CHAPTER 3: IMPULSIVE PHASE TRANSPORT

Richard C. Canfield, University of California, San Diego
Francoise Bely-Dubau, Nice Observatory
John C. Brown, University of Glasgow
George A. Dulk, University of Colorado
A. Gordon Emslie, University of Alabama in Huntsville
Shinzo Enome, Nagoya University
Alan H. Gabriel, Rutherford Appleton Laboratory
Mukul R. Kundu, University of Maryland
Donald Melrose, University of Sydney
Donald F. Neidig, Air Force Geophysics Laboratory
K. Ohki, Tokyo Astronomical Observatory
Vahe Petrosian, Stanford University
Arthur Poland, Goddard Space Flight Center
Erich Rieger, Max-Planck-Institut fur Physik und Astrophysik
Katsuo Tanaka, Tokyo Astronomical Observatory
Harold Zirin, California Institute of Technology

3.1 INTRODUCTION

3.1.1 Motivation for Transport Studies

In the astrophysics community, 'the solar flare problem' is generally considered to be how to accumulate sufficient magnetic energy in one active region and to subsequently release it on a sufficiently short time scale. Satisfactory solution of the solar flare problem will require at least two achievements by the solar physics community: first, convincing theoretical demonstration that one or more mechanisms of energy storage and release can occur; second, convincing observational demonstration that one (or more) of these theoretical processes actually does occur in the solar atmosphere. The contents of this Chapter essentially relate to the second problem, being largely concerned with how the energy released from magnetic field is transported through the solar atmosphere before escaping in the form of the radiant and mechanical energy signatures which we must interpret.

A central point of general agreement concerning all mechanisms suggested for dissipation of magnetic energy in flares is that the actual sites of reconnection must involve scale lengths well below currently, or forseeably, achievable spatial resolution. Consequently, observational evidence in support of a flare theory is necessarily indirect, in the sense of not involving measurement of plasma parameters in the primary dissipation regions. Such indirect evidence may be of several kinds. Firstly, circumstantial evidence may be obtained by spatial resolution of the geometry, on a larger scale, of the magnetic environment in which the mechanism operates. This may permit distinction between such options as emerging flux models and twisted arch models or between mechanisms driven by currents parallel to the magnetic field, as opposed to perpendicular. Spatial resolution also permits mapping of the paths of flare products. Secondly, temporal evidence can be obtained by use of high time resolution to set limits on instability growth rates, to imply the occurrence of repetitive or multiple dissipation, to indicate the production sequence of the various flare manifestations and, by causality arguments, to set an upper limit to the size of the primary dissipation site. Thirdly, thermodynamic evidence on the nature of primary dissipation is obtainable from the distribution, particularly in the impulsive phase, of the flare energy release over its various modes, i.e. fast particles, conduction etc. For example the Petscheck mechanism releases a major fraction of the magnetic energy directly into bulk mass motion while the tearing mode initially results chiefly in plasma heating and particle acceleration.

Transport of energy away from the primary sites, and its ultimate thermalization, depend not only on the primary mechanism itself but also on the larger-scale structure of the active-region atmosphere in which the transport occurs. Consequently the study of energy transport as a diagnostic of flare mechanisms involves extensive theoretical modeling of the transport processes, as well as observational input, to provide the framework for interpreting the observations. Use of such transport studies to infer properties of the initiating disturbance is essentially an inverse problem, and so carries the danger of indeterminacy through mathematical ill-posedness. For example, the thermal structure of a conductively evolving atmosphere rapidly becomes only very weakly dependent on the heating function which initiated it, and so is a poor signature of this function. In such situations, the best strategy is to utilize jointly as many as possible independent signatures of the process to minimize indeterminacy. It is just such a combination of independent signatures of the flare process (much more compelling when taken together) that the coordinated observational approach of SMM has rendered possible.

In addition to these flare-oriented objectives, of course, the study of flare energy transport has contributions to make to the broader field of transport studies per se, such as in plasma and atomic physics.

3.1.2 Historical Perspective

Energy transport studies have become an increasingly prevalent means of investigating solar flares since around 1970 with the accompanying steady improvement in observational coverage and resolution in space, time, and spectrum, from the start of the Orbiting Solar Observatory (OSO) period onward. By the time of the Skylab Apollo Telescope Mount (ATM), considerable progress had been made toward obtaining the instrumentation needed for acquisition of high resolution data over as wide as possible a variety of
wavelengths from optical to $\gamma$-rays. Thus the ATM package achieved spatial resolution of the order of arc seconds in the soft X-ray and ultraviolet ranges together with extensive UV line spectroscopy. Contemporarily though not simultaneously, other satellites (notably ESRO TD1A, OSO-7 and the Intercosmos series) were improving the quality of hard X-ray measurements and extending the spectral range upward in energy resulting in the detection of solar $\gamma$-ray lines. Though ATM itself was never designed as a flare mission it nevertheless made major contributions to progress in flare observations, in addition to its pioneering discoveries in relation to solar coronal structures and active regions (cf. Zirker 1977, Orrall 1981). In particular, ATM established the importance of loop structure in the magnetic configuration of many flares, and the compactness of bright XUV flare kernels. In addition, the use of ATM for flare studies (Sturrock 1980) delineated the limitations of such a package for answering some of the key questions concerning flares, and thereby provided important guidelines for the planning of subsequent missions, especially SMM.

