An overview of the solar flare phenomenon

Facts are stubborn things

Gil Blas, A. R. Lesage (1668–1747)

The first recorded observation of the flare phenomenon was made independently by two observers, R. C. Carrington and R. Hodgson, on 1 September 1859. While engaged in the routine survey of sunspots on the solar disk, they witnessed an intense brightening of regions in a complex sunspot group; the event lasted only a few minutes. This was an example of a relatively rare event – a large white-light flare – in which the optical continuum is enhanced sufficiently over the background photospheric field to be visible in contrast, in this case with a relatively crude instrument. Most flares are not so conspicuous in visible light; they reserve their strongest enhancements for spectral lines such as Hα, and they also radiate copious amounts of energy in extreme ultraviolet (EUV) and soft X-ray wavebands. With the advent of spacecraft observations from about 1960 onward, our observational data base on solar flares has increased by orders of magnitude. Flares have been observed across the electromagnetic spectrum, from decametric radio emission to γ-rays in excess of 10 MeV. However, it is fair to say that our theoretical understanding has failed to keep pace with this explosive growth in our observational knowledge. Some of the reasons why this is necessarily true will be dealt with as they arise, and we will, where applicable, point to areas now ripe for rapid development.

1.1 History of observations

Until the time of the Orbiting Solar Observatory series of satellites and the Skylab Space Station with its Apollo Telescope Mount cluster of solar instruments, flare observations were almost uniquely carried out photographically in one of a small number of optical spectral lines, especially in the hydrogen Hα line at 6563 Å. An overwhelming amount of data on different flares therefore exists in the form of Hα only, and general conclusions as to the size, shape, lifetime, intensity, etc., of flares have historically been drawn from these data. In particular, flare classifications,

pertaining to an assumed 'importance' of the flares, have been devised. Table 1.2 below shows the dual form for importance classification adopted by the International Astronomical Union (IAU) (1966). The classification describes fairly well the low-energy, cool optical flare whose importance correlates well with many terrestrial, flare-induced effects such as certain geomagnetic storms, auroral displays, etc. On the other hand, such classifications do not, nor can they, reveal the basic physical processes responsible for the flare. Research during the last several years has shown that while the optical flare, as portrayed in Hα, is an impressive display, there is a hot, high-energy part of most flares, and the study of this part reveals much of the underlying physics of flares. Nevertheless, observations of the cool component, both its spectral signatures and its morphology, offer important clues to its physical character, as stated by Smith and Smith (1963) in their book, which for many years served as an important source of information on the 'classical' flare. While the earlier spectral observations of flares mainly utilized hydrogen lines and a few helium lines to derive physical quantities such as temperature and density, they also provided information on a large number of lines from neutral and singly ionized metals. These latter lines provided information on excitation conditions in flares and, furthermore, served to classify flares and related prominences (Waldmeier, 1951; Zirin and Tandberg-Hanssen, 1960; Tandberg-Hanssen, 1963) in systems relating to the low-energy part of flare plasmas.

Particular interest has always been attached to the observation of white-light flares [e.g., Švestka (1966*a*); McIntosh and Donnelly (1972); Uchida and Hudson (1972)]. Once considered a rare phenomenon, white-light flares have recently been studied in considerable detail and are now thought of as a fairly common occurrence during sunspot maximum years (Neidig and Cliver, 1983). The spectral signature of these flares is a continuum emission so short-lived that it is difficult to record (Grossi-Gallegos *et al.*, 1971; Rust, 1973). Easier is the observation of parts of the flare, generally small patches, seen in integrated light. The existence of the continuum emission indicates a dense flare plasma and thorough studies of this phase of larger flares provide valuable information on their nature.

