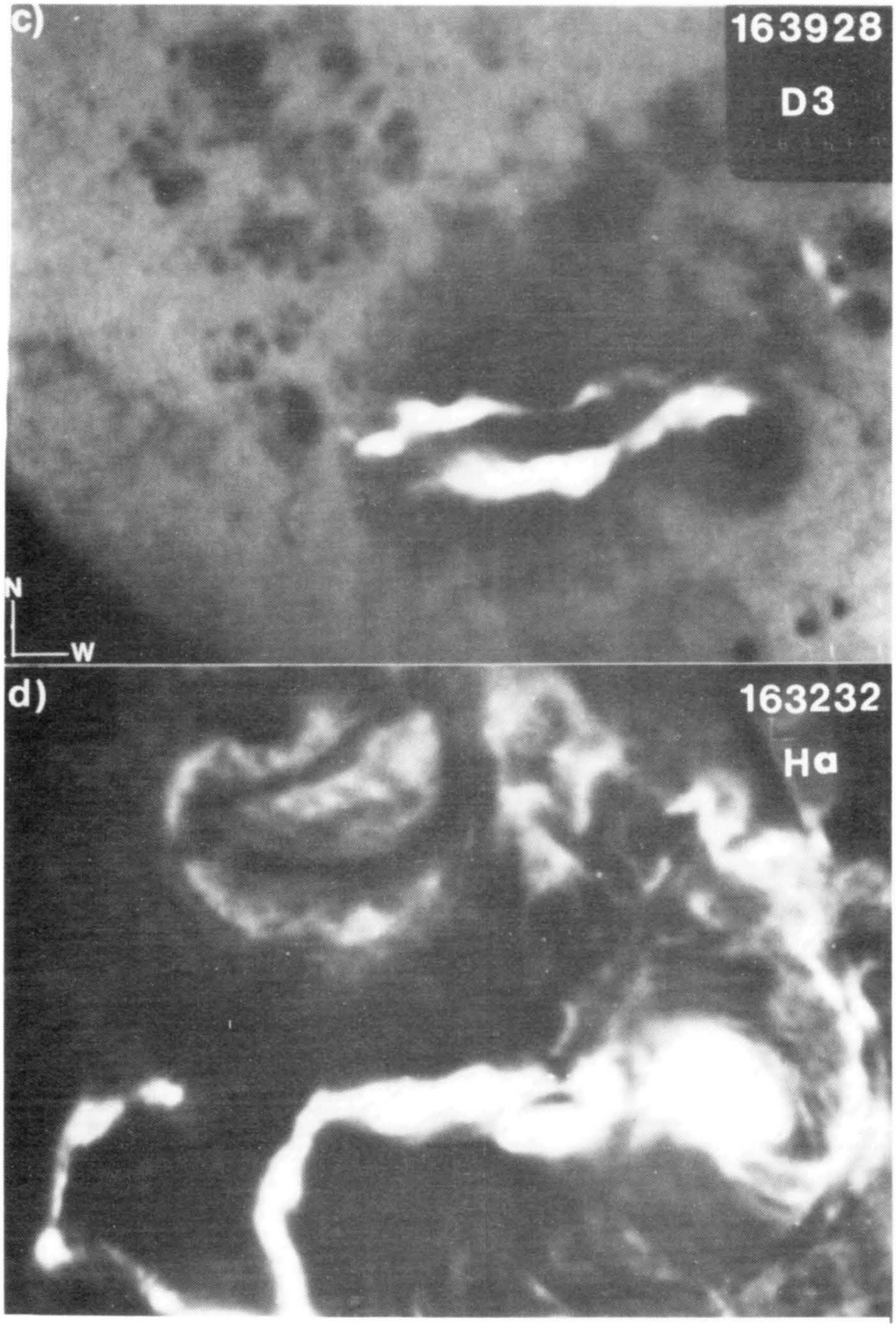

Fig. 1 - continued
(c) D3 16:39:28 UT -- Mainphase: bright loops now connect the two main kernels, which are spreading and elongated; the moving spot at right is also very intense and new emission appears at far right where an associated flare developed.
(d) Hα + 0.5 A 16:32:22 UT -- Note the intense emission over the p spot, corresponding to several small kernels in (b).

calculated. The appearance of Hα and coronal line loops high above the surface, visible on Lyot's early films, is an important fact in understanding the thermal flare process. Our flare models must produce a strong departure from hydrostatic equilibrium in the high density region at the loop tops, where the reconnection in the thermal phase appears to take place.