The most important obstacles to progress before the launch of SMM were the lack of data of sufficiently high time resolution, the lack of data sufficiently early in (or prior to) the flare, and the lack of data coordination over a wide enough spectral range (Brown and Smith, 1980). Examining ATM in the light of these obstacles, we can see with hindsight that ATM did not respond sufficiently quickly to enable systematic studies of flare onset or impulsive phases. During ATM much information on flare spatial structure was recorded photographically in most characteristically ‘thermal’ wavebands — from optical to soft X-rays — with resulting limitations on time resolution, on simultaneous spectral coverage, and on calibration. Typically ‘non-thermal’ emissions such as hard X-rays, $\gamma$-rays, and microwaves were recorded with comparatively low sensitivity, and little or no coordination with the ATM experiments. A typical consequence of these problems was that testing of electron heated models of flare atmospheres involved construction of a spectroscopic model atmosphere from data obtained from a variety of places in a variety of flares, and use of hard X-ray data, devoid of spatial information, from yet different flares.

By contrast, SMM formed a coordinated package of instruments dedicated to study impulsive phase phenomena at high time resolution using pre-planned targets and automated response to flare onset triggers. In addition, the spacecraft carried short wavelength instruments of unprecedented sensitivity and time resolution (the Gamma-Ray Spectrometer, GRS, Forrest et al., 1980, and the Hard X-ray Burst Spectrometer, HXRIs, Orwig et al., 1980) and spatial resolution (the Hard X-ray Imaging Spectrometer, HXIS, van Beek et al., 1980) and the facility for high resolution atomic X-ray line Spectroscopy (the Soft X-ray Polychromator, XRP, Acton et al., 1980), with digital data recording, as well as the Ultraviolet Spectrometer and Polarimeter (UVSP, Woodgate et al., 1980). Furthermore, the package was supported by a wide range of other spaceborne instrumentation and an international network of ground based observations coordinated through the SMY (Svestka, Rust, and Dryer 1982). Ground support included rapid arc-second resolution in both microwaves, by the VLA, and in spectrally resolved optical lines by the Sac Peak Vacuum Tower. Hard X- and $\gamma$-rays were still observed without spatial resolution but with such sensitivity as to permit close temporal correlation with features in the longer wavelength images, as well as $\gamma$-ray nuclear abundance spectrometry with some simultaneous millimetric coverage at ultra high time resolution by Itapetinga. On the debit side, SMM was limited in soft X-ray spatial resolution compared to ATM, and in the restricted range of UV spectral coverage, of particular importance in modeling transition-region lines.

A significant contribution to our knowledge of impulsive-phase transport has come about as a result of the post-SMM launch of the Japanese Hinotori spacecraft, whose flare instruments (Solar Gamma-Ray Detector, SGR, Hard X-ray Monitors, HXM and FLM, Imaging X-ray Telescope, SXT, and Soft X-ray Crystal Spectrometer, SOX) are described by Kondo (1982). Early Hinotori results have been described in the Hinotori Symposium on Solar Flares (Tanaka et al., 1982) and Recent Advances in the Understanding of Solar Flares (Kane et al., 1983); later Hinotori results play an important role in this chapter.

In addition to the observational requirements already mentioned, there is obviously a need for improved theoretical modeling, particularly in the direction of making predictions which would be testable in terms of realizable data (Brown and Smith, 1980). The post-ATM period has indeed seen a major increase in the amount and sophistication of flare modeling work, particularly in respect of relaxation of earlier simplifying assumptions (such as hydrostatic equilibrium and optical thinness) in describing the atmospheric response, and of basic electrodynamic and plasma collective effects on the transport of charged beams.

The perceptive reader will find that many of the questions posed in the Skylab era have been answered in this chapter. Kane et al. (1980), in the impulsive phase chapter of Solar Flares: A Monograph from Skylab Solar Workshop II (Sturrock, 1980), posed three key questions:

1. Is the distribution of energetic electrons thermal or nonthermal?
2. Do the energetic particles (electrons), produced during the impulsive phase, provide the energy for the whole flare?
3. Among the models of the impulsive phase suggested so far, which ones are most consistent with observations?

The answers to these questions are fundamental to our understanding of space plasmas; if the models that are best supported by the observations require particle acceleration efficiencies $\geq 0.1\%$ (Hudson, 1979), a substantial challenge is presented to the solar flare theorist.
Our work answers primarily the first and third questions, while the second motivated the Solar Flare Energetics group. Certainly there are still major gaps in our theoretical understanding of how energy is propagated by electron beams, as well as our observational understanding of spatial, temporal and spectral scales. However, one impulsive phase transport model now stands out above all others: the nonthermal-electron thick-target model, in which the dominant role in the transport of energy on impulsive-phase timescales (usually \( \leq \) tens of seconds) is played by beams of electrons, mostly in the deka-keV range, whose velocity distribution function cannot be described by a single-temperature Maxwellian. These electrons are guided along a loop-like magnetic field structure, from an acceleration site in the corona; they heat and cause both thermal and nonthermal emission as they are fully thermalized in the loop plasma, a thick target. The reader should not get the impression that this model describes all flare energy transport, particularly on longer timescales, or that the nonthermal-electron thick-target model passes all the observational or theoretical tests we impose upon it below. Surely there is ample evidence for the impulsive-phase existence of nonthermal particles other than electrons, for example, and for thermal domination of later flare phases. However, our preoccupation with the nonthermal-electron thick-target model in this chapter is testimony to our finding that, of the models available, it does the best job of explaining the wide variety of impulsive-phase observations we have studied.