In the history of morphological flare observations, two concepts which have played an important role in directing certain trains of research are *homologous* and *sympathetic* flares. With more than one active region visible on the Sun, flares may occur more or less simultaneously in different locations (Richardson, 1951) and there seems to be more of these so-called sympathetic flares than can be accounted for under the assumption of random occurrence of flares. Skylab data show that active regions are often

connected by arch or loop structures, visible in soft X-rays and outlining magnetic field configurations. These coronal magnetic loops provide channels for different types of disturbances that can thereby travel from one active-region flare site to another. The concept of sympathetic flares thus leads us to consider the role played by magnetic fields. A similar emphasis on the importance of magnetic fields in the flare phenomenon is provided by the notion of homologous flares. It is often observed that a flare is brightening in exactly the same position and exhibiting the same geometrical outline as a previous flare in that region. This repetitive character was first discussed by Waldmeier (1938*b*). Such flares have been studied extensively, based on the point of view that the form of the magnetic field configuration is a determining factor in the processes that lead to a flare, and that there is a rebuilding of the stressed magnetic field after each successive flare. We are therefore again led to consider the role played by magnetic fields in flares.

It has been known for a long time that nearly all flares occur in active regions with sunspots, and the more magnetically complex the sunspot group, the higher the frequency of flare occurrence [e.g., Bell and Glazer (1959); Dodson-Prince and Hedeman (1970)]. This observation clearly points to the magnetic field as an important, maybe crucial, ingredient in the flare process. However, early observations showed that 'bigger is not better' in the sense that the flares are normally not found above the spot umbrae where the magnetic field is strongest (Švestka *et al.*, 1961). Evidently, aspects other than strength of the magnetic field configuration are important. On the other hand, when a flare occurs over a spot, new signatures appear as the flare plasma becomes the generator of radio microwave radiation and X-rays (Dodson-Prince and Hedeman, 1960; Ellison *et al.*, 1961; Martres and Pick, 1962). These early observations point to the hot, high-energy part of flares, the study of which has become increasingly important for our overall understanding of flares.

With the advent of solar magnetographs, direct comparison of flare position with pre-existing magnetic structures became possible (Bumba, 1958; Severny, 1958). Of particular interest is the location relative to the so-called 'magnetic neutral line' – the locus of points with zero longitudinal component of the magnetic field – a situation we shall further explore later. Martres *et al.* (1966) investigated this behavior in relation to the occurrence of small bright points of emission where the flare first occurs. As we shall see later, the study of these bright points has led to additional information on the flare phenomenon.

While most early magnetic field observations were restricted to measuring the longitudinal component, sporadic information was also gathered relating

relating to the transverse component [e.g., Zvereva and Severny (1970)]. During the big flare of August 1972, Zirin and Tanaka (1973) and Tanaka and Nakagawa (1973) discussed the importance of the observed sheared structure of the magnetic field in the region where the flare occurred. We shall see later that certain plasma motions that lead to sheared magnetic field structure do indeed seem to play a major role in flare production.

In the 1960s, our knowledge of solar flares increased dramatically with the Orbiting Solar Observatory (OSO) series of satellites which allowed us for the first time to study in detail the characteristics of flares at wavelengths inaccessible to ground-based observatories. These early spacecraft observations were soon supplemented by a host of others, such as from the European Space Research Organization TD-1A in 1972, NASA's Skylab Apollo Telescope Mount (ATM) in 1973–74 and, more recently, from the International Sun–Earth Explorer 3 (ISEE 3) and the US Air Force P78-1 satellite. Of particular interest are the data from NASA's Solar Maximum Mission (SMM) and the Japanese Hinotori spacecraft, with their emphasis on the high-energy aspects of solar flares. It is inappropriate here to discuss the vast amount of observational facts gathered from these spacecraft observations or from the host of rocket and balloon flights launched during the same period; we will have ample opportunity to discuss these observations throughout the course of the book.

We close this brief discussion by showing in Figure 1.1 the early time history of a fairly typical solar flare in various wavebands. We note that the manifestations of a flare are indeed complex and that, although there is some agreement between the light curves at different wavelengths, these manifestations are sufficiently diverse that no single one can fully describe the evolution of the flare. Added to this the fact that no spatial, polarization, or directionality information at any wavelength is presented in the diagram and that only a crude representation of a solar flare spectrum can be ascertained from it, one gains an idea of the degree of complexity of a solar flare and the potential wealth of information that we can extract from it.

1.2 Classification of flares

As with most complex phenomena, scientists have tried to devise schemes to classify flares and make some order of their seemingly chaotic and bewildering manifestations. The hope was that by systematically classifying certain observables, one would be able to gain some insight into the physics of the phenomenon.