The two strand phenomenon is a result of the fact that all closed field lines must intersect the solar surface in two points. The rapid spread of the footpoints probably signals the rapid spread of flare energy release from its inception at a low height along the neutral line. The loops trace the field line connection between points of opposte polarity. Whether these connections existed before the flare or are new cannot presently be proven, but, at least in the 6/6/82 flare, it seems clear that the field lines previously ran between the two large spots that appeared together. After the flare the loops show them connecting the largest spots directly across the neutral line.

So much emphasis has been placed on the "two-ribbon flare" that the question is often asked "Where are the other two ribbons?" The Petschek theory, as well as other models of magnetic merging require the involvement of four magnetic poles, and energy from the reconnection area must flow down to each pole and produce surface brightening. In fact, in a complex flare in Hα there are plenty of ribbons (often more than four) because the line is easily excited by any energetic particles that get into a particular line of force; we have many examples of brightening 100,000 km away. Thus the main ribbons must be picked by using less sensitive lines, like D3 or Hα, and a reasonable magnetic connection. For example, two bright D3 ribbons connected by loops must be the two poles whose connection results from the flare; an area connected to the kernels by Hα fibrils can contain one of four ribbons. One important means of determining transverse field links is to study small flares in a region; these usually have 2 bright points which must be at the ends of a single line of force.

But the main phase of most big flares shows two long ribbons connected by an arcade of loops! If we examine the D3 pictures of the 6/6/82 flare (Figure 1), we see in the impulsive phase 11 separate kernels in four general areas, two of each polarity. The neutral line has been so stretched by spot motion that the poles are all aligned roughly along these strands. When the main phase of the flare begins, the peak intensity rapidly spreads from the kernels to the entire string of poles. The spreading apart of the two strands corresponds to the spread of the energy release from the critical neutral line region and is of course matched by the increasing height of loops above the region.

Finally, the three-ribbon flare is an event which is much more common than previously thought. An example is the flare of July 1, 1980 (Zirin and Neidig, 1981). A few minutes after the impulsive peak a third region brightens and remains bright long after the two ribbons near the neutral line fade. Feldman et al. (1982) found a number of cases where the third strand was prominent. The effect is not prominent in Hα but quite marked in D3, and even more prominent in the continuum, judging by the July 1, 1980 event.

4. OTHER LINES

Interesting perspectives on flares may be obtained in other lines, which may be observed with the universal filter or specialized filters. Hβ is particularly interesting because it is not as deep and saturated as Hα but the atmosphere phenomena are easily seen. However, exposures with the universal filter are long because the polaroids have poor transmission in this region.

It is surprising that the most intense Balmer line is Hγ, both in solar and stellar flares, i.e. there is a Balmer increment. Probably this is due to the fact that the lines are very deep (Neidig estimates τ (Hα) = 10^4) and reach the Planck limit. If the temperature where the flare emission occurs is 10 000° the Planck peak will of course be in the blue.

As noted above, D3 is a beautiful line to study flares; no chromospheric background is seen and the emission may be quite intense. With the exception of post flare loops, which are in emission only in the most intense events, all atmospheric phenomena (surges, sprays, active filaments) are in absorption. This is because D3 is seen against the 6 000 degree photosphere, and densities greater than 10^{13} are required to produce sufficient collisional excitation for the line to be seen. Feldman et al. showed that the emission in D3 appears in two ways: scattered transient brightenings along the neutral line during the impulsive phase, and intense brightening of two ribbons on either side of the neutral line during the thermal phase. This behavior allows us generally to separate impulsive and thermal phases and to identify footpoints of the core of the flare. The most intense D3 occurs during the peak of SXR emission from a 2 keV thermal plasma (Figure 1). It is intriguing to compare the transient brightenings along the neutral line with the "islands" that appear in the first phase of tearing mode instability (van Hoven, 1981); each island could generate the particles for each pair of footpoint kernels.