### 3.1.3 Overview of the Chapter

Some of the most striking recent results have come from observations in the highest energy ranges. Interpretation of SMM \( \gamma \)-ray line data has established the presence of protons \( (E \geq 10 \text{ MeV}) \) in regions of density \( \geq 10^{13} \text{ cm}^{-3} \), while the ISEE-3/PVO occultation data demonstrated that electrons of \( E \geq 150 \text{ keV} \) are stopped deep in the chromosphere. Both these results suggest a thick target beam interpretation, in which the radiation is generated in the course of fully stopping the particles, in the lower solar atmosphere. On the other hand, ISEE-3/PVO data surprisingly show very little directivity at 350 keV, contrary to purely collisional transport models of the electron beam. Most notable of all high energy data is the striking demonstration, from limb brightening studies, that directivity is present in the continuum around 10 MeV.

The first-ever images in the deka-keV range (10 – 100 keV) stimulate much of the work of this Chapter. While there remains considerable debate over their implications for theoretical models (in particular beams), the presence in some flares of impulsive hard X-ray footpoints, coincident in time and space with chromospheric emissions, has been clearly established by SMM and Hinotori. Hinotori results have also been used to suggest a provisional classification of hard X-ray flares, of which footpoint events are only one class. Simultaneous microwave data have permitted comparative morphology studies of hard X-rays and microwaves with resulting constraints on the electron and magnetic field distributions. Computations of hard X-ray polarization incorporating the effects of magnetic field curvature show that even the most recent low polarization results from Shuttle experiments are not incompatible with a beam model.

While the greatly improved spatial and temporal resolution and coordination of SMM data in the hard X-ray and EUV further support the view that these radiations have closely related origins, the relationship between the fluxes in these bands remains a theoretical enigma in terms of energy balance.

The inclusion of high resolution X-ray spectroscopy allowed detailed diagnostics of the hot plasma during the onset of the thermal phase, using recent developments in the atomic physics of dielectronic satellite spectra. Line profiles and shifts enabled the determination of turbulent energy content and upward velocities during the impulsive phase, resulting from the chromospheric evaporation process. Observations of X-ray and K\( \alpha \) line excitation, together with data on the soft X-ray plasma as a whole and on the hard X-ray source electrons, have demonstrated that electron beam excitation is not essential to explain the K\( \alpha \) observations, though not precluded by it.

Improved calculations of the production of the nonthermal red shifted L\( \alpha \) line by capture processes on a descending proton beam have placed severe constraints on the flux of such beams and hence on flare models where protons play a central role energetically and in which there has been renewed interest recently.

At the time of ATM, white light flares were still regarded as a rare phenomenon suggestive of an exotic explanation. Studies with improved observational methods have shown that they are in fact of common occurrence. Coordinated temporal, spatial, and spectral information shows that the impulsive phase white light emission is spectrally compatible with a chromospheric origin, and temporally and energetically compatible with a thick-target electron beam as the power supply.

Interpretation of H\( \alpha \) profiles represents a particularly good example of the progress SMM achieved toward the ideals of rapid response, data coordination, and detailed modeling. Full radiative transfer models of the flaring chromosphere have permitted comparison on a pixel by pixel, instant by instant basis of H\( \alpha \) profiles observed at Sacramento Peak with those predicted according to chromospheric evaporation models driven respectively by thick target electron beams (observed by HXIS) and by thermal conduction. Results provide direct evidence in support of electron beam heating in some bright H\( \alpha \) kernels.

Finally, prior to the SMM era decimetric radio waves were considered to be a plasma diagnostic emission of little importance energetically. At this time serious consideration
is being given to their role in heating both the corona (by reabsorption) and the chromosphere (by electron precipitation) in flares where conditions favor the onset of coherent amplification of these radio waves. The full ramifications of this process for flare energy transport are only now beginning to be appreciated.

3.2 IMPULSIVE PHASE OBSERVATIONS AND THEIR INTERPRETATION

3.2.1 Gamma-Ray Emission Above 10 MeV.

Photon emission above 10 MeV during solar flares was observed by the Gamma Ray Spectrometer (GRS) on board SMM. As of February 1984 the highest energy photons detected, energy ~ 80 MeV, came from an intense flare on June 3, 1982. In this contribution we discuss the timing between hard X-rays and gamma rays and present evidence for directivity of the highly energetic particles that give rise to the emission above 10 MeV.

3.2.1.1 Relative Hard X-ray and Gamma-Ray Timing

As a typical example we show in Figure 3.1 the time history of the flare of June 15, 1982 in different energy bands. The event has a simple time structure consisting mainly of one impulsive burst. The peak of the emission is simultaneous within ± 2 sec over more than three decades of energy from 30 keV - 50 MeV. The hard X-ray flux shown in the upper two panels is assumed to originate from electron bremsstrahlung. The emission from 4.1 – 6.4 MeV (the nuclear energy band) is from nuclear lines and from bremsstrahlung of relativistic electrons (Forrest 1983). At energies greater than 10 MeV the gamma rays are expected to be produced by bremsstrahlung of very highly energetic electrons and by the decay of pions (π⁻ and π⁺), because the contribution of nuclear lines above 8 MeV is negligible (Crannell, Crannell, and Ramaty, 1979). The relative importance of the two processes for solar flares is discussed by Ramaty et al. (1983) and by Rieger et al. (1983).