1.2.1 *X-ray classification*

While all flare classifications prior to the 1960s have relied on observations in the visible part of the spectrum, more recent data suggest that the X-ray signature of flares may give as good, or better, criteria for classification purposes, so that we can deal with better coverage and rely on quantitative, more objective data. In addition, the X-ray criteria may simultaneously provide a more profound physical insight into the flare phenomenon itself.

Probably the simplest classification used is based on the global output of soft X-ray photons during a flare. This soft X-ray classification utilizes the flux in the 1–8 Å range of the spectrum, and, depending on the measured X-ray flux, one classifies a flare as a C, M, or X flare according to the scheme in Table 1.1. The classification letters represent the order of magnitude of the X-ray flux, and a subsequent numerical value indicates the multiple of the order of magnitude (e.g., M3 = 3×10^{-2} erg cm^{-2} s^{-1}). To accommodate flares smaller than class C, referred to as subflares, a class B has been added.

Fig. 1.1. Time profiles of the 22 GHz, hard X-ray, O V (1371 Å) and Fe XXI (1354 Å) emissions through the impulsive phase of a flare on 1 November 1980 (from Tandberg-Hanssen *et al.*, 1984).

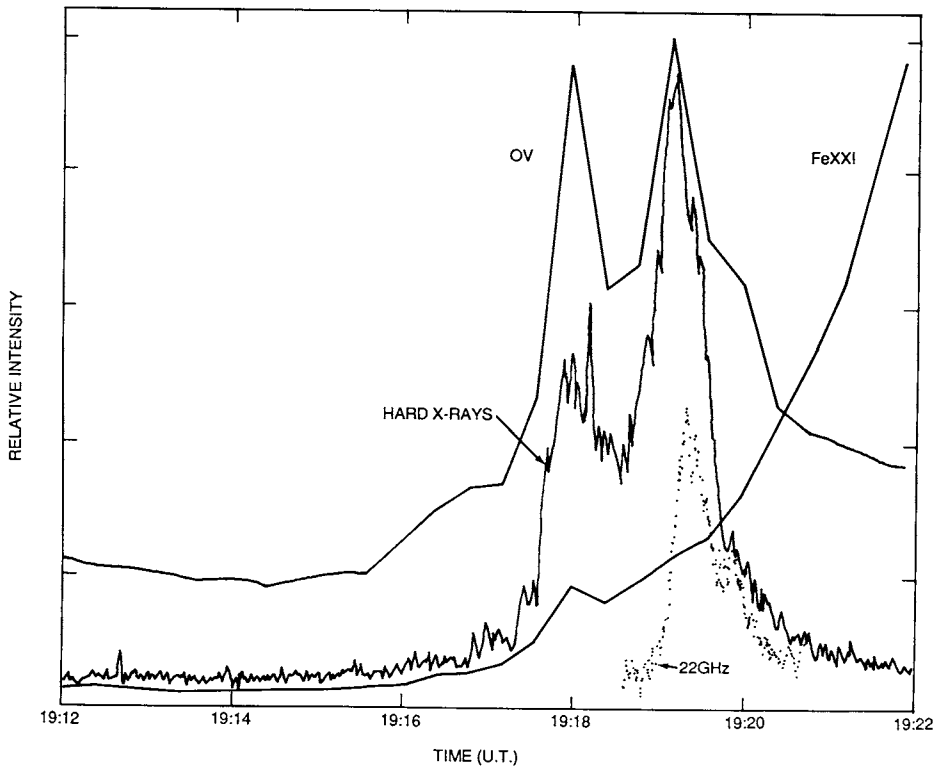


Table 1.1. *Soft X-ray flare classification*

Class	Intensity (erg $cm^{-2} s^{-1}$)
B	10^{-4}
C	10^{-3}
M	10^{-2}
X	10^{-1}

Measurable X-ray levels below 10^{-4} erg $cm^{-2} s^{-1}$ are indicated as decimal values of B (e.g., B.7 = 7×10^{-5} erg $cm^{-2} s^{-1}$).