Other lines, such as MgI b and NaD appear virtually identical to D3 but less intense (partly because they are in absorption lines) and without absorption from surges, sprays, etc. Even weak lines such as FeI 5324 and CaI 6103 appear in emission in high resolution images, with intensities about twice the line center (i.e., about 0.2 x the photosphere) and distribution somewhat similar to D3. All these emissions appear to be footpoints, either of streaming impulsive particles or conduction from thermal loops (Figure 2).

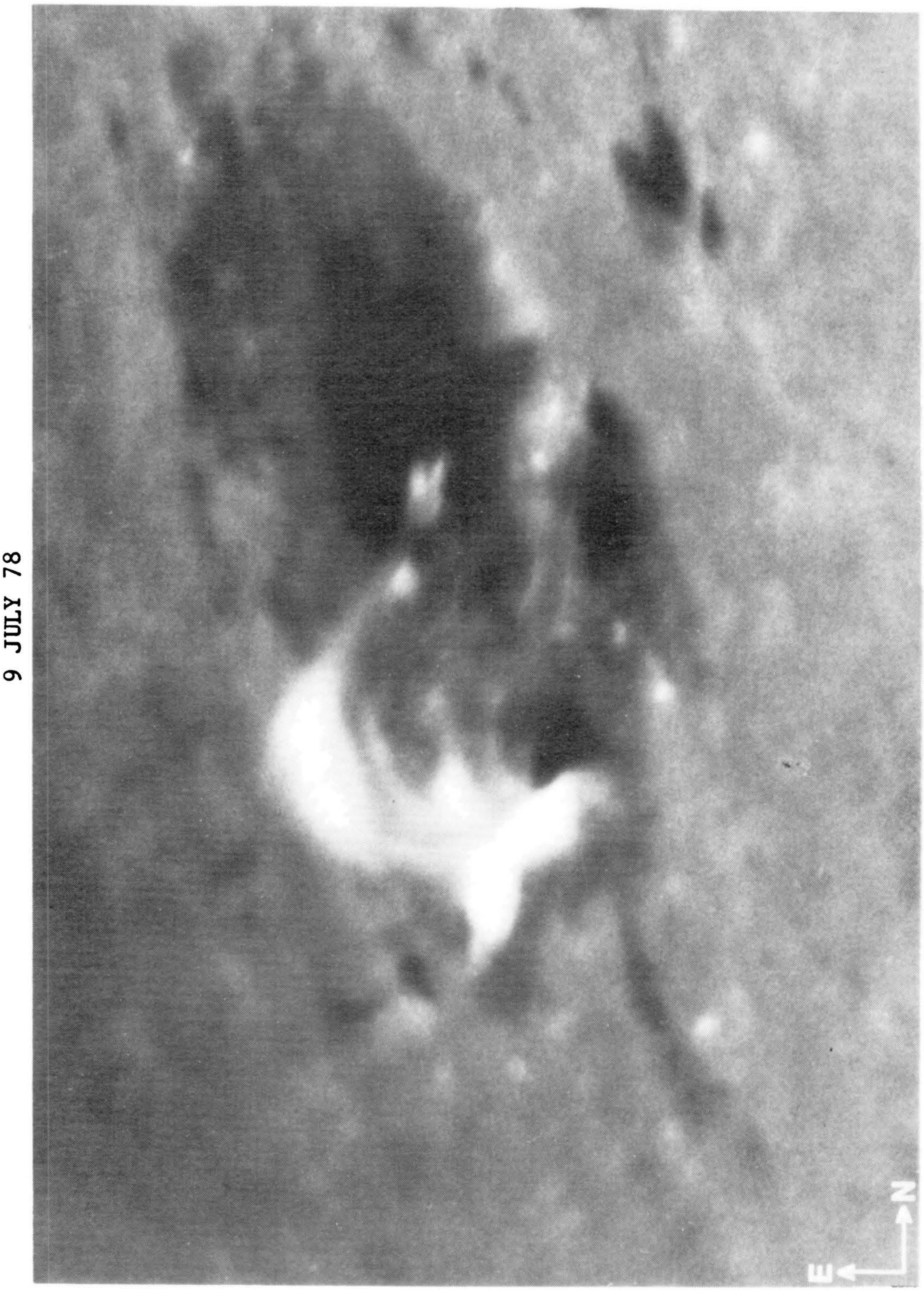

Fig. 2: Hα-1.5Å filtergram of a flare in an active region not far from the east limb. Note the loops connecting the footpoints of the flare.