Until February 1984, 14 flares were observed with emission above 10 MeV. Their time history has the following general characteristics: They are of short duration (~ 1 min) and very impulsive. Rise and fall times are on the order of seconds. They exhibit single or multiple peaks. The peak emissions of the hard X-ray and gamma rays are simultaneous within about ± 6 sec. There is no systematic delay of the gamma rays with respect to the X-rays.

Because of the simultaneity of the gamma- and X-rays, the energy loss of the highly energetic particles (electrons and/or ions) has to take place at ambient densities > 10¹³ cm⁻³, if we assume a thick target situation (Bai and Ramaty 1976).

Figure 3.1 Time history of the flare of June 15, 1982 in various energy bands.

3.2.1.2 Directivity of Highly Energetic Particles

The flares with photon emission above 10 MeV are at heliocentric angles of > 60° (Rieger et al., 1983). This is shown in Figure 3.2, where the location of the flares, known from Hα observations (NOAA: Solar Geophysical Data) is plotted. The arithmetic mean heliocentric angle of all 14 flares is 79°. If the radiation is isotropic the probability for a chance coincidence of such a distribution is ~ 2 × 10⁻⁷. Therefore we conclude, that this “limb brightening” is the result of directivity of the radiation. This directivity must exist also in the primary highly energetic particles, because photons with energies above 10 MeV, if created by electrons (> 10 MeV) via bremsstrahlung or by protons (> 100 MeV) via pion decay, are emitted preferentially in the direction of the motion of these particles. At this time the most likely interpretation appears to be that the emitting particles are indeed travelling roughly normal to the sun-center direction (J. Cooper and V. Petrosian, private communication), as if they were near their mirroring point, for example.

To study this phenomenon in more detail it would be necessary to make stereoscopic observations with two de-
tectors of the GRE type widely spaced in solar aspect angle, as was done by Kane et al. (1980, 1982) at lower energies.

3.2.2 Hard X-ray and Microwave Morphology

3.2.2.1 Observations

The strong correlation in the temporal evolution of hard X-rays and microwave radiation established over the years has led to the belief that the same or very closely related populations of electrons are responsible for both of these two different flare radiations. High spatial resolution observations at both wavelength ranges provides further information on the details of this correlation.

Since the first fanbeam observation of flares by Eno M et al. (1969) with resolution 24" at 9.4 GHz, the new instruments such as the NRAO 3 element interferometer with resolution 10" at 2.7 GHz (Aliassandrakis and Kundu 1975), the WSRT array with one dimensional resolution of 6" at 5 GHz (Aliassandrakis and Kundu 1978, Kattenberg and Allart 1981) and finally the VLA with two dimensional resolution of 1" at 15 GHz (Marsh et al., 1980, Lang et al., 1981 and Kundu et al., 1981) have steadily improved the resolution of microwave observations. VLA observations (Kundu et al., 1982, Dulk et al., 1983) and fanbeam observations at 35 GHz (Nagoya) and 17 GHz (Nobeyama) have continued, while the HXIS on SMM and the SXT on Hinotori have begun for the first time to provide spatially resolved images with resolution of about 10" at X-ray energies of less than 40 keV. Unfortunately, simultaneous microwave and hard X-ray high resolution images exist for only a few flares. Consequently, to some extent, we must still rely on the data of more numerous events with only either hard X-ray or microwave images.

We first summarize the results of observations where there exists only high resolution microwave data, then those with only X-ray data and, finally, the few cases with both X-ray and microwave observations.

a. Microwave Morphology. Studies of spatial structure of microwave radiation from flares have been summarized by Marsh and Hurford (1982) and Kundu (1983). It should be noted at the outset that, unlike the X-ray studies, the microwave observations (in particular the high resolution one-dimensional results from WSRT and the two-dimensional results from the VLA) have not been as systematic, continuous and extensive as the X-ray observations. Consequently, it is difficult to classify the microwave structure in well defined categories. In spite of this, these observations have shown some common features on which we will concentrate here, remembering that there may be more complex structures yet to be studied, analyzed, and classified.

The most general statement that can be made is that the brightest point of the microwave radiation during the impulsive phase occurs near a magnetic neutral line and not on an Hα kernel, and that in cases with simple field geometries there is a single dominant source. Whenever a secondary source is detected, that source lies near another neutral line (e.g., 26 May 1980 flare, Kundu 1983). It can be concluded that with few exceptions (see Kundu et al., 1982), the predominant microwave emission does not come from the footpoints of the flaring loop. In some flares the microwave source is approximately midway between the footpoints and is nearly equally but oppositely polarized (circular) on both sides of the maximum brightness spot (see Figure 3.3). The simplest interpretation of these observations is that the emission comes from the region around the top of a loop. Spectral variation of the structure and polarization then tell us something about the geometry of a loop and the pitch angle distribution of the radiating electrons (Petrosian 1982).

However, there are flares where only one sense of polarization is observed (cf. Kundu 1983), and there are cases where the source is not located very high up in the corona (Kai et al., 1982). These features have been attributed to asymmetric magnetic loops (Kundu and Vlahos 1979) but, as we shall discuss below, such structures could also arise when a non-circular or sheared loop is viewed from an angle away from the line perpendicular to the field at the top of the loop.

b. X-ray Morphology. Prior to the era of SMM and Hinotori, the only information on hard X-ray spatial structure was obtained from stereoscopic observation by the PVO and ISEE-3 satellites (Kane et al., 1979, 1982). These ob-
Observations do not provide images, but give easily-interpreted information on the height of the hard X-ray emitting regions. Furthermore, they extend to much higher X-ray energies than the imaging instruments and tell something about the variation with height of X-ray spectra, which turns out to be very useful in modeling (cf., Brown et al., 1981, Leach and Petrosian 1983).