While class C flares are a common occurrence during years near sunspot maximum, the frequency of X flares is always low. As an example, statistics (Smith, Jr, 1986) for the maximum years 1979–80 show that 2834 class C flares were observed, while the numbers for classes M and X were 554 and 59 respectively. For the minimum year 1976, only 75 class C flares were reported versus five class M flares. There was no class X flare. For the whole period 1976–85, there were reports of 17 986 flares, of which 74% were class C and 1% class X.

For most flares, a good correlation exists between the X-ray classification of Table 1.1 and the time-honored visual classification of Table 1.2, which we will discuss below.[1] Consequently, for these flares, not much is gained by changing (or adding to) the familiar visual classification scheme to the B, C, M, X notation of Table 1.1. However, some flares rate as very important in the X-ray domain, while their visual output would classify them as minor events – and vice versa for other flares. In these cases, we have added information that can tell us something about the nature of the flare in question, e.g., regarding the temperature of the radiating plasma, the location of the flare site, etc.

Furthermore, the X-ray signature can be used to describe a flare beyond giving its X-ray flux as indicated in Table 1.1. Several attempts have been made to use characteristics of the X-ray emission to classify flares in ways that reveal certain aspects of the flare physics. Time profiles of the flare emission at several frequencies (energies) enable us to divide flares into impulsive or gradual events and into flares with hard and/or soft X-ray spectra. From Skylab data, Pallavicini *et al.* (1977) studied soft X-ray flares seen at the limb and classified them into (*a*) compact flare loops, (*b*) point-

[1] This is not too surprising, and is an example of what has come to be known as the 'Big Flare Syndrome', which simply states that the bigger the flare, the brighter it is at all wavelengths.

like flares, and (*c*) large, diffuse flare loop systems. We shall return to some of these characteristics in Chapter 5.

Hard X-ray flares have been classified by Tanaka (1983) [see also Ohki *et al.* (1983) and Tsuneta (1983)] into three categories, viz.

Type A: thermal flares, in which a very hot ($T \approx 30$–50×10^6 K) plasma produces a smoothly varying thermal X-ray emission; these flares are small (< 5000 km) and compact and occur at low altitudes;

Type B: impulsive flares whose plasma, which is confined in long ($> 10^4$ km) sheared loops, produces impulsive spikes or bursts in the X-ray time profile;

Type C: gradual flares, whose plasma produces smoother, long-duration X-ray time profiles; these flares occur at high altitudes, $\sim 5 \times 10^4$ km, and the X-ray spectrum hardens with time.

Types B and C are generally considered nonthermal X-ray sources. We will return to this important point in Chapters 7 and 8.

1.2.2 *Visible light classifications*

A flare on the solar disk is observed as a temporary emission within some dark Fraunhofer line. The spectral line most frequently used is Hα, and in great flares the emission excess inside Hα may be several times the intensity of the adjacent continuum. On pictures of the disk, e.g., on a spectroheliogram or a Lyot filtergram, a flare is observed as a brightening of parts of the solar surface. The area covered by a great flare may be several times 10^9 km^2. Flares smaller than about 3×10^8 km^2 have generally been referred to as *subflares*.

It is this area that serves as the primary basis for visual light classifications. It is measured as the projected area in terms of the unit 1 square degree heliographic at the center of the disk, and should refer to the time of maximum brightness of the flare:

$$1 \text{ degree heliographic} = \tfrac{1}{360} \text{ solar circumference}$$
$$= 12\,500 \text{ km}.$$

One normally corrects the area for foreshortening and expresses it in millionths of the disk:

$$1 \text{ millionth of disk} = \tfrac{1}{97}\ (= 0.0103) \text{ square degree},$$
$$1 \text{ square degree} = 1.476 \times 10^8 \text{ km}^2 \text{ of solar surface}.$$

The above classification has the obvious drawback that it ignores the brightness of the flare. A unit area of flare emission is given the same weight whether it emits copious amounts of photons (i.e., is bright) or is a poor