5. CONTINUUM OBSERVATIONS

The continuum emission from flares is of great interest and represents one of the major energy outputs of the flare. The morphology of the continuum is similar to that described for D3, except that the impulsive kernels are relatively more intense and the emission decays more rapidly. Until recently we thought that white light flares could only occur on the disk and that they were essentially due to heating of the photosphere by streaming particles or conduction. But Harvey's observation (unpublished) of a white light flare above the limb shows that flares can emit continuum comparable in intensity to the photosphere from regions some distance above it. A simple calculation shows a hot very dense loop ($N_e = 10^{14}$, T 100,000°) might produce the observed effect. However as I noted above, most white light flares are observed at footpoints, and we must consider the general case as heating of footpoints by either streaming particles or conduction.

The problem of the blue continuum remains unresolved. A strong increase in flare continuum intensity is observed below 4000 Å in both solar and stellar flares. For example, Zirin and Neidig (1981) found no detectable increase at 5000 Å in flares where the continuum intensity doubled at 3862 Å. Neidig (1982) found I/Iphot of 3.6 at 3610 Å, 0.65 at 4275 Å and 0.4 at 6203 Å. Although Neidig (1982) has shown that some of this effect is due to the drop in the background intensity, he points out this cannot explain kernels where the intensity increases to 3 or 4 times the photosphere in H^-. Opacity at these wavelengths also seems unlikely in view of the known shape of the H^- absorption. Similarly, it is hard to explain the spectral dependence of the white light flare by hot black body emission, as was done in stellar flares by Mochnacki and Zirin (1980). Any hot black body will produce considerable emission in the green which is not observed. The effect of lines appears to me inadequate. If we filled in all the lines we would only increase the intensity at 3800 Å by 1.9 and 4300 Å by 1.5. The metallic lines are in fact narrow (about 0.1 Å wide) and it is hard to understand how they could provide enough intensity to double the continuum at 3610 while producing an only 5% increase at 4275 as occurred in the 1 July 1980 flare. Perhaps there are autoionization continua in this region.

The spectral evidence now favors strong Balmer continuum emission in flares. This conclusion is based on Neidig's data at 3610 Å (1982) and Zirin and Neidig (1981) and the work of Mochnacki and Zirin (1980) who found strong Balmer continuum in stellar flares. My experience with stellar flare spectra is that the Balmer continuum disappears if the detector has low sensitivity below 3800 Å, and I attribute earlier failure (Svestka, 1965) to detect Balmer continuum to that cause.

In summary, the different behavior of the continuum and D3 emission during impulsive and hot thermal events shows that heating at various low levels in the atmosphere is produced by both streaming and conducting electrons at footpoints. The amount transmitted is greatest in the hot thermal phase, and large increases (2 to 4 times) above the photospheric background can occur in D3 and the blue continuum.

6. THE LIMB FLARE

Most flares are seen on the disk and give useful information on the flare location relative to the magnetic fields. But since there is good evidence that the flare energy release takes place above the photosphere, it is important to study the limb flare, particularly in Hα where we have enough opacity to give hope of seeing the flare source.

What we see on the limb bears little relation to the flare on the disk, except for loops and sprays. Two common types of limb flare are (1) a bright condensation (usually the neutral line filament) appears above the limb and blows off in spray, followed by loop condensation, (2) ragged bits of emission, much brighter than the disk in Hα are seen. Except for the post-flare loops, the Hα light curve is short and relatively similar to the hard X-ray time profiles (Zirin, 1979). This is further evidence that (1) the HXR come from the Hα-active region just above the limb and (2) the main flare brightening after the impulsive burst is a low chromospheric phenomenon heated by conduction from the thermal flare.