The first hard X-ray images of flares were obtained by the HXIS on SMM. The analysis of the few early flares emphasized X-ray structure consisting of two sources which were identified as the footpoints of flaring loops (Hoyng et al., 1981). But as indicated more recently (Duijveman and Hoyng 1983), there are many flares where the bulk of the hard X-rays (\(>16 \text{ keV}\)) come from a single source.

Analysis of the data from SXT on Hinotori (Ohki et al., 1982, Tsuneta et al., 1982, Takakura et al., 1982, Ohki et al., 1983, Tsuneta et al., 1983, Takakura et al., 1983 and Kosugi et al., 1983) indicates the predominance of flare images consisting of a dominant single source which sometimes expands and become elongated, and sometimes shrinks.

There are also rarer occurrences of double sources (not necessarily of equal intensity) which merge into a single source located between the original two sources as the flare progresses. Similar structures and evolution have now been shown to be present in the HXIS data (Machado 1983, and Machado et al., 1983).

Based on light curves and spectral images measured by instruments on board Hinotori, various classifications of the flares have been given by Ohki et al. (1983), Tanaka (1983) and Tsuneta (1983). For the following reasons, the classifications are preliminary at this time. They are based on only the strongest flares, which may not be representative of all flares. The majority are weaker and shorter-lived than the ones used for this classification. In addition, relatively few flares have been classified; the number of flares in each class is small. It should also be noted that they are based on only qualitative differences between spatial structures, spectra and the impulsiveness of the light curve. For a more thorough analysis the parameters describing these characteristics should be quantified and flares binned accordingly. The flares

Figure 3.3 VLA map at 15 GHz in right circular (black) and left circular (white) polarization superimposed on Hα frames. The dashed line depicts the location of the magnetic neutral line (from Hoyng et al., 1983).

3-6

Provided by the NASA Astrophysics Data System
studied may be the extreme cases, so that further analysis may show a continuum of classes with no clear dividing line between the present categories.

The Hinotori data have been classified into three types, A, B, and C, as described in Table 3.1. Type A events are defined as those which have a smooth light curve below 40 keV (no discernable spiky features), a soft spectrum (spectral index > 6 for a power law fit, or an exponential spectrum within observational uncertainty) and consist of a single compact source. Two examples of such flares have been extensively studied (Tsuneta et al., 1984), the April 2nd and July 17th flares of 1981. The April 2nd flare (Figure 3.4) shows a single steady point source (perhaps barely resolved at 15'). Strong Fe XXVI lines are also observed throughout these events.

Type B flares are those with two distinct hard X-ray bright spots during initial (impulsive) part of the flare, which evolve into a single source located somewhere between the initial bright regions. During this evolution it is observed that spikiness of the light curve disappears and the spectrum softens. Hinotori events of July 20 (Figure 3.5) and October 15 of 1981 are two such events.

Finally, there are Type C flares which, like Type A flares, consist of a single source, but the source is more diffuse and is clearly displaced from the Hα kernel and, like the microwave sources described above, could be emanating from the top of a set of loops. The light curve is also smooth, even at energies greater than 49 keV, and the spectrum is hard.

The May 13 (Figure 3.6) and April 27 flares of 1981 are two such events.

As mentioned above, the HXIS data also show flares with spatial structures similar to those described by Hinotori. Figure 3.7 shows some such examples.

It should be noted that some of the Type A flares may actually be of Type B, but unresolved, either because they are intrinsically compact or because they are elongated sources viewed along the major axis.

c. Simultaneous Hard X-ray and Microwave Morphology. There are less than a dozen flares which are resolved (in one-dimension or two) at both hard X-ray and microwave energies. Here we describe the brightest of these and concentrate on the relation between the hard X-ray and microwave structure of the most important features. We begin by summarizing Hinotori observations of four events representing the three types described above, and then four SMM X-ray events that have also been observed in microwaves by the WSRT or the VLA, and fit them into categories A, B or C.

The August 11, 1981, Hinotori flare is classified as Type A, but as far as the X-ray image is concerned, it falls on the border line between Types A and C. Without a quantitative measure of the distinguishing parameters of the various types, one cannot resolve this ambiguity. In any event, the X-ray image appears to be a steady single elongated source (with a hint of a double source at the rising phase). The one-

<table>
<thead>
<tr>
<th>Table 3.1 Hinotori Flare Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>OBSERVATIONAL CHARACTERISTICS</strong></td>
</tr>
<tr>
<td><strong>TYPE</strong></td>
</tr>
<tr>
<td>A</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>B</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>C</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

Provided by the NASA Astrophysics Data System
HINOTORI SOLAR X-RAY IMAGE
SXT-2 30-50 KEV YR:MM:DD = 81:8:11
ONE PIXEL = 9.958 (ARCSEC)
CONTOUR: MIN = 1.0000 STEP = 1.4142

Figure 3.4 Example of a Type A flare, Aug. 11, 1981. The lower right-hand panel shows a sketch of the Hα flare at 01:55:45 UT and the direction of the peak of the one-dimensional brightness at 17 GHz at 01:45:10 UT. The other five panels show five different X-ray (25-50 keV) images (Takahara et al., 1983).

dimensional 17 GHz tracing shows a near coincidence of the peak microwave and X-ray positions (Figure 3.4). Both maxima seem to be located over a neutral magnetic line. (For details see Takahara et al., 1983.)