Table 1.2. Hα *flare classification*

Corrected area		Relative intensity evaluation		
In square degrees	In millionths of hemisphere	Faint (f)	Normal (n)	Brilliant (b)
< 2.06	< 100	S f	S n	S b
2.06–5.15	100–250	1 f	1 n	1 b
5.15–12.4	250–600	2 f	2 n	2 b
12.4–24.7	600–1200	3 f	3 n	3 b
> 24.7	> 1200	4 f	4 n	4 b

emitter (i.e., is faint). To remedy this, a dual form for importance classification has been adopted (IAU, 1966). This importance evaluation consists of two elements, a number and a letter. The number describes the size of the area. The letter indicates whether the intensity of the flare area is faint (f), normal (n), or brilliant (b), and considerable subjectivity is associated with this part of the classification. The dual form for the importance classification is given in Table 1.2.

Radio observations have greatly added to our information of the flare phenomenon, and routine monitoring records flare emission at a number of frequencies. In general, the stronger the optical or X-ray emission from a flare, the stronger the radio flux, S. The latter is given in solar flux units, sfu, where $1\,\text{sfu} = 10^{-22}\,\text{Wm}^{-2} = 10^4$ Jansky. Using such flux measurements, we can assign a certain flux to the importance of a flare. As a rule of thumb, we may say that at 5 GHz, which corresponds to a radio wavelength of 6 cm, the microwave flux, S, from an optical Hα subflare is approximately 5 sfu, increasing to 30 sfu for a class 1, 300 sfu for a class 2, 3000 sfu for a class 3, and 30 000 sfu for a class 4. An approximate relation between importance class (Hα) and radio flux at 6 cm, expressed in sfu, is therefore

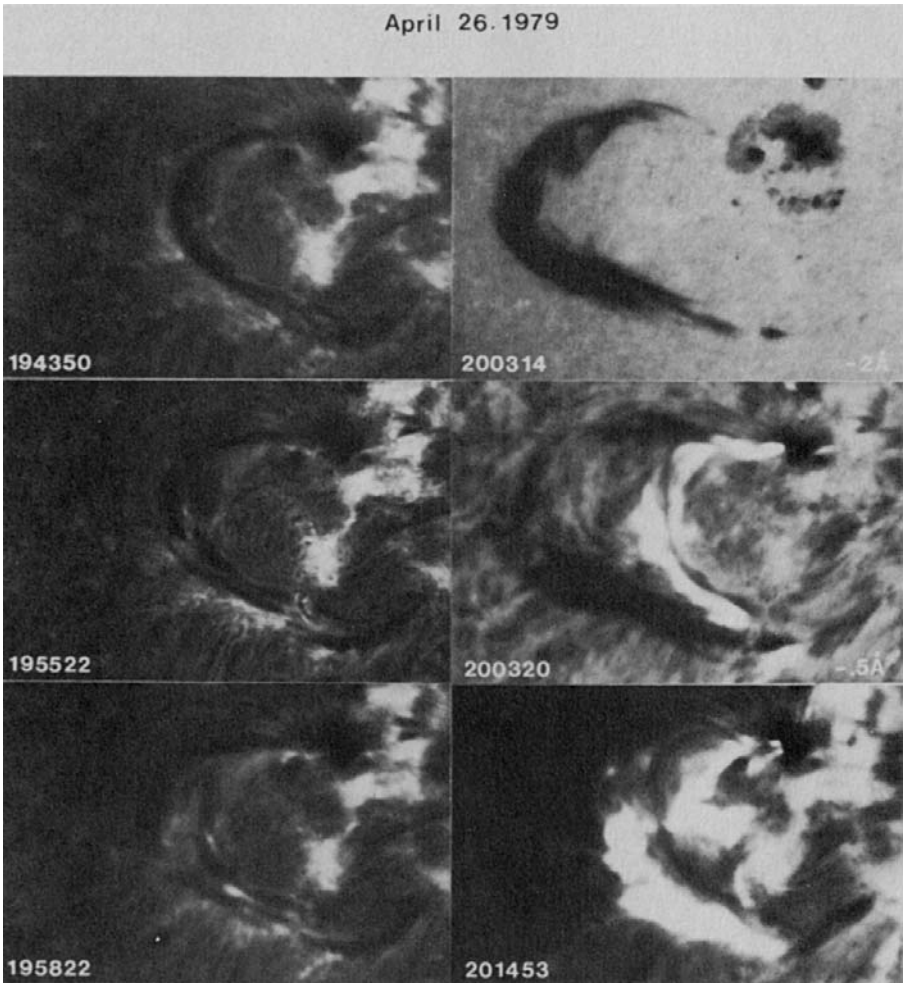
$$\text{Imp} = \log[S(\text{sfu})] - 0.5.$$