We do not know exactly what the bits of Hα emission are, although in some cases they clearly match the erupting neutral line filament, in others that is not the case - they brighten and fade without motion and may or may not be loops. In other cases, the flare appears to be a set of low lying loops which grow upward.

7. FOOTPOINTS: A PEDESTRIAN SUBJECT

Any doubt about the connection of loops and footpoints should be removed by Figure 2, a frame in Hα-1.5 Å obtained about 20 minutes after a class 2 flare. The footpoints seen here were emitting in white light 15 minutes earlier; the impulsive flashes were at other points nearby. We clearly see how the bright loops connect the heated footpoints.

Although the role of footpoints is clear, the formation is hard to understand. All indications from optical data are that the heating at footpoints goes much deeper than current models. The He D3 emission requires $T > 15000°$ at $N \sim 10^{14}$. The white light data requires $T > 8000°$ at the photosphere. The HeII data of Linsky et al. (1976) require 85000° at a density ($\sim 10^{12}$) where 304 is optically deep. The fact that much of this heating occurs in the thermal phase is just as surprising as the instantaneous production of the little kernels in the flash phase.

8. THE RIDDLES OF FLARE SPECTRA

I would like to summarize some riddles of the optical flare:

1) The blue continuum.

2) Hγ is the strongest line of the Balmer series. This is possibly explained by black body emission in the Balmer lines at 7000°.

3) The integrated intensity of Hα equals that of Lyα. This could be explained by black body emission in the Balmer lines at 15 000°, which obviously contradicts (2). I do not accept published models which find a set of parameters that fit this observation; we require a more physical explanation of why this happens so often in nature.

4) To what feature of the disk flare does the limb flare correspond?

5) Jefferies, Smith and Smith (1959) found HeII 4686 in the spectra of disk flares. Although the brightness is not great, high loop densities (at least $10^{12.5}$) are required. The 4686 intensity measured by them exceeds the intensity of λ304 measured by Linsky et al. (1976) in a similar flare.

6) The D3 line is observed in emission at footpoints and in loops. The footpoint emission can reach 4 times the photosphere. The loops are at least 10% above the photosphere. Densities of $N_e > 10^{13}$ are required.

7) Post flare loops also appear in emission against the disk in Hα.

Where does the future lie in optical flare observations? Although we have considerable good data, we still need:

1) D3 flares near the limb.

2) Flare images in HeII 4686.

3) Polarization measurements (discussed by Dr. Henoux).

4) Flare spectra of the real kernels.

5) Definitive measurements of the Balmer continuum.

6) Better measurements of the transverse field.

7) Magnetograms in Hα.

This survey has been prepared during my stay at Meudon. I thank Dr. Henoux for his hospitality, and the Ministère d'Etat à la Recherche et à la Technologie for a grant supporting my visit. The research at Big Bear was supported by the National Science Foundation under ATM-8112698 and by NASA under NGL 05 002 034.

REFERENCES

Feldman, U., Liggett, M., and Zirin, H.: 1982, BBSO #0214, submitted to Ap.J.

Jefferies, J. J., Smith, E. V., and Smith, H. J.: 1959, Ap.J. 129, 146.

Linsky, J. L., et al.: 1976, Ap.J. 203, 509.

Mochnacki, S. W., and Zirin, H.: 1980, Ap.J. 239, L27.

Neidig, D.: 1982, submitted to Ap.J.

Severny, A. B.: 1958a, Izv. Krym. Astrofiz. Obs. 19, 71.

Svestka, Z.: 1965, Adv. Astron. Astrophys. 3, 119.

Svestka, Z.: 1976, Solar Flares, Reidel.

van Hoven, G.: 1981, in E. R. Priest (ed.) Solar Magnetohydrodynamics, Gordon and Breach, London.

Zirin, H., and Neidig, D.: 1981, Ap.J. 248, L45.

Zirin, H.: 1978, Solar Phys. 58, 95.

Zirin, H.: 1980, Ap.J. 235, 618.

DISCUSSION

HIEI: Can we identify the loop prominence systems which connect all the Hα bright regions of two-ribbon flares?

ZIRIN: In principle, yes, but complete identification is limited by signal-to-noise ratio. On Sept. 10, 1974, we see loops connecting almost everything.