The April 1 and May 13, 1981, Hinotori flares are both classified as Type C. The May 13 flare has a very smooth time profile even above 100 keV, while the X-ray light curve for the April 1 flare is midway between the August 11 and May 13 flares. As shown in Figures 3.6 and 3.8, the 35 GHz microwave position agrees with the position of a single dominant X-ray source. Both emissions appear to come from the top of a loop or arcade of loops. However, there is some asymmetry in the one-dimensional microwave image. This asymmetry becomes stronger (indicating the existence of two sources) in the later phases of the flare, and in the case of the April 1 event (Figure 3.8) it may be interpreted as a double source. Unfortunately, there is no X-ray image for these periods.

The October 15, 1981, Hinotori flare may be classified as Type B because initially it consists of two sources (one much brighter than the other). However it is not clear if these sources correspond to any Hα kernels. The 35 GHz image can be decomposed into a double source with the stronger microwave source corresponding with the weaker X-ray source (Figure 3.9). There is a secondary peak in the light curve during which the X-ray structure changes rapidly. The brighter source moves about 30 arcsec southward and the weak source becomes stronger. Finally, during the decaying phase these sources disappear and a third source appears in a new position. There are no microwave images for the second peak or the decay phase of this flare. Probably this is a Type B flare like the July 20, 1981, limb flare. However, one cannot rule out the less likely possibility that the four different bright regions correspond to the top of four different loops.

The November 5, 1980 SMM event (Figure 3.10) described by Hoyng et al. (1983) has excellent overlapping HXIS and VLA coverage. As is typical, the 15 GHz radiation comes from above a neutral line, as shown in Figure 3.10. The hard X-ray and 9.4 GHz light curves show a Type B time profile and consist of three peaks, with the third peak being a soft, gradual one. The 16 to 30 keV HXIS counts in various pixels conform to the Type B character of this event in that, initially, there is less flux from the region over the neutral line than from the Hα bright spots, and during the second and third peaks most of the X-rays come from the approximate center of gravity of the bright region and coincide with the bright microwave source.

The July 13, 1980, SMM event described by Kattenberg et al. (1984) shows two sources in coincidence with the Hα bright patches. Unlike the Type B flares there is no gradual peak (or phase) in the light curve and no emergence of a single source in the middle. Unfortunately, the baseline of the WSRT one-dimensional microwave tracing is perpendicular to the line connecting the two X-ray sources so that we do not know if the microwave source is also double. However, the usual centrally-located microwave image is consistent with the data.

On June 24, 1980, two bursts at 15:20 and 19:57 were extensively mapped at 6 cm by the VLA with some coverage by the HXIS on SMM (Kundu et al., 1983). The 15:20 event shows a hard X-ray image displaced by 20 to 30 arcsec from the microwave source, the latter being typically bipolar (separated right and left circularly polarized structures, as in Figure 3.3). The softer X-ray and microwave
emissions are co-spatial, presumably coming from the loop top, while the hard X-rays may come from a footpoint. This may therefore be a Type B event.

d. Summary. The sample of flares described above is obviously incomplete in many ways. In general, any sample is limited by selection effects associated with observing instruments. However, certain inferences about flare phenomena are affected less than others by such selection criteria. We hope that by concentrating on the gross features of the flare phenomenon the influence of the selection effects can be minimized.

The prominent features of the X-ray and microwave morphology on the basis of the small number of events presently analyzed, can be summarized as follows:

1. For almost all flares where there exist simultaneous microwave and Hα images the microwave emission comes from a region above the neutral line (presumably around the top of a loop) between Hα bright patches (presumed to be the footpoints of loops). There are exceptions, when a minor microwave feature may be near an Hα bright patch (Dulk et al., 1983). Also, there are flares (e.g., Oct. 10, 1981) where because of coincidence of the microwave and hard X-ray images, it is assumed that all the emission is from footpoints. However, there is no direct evidence for microwave emission from the footpoint of loops.

2. In the majority of cases the X-ray emission comes from a single dominant source (Types A and C), which is found to be located above a neutral line and not on an Hα bright patch, except that for some compact, Type A bursts higher resolution is needed for confirmation of this picture. The structure of Type B bursts with two bright X-ray sources (normally coincident with Hα patches) evolves into a single source located between the Hα patches.

3. The limited hard X-ray and microwave data obtained so far are consistent with, but are not a direct evidence for, the conclusion that such large scale changes in the images occur on a hydrodynamic time scale, with velocity v ~ 50 to 200 km/sec.

4. The hard X-ray and microwave images, even though different in detail, have roughly the same location whenever both images are dominated by a single source. For bursts whose X-ray images are initially double, the single source that develops later most probably coincides with the microwave source, presumably located near the top of a loop or arcade of loops.

As we shall show in the next section, the above general features are consistent with the nonthermal model, whereby semi-relativistic electrons are injected in a closed loop somewhere in the corona. One aim of such observational and theoretical work is to determine what constraints the observations provide on model parameters such as field geometry, loop
size and density, and the spectrum and pitch angle distribution of the accelerated electrons.