1.3 What does a flare look like?

It will be helpful for later discussions on flares to have in mind certain facts about the flare emission as observed in various wavelength regions. For a succinct overview see Bhatnagar (1986). Flares are roughly – and for practical purposes – divided into two groups (Pallavicini *et al.*, 1977), viz. small, compact flares and large, two-ribbon flares. However, we shall see later that it is not the distinction between compact, in the sense of small, versus two ribbons that is important from a physical point of view,

but rather whether the flare is confined or not (Švestka, 1986). A compact flare probably takes place in a small loop in the lower corona, and the flare emission is largely confined to the plasma in the loop where it eventually dies away. The triggering, on the other hand, seems to be similar to the triggering of large flares, namely due to a large-scale eruption and reconnection of sheared magnetic fields (Sturrock *et al.*, 1984; Machado *et al.*, 1988*a*). By contrast, a two-ribbon flare is always associated with an erupting prominence, and the flare emission occurs in an arcade of 'post-flare loops' along the prominence with the individual loops oriented more or less at

Fig. 1.2. Development of a two-ribbon flare on 26 April 1979 following ejection of prominence material. Mhe panels at 20:03:14 and 20:03:20 UT show pictures taken in the blue wing of Hα where the ejected material is more visible (from Tang, 1986).



right-angles to the long axis of the prominence. Emission also occurs at the feet of the loops, thereby forming two ribbons on either side of the prominence. These ribbons of emission move apart as the flare progresses with larger and higher loops continuing to bridge them. The erupting prominence may eventually be confined or break away. In the latter case, we witness a coronal mass ejection (see Chapter 9). Figure 1.2 shows the development of a two-ribbon flare recorded in Hα. In this case, the field configuration allowed prominence material to escape to the outer corona, and a well-observed mass ejection resulted (Moore and LaBonte, 1980). Švestka (1986) refers to this class of flares as *dynamic*; see also Machado *et al.*, (1988*a*). During the development of two-ribbon flares the post-flare loops, when seen as loop prominences above the solar limb, exhibit material falling down from the loop apex along their legs and into the chromosphere. These post-flare loops are among the most distinguished features of large loop manifestations (see Figure 1.3(*a*)) and show better than nearly anything else the decisive influence of magnetic fields on mass motions in the solar atmosphere. Seen on the disk the loops – which then show up in absorption – again provide an interesting display as they connect the two ribbons of the flare emission (see Figure 1.3(*b*)).

Sometimes it is difficult to identify the two (or more) ribbons of a large flare, and at other times the ribbons exhibit exotic forms. One should keep this in mind when trying to model solar flares – often we make such simplified assumptions as to their morphology that we may invalidate the approach. An interesting and curious example of such a two-ribbon flare is shown in Figure 1.4 (Tang, 1985), where a loop-shaped, nearly circular prominence occupied the inversion line of magnetic polarity. One flare 'ribbon' filled part of the interior dominated by one polarity, while the other ribbon formed a bright band outside the filament in a region of the opposite polarity.

The smallest flares observed in Hα appear as unresolved bright blobs. Since larger flares that can be resolved generally show two footpoints (or two ribbons), it is reasonable to speculate that even the very smallest flares also have two footpoints, if they are flares in very small magnetic loops. Tang (1985) has been able to observe the footpoints in such small flares in the wings of Hα, where two distinct flare kernels can be detected, while observations in the center of Hα only reveal an unresolved bright blob, obscuring the footpoints; see Figure 1.5. The reason for this difference is that in the wings of Hα where the absorption coefficient is small (Section 2.1), we see down to the cooler chromospheric parts of the flare loop; whereas, in the center of Hα, we observe its higher-lying, hotter, and more disturbed parts.