3.2.2.2 Interpretation

We now interpret the above data in the framework of the nonthermal models whereby electrons are accelerated to energies well beyond the thermal energy of the coronal plasma. These models are variously classified as thick target, thin target, trap, etc.

The traditional trap models require both a rapid convergence of the field lines and a low density. For electrons of energy $E$ to mirror back and forth in the corona, one needs $\delta B/B \gg \delta L/\lambda$, where $\delta L$ is the length scale of the magnetic field variation $\delta B$, and $\lambda$ is the mean free path of electrons of energy $E$. For the X-ray-producing electrons ($E \sim 20$ keV) to survive for time $t$ one needs a value of density $n < 10^4 (E/20 \text{ keV})^{3/2} (100/t \text{ sec}) \text{ cm}^{-3}$.

The thin target model will be realized only in an open magnetic field configuration. Isotropic or outward streaming injections of electrons can be dismissed because, contrary to observations (Lin and Hudson 1971, 1976), they require as many electrons to reach the earth as are needed to produce the hard X-rays. If the electrons are highly beamed toward the chromosphere, the result will be similar to the thick-target closed-field configurations discussed below.

Consequently, we will consider only the general thick-target model where electrons are injected in a closed magnetic loop (Figure 3.11). We shall need to specify the site of injection of the electrons. If the site of injection is very deep, well below the transition zone (point 1 of Figure 3.11a), the electrons lose energy quickly, causing direct heating and evaporation of chromospheric plasma, and produce hardly any hard X-rays above the transition region. We shall not discuss this possibility here because it will fail to describe Type C bursts discussed above. On the other hand, if the injection is at lower densities (point 2 of Figure 3.11a), so that all of the outgoing particles eventually reach the top of the loop, then the situation is qualitatively similar to the injection at the top of the loop (point 3), which is what we shall assume.

**a. Transport of Electrons.** The transport of the nonthermal electrons is affected by many parameters. In particular, if the pitch angle distribution of the electrons is highly anisotropic (beamed electrons), they are subject to some instabilities which change the distribution on a time scale of the order of plasma oscillations. We shall not consider such cases here. Furthermore, if the beams constitute a high current, a reverse current is set up. The electric field which drives the reverse current may then decelerate the original beam. These aspects of the transport are discussed in Section 3.3.1. Here we concentrate on the collisional effects and the role of large-scale static magnetic fields.

The solar flare plasma conditions allow for various simplifying assumptions, so that the variation with depth of the distribution of electrons $F(E, \mu, \tau)$, where $F \, dE \, d\mu \, d\tau$ is the electron flux in the energy range $dE$ (in units of $mc^2$), pitch angle cosine range $d\mu$ and dimensionless column depth range $d\tau = dN/N_0$ ($N_0 = 5 \times 10^{22} \text{ cm}^{-2}$ is the column depth required for stopping an electron with $E = 1$) is determined by a single parameter $dN/d\tau$, which describes the variation of the magnetic field $B$ along the loop, with respect to the distribution of plasma. Parameters of the injected electron spectrum $F(E, \mu, \tau) = AE^{-\delta}/(\mu^2 - 1) \mu^2 \, d\mu$ are the spectral index $\delta$ and the pitch angle dispersion $\alpha_\phi$. Leach and Petrosian (1981) applied the Fokker-Planck method to electron transport in flare loops. Here we point out an aspect of their work that is relevant to the relationship between hard X-ray and microwave emission.
Let us initially consider loops with uniform magnetic field, i.e. $d\eta B/d\tau = 0$. For nonrelativistic electrons, pitch angle diffusion and energy loss take place on comparable time scales. This means that models with widely different values of $\alpha_3$ appear different only at the top of the loop. Such models are virtually indistinguishable models lower down, where the electron distributions become nearly isotropic. Consequently, the radiation signature of spatially unresolved loops will tend to be closer to what one expects from an isotropic rather than a highly beamed pitch angle distribution. Note, however, that this is not true at relativistic energies (i.e. electrons producing the high-frequency microwave radiation and the continuum $\gamma$-rays). Relativistic electrons lose energy more quickly than they diffuse in pitch angle.

The effect of magnetic field convergence, i.e., of a non-zero $d\eta B/d\tau$, is to mirror some of the electrons which do not reach to the chromosphere and confine them to the top of the loop. However, such electrons also eventually scatter into smaller pitch angles, penetrate deeper, and thermalize. The overall effect of the non-zero $d\eta B/d\tau$ is to isotropize the pitch angle distributions more quickly and enhance the effects described in the previous paragraph.

**b. Non-thermal Emission.** The most prominent aspects of the hard X-ray and microwave observations can readily be understood in the framework of a simple semi-circular loop model, as described above. Although this model is too simplified to use in interpreting the observations in detail, it shows that the geometry of the loop and the physical parameters of the plasma play a more significant role than emphasized in previous treatments. Following Leach and Petsosian (1983) and Petsosian (1982), we assume a semi-circular loop (radius R) in the corona with variable magnetic field which becomes vertical and uniform in the chromosphere (see Figure 3.11a). We also assume that the density $n_0$ of the preflare plasma in the loop is constant in the corona and rapidly increasing in the transition region. The important parameters for our consideration are the column depth of the transition region, $N_{\tau r} = n_0 \pi R/2$, and $d\eta B/d\tau$.

We begin with the spatial structure of X-rays in the range 16 to 50 keV. The X-ray emission is obtained from the electron flux $F$ given by the transport analysis, and the bremsstrahlung radiation cross-section. At these energies the electrons are non-relativistic, and their radiation is nearly isotropic. The scaling between electron energy $E$ and column depth $\tau$ directly translates into a similar scaling between $\tau$ and the photon energy $k$ (in units of $m_e c^2$). Leach (1984) showed that the variation of the X-ray intensity $I$ (normalized photon counts per energy interval $dk$ and depth interval $d\tau$) with $k$ and $\tau$ can be described to within 20% by

---

*Figure 3.7 Examples from HXIS of double sources, single sources and evolving structures similar to those seen by Hinotori (from Duijveman and Hoyng 1983 and Machado 1983).*
Figure 3.8 One-dimensional microwave tracing and the hard X-ray image from SXT of the April 1, 1981, flare (from Kawabata et al., 1982).

\[ I(k, r) = \frac{I(k, r) \, dr/\, \tau = k^{-3/2} \, (1 + \tau / k^2)^{3/2} \, dr/k^2. \]  

(3.1)

This simple picture becomes complicated when we translate it into observables. Let us consider the circular loop of Figure 3.11a viewed from directly above. For comparison with observation we need the variation of the intensity projected on the solar disk \( I(k, r) \) with the projected distance \( r \) from the center of the loop:

\[ I(k, r) = I(k, r) dr/\, \tau = I(k, r) \left( n_o/N_o \right) \left( 1 - \tau^2 / R^2 \right)^{-1/2} \]  

(3.2)

The last geometric term plays a significant role, as shown by ds/dr in Figure 3.12. Note, however, that even for non-circular geometry ds/dr (and dr/dr) can increase rapidly as the footpoints are approached. At a given photon energy \( k \), if \( N_o \) is small, or more precisely if \( \tau_o = N_o/N_o < < k^2 \), then \( I(k, r) \) will be nearly constant for \( 0 < r < R \). However, because of projection effects \( I(k, r) \) will rise with \( r \), such that the footpoints at \( r = R \) will appear brighter than the top of the loop at \( r = 0 \) (cf., Figure 3.12b). For \( N_o \) >> \( k^2 N_o \)

Figure 3.9 A comparison of the SXT hard X-ray image with the fanbeam radio brightness after restoration for the October 15, 1981, event (from Kawabata et al., 1982). Whether the two images belong to footpoints of a single loop or emanate from top of the distinct loop is not known.

Figure 3.10 The microwave and hard X-ray images of November 5, 1980, flare. The lower panels are 8" x 16" in size and a dot indicates significant flux in excess of a single temperature fit (from Duijveman and Hoyng 1983). (cf. Figure 3.3)
Figure 3.11 Description of the geometry of loops: (a) semi-circular loop; (b) asymmetric sheared loop. Circled points (1), (2) and (3) refer to the possible locations of injection of electrons. The wavy arrows show the line of sight and the value of $\theta$ used in equation (3). Note that $ds$ and $d\tau = n ds/N_0$ are measured along the loop from the point of injection, assumed to be at (3). $R$ is radius of the loop and $r$ is a projected distance from the center of the loop.

Figure 3.12 (a) Variation with projected distance (half of the loop) from the loop centers of various quantities. (b) Variation of expected hard X-ray brightness for the indicated values of $N_{tr}/N_0k^2$.

the brightness will decrease steadily from the top to the footpoints. The changeover of the location of the maximum brightness from $r = 0$ to $r = R$ will occur quickly around $N_{tr} = N_0k^2 = 10^{30}$ cm$^{-2}$ ($k/22$ keV)$^2$.

At photon energies less than 10 keV ($k < .02$) a low pre-flare value of $N_{tr} = 10^{28}$ cm$^{-2}$ is sufficiently high to ensure that the maximum brightness occurs at the top of the loop. But such a low value of $N_{tr}$ is not sufficient to confine higher-energy (say $> 22$ keV) electrons to the top of the loop. Such electrons will penetrate below the transition region so that their radiation at 22 keV will be concentrated on the footpoints. Only if $N_{tr} > 10^{29}$ cm$^{-2}$ does the top become brighter than the footpoints at these energies also.

This general picture is nearly independent of the viewing angle and the field geometry because in general $d\tau/dr$ becomes very large near the footpoints. However, the details of the shapes of the curves drawn on Figure 3.12 do depend on the field geometry and the viewing angle.

This behavior therefore suggests that for the large diffuse Type C flares described in the first part of this section, $N_{tr} > 10^{30}$ cm$^{-2}$ (perhaps $n \sim 10^{18}$ cm$^{-3}$, $L \sim 10^9$ cm), so that one sees the 10 - 40 keV photons from the tops of the loops. On the other hand, for the Type B events, initially $N_{tr} < 10^{29}$ cm$^{-2}$ and the emission is concentrated in the footpoints. The asymmetry of the radiation from the two footpoints as commonly observed must then be attributed to an asymmetric field geometry (e.g., Figure 3.11b). As observed by Hinotori (Figure 3.5) and in some events by HXIS (Figure 3.7), the double images merge into a single coronal source as the flare progresses. This means that the column depth has increased from its initial low value to a value larger than $10^{20}$ cm$^{-2}$. As mentioned in connection with observations, such changes seem to occur on a hydrodynamic time scale, suggesting that the filling of the loop by the evaporated flare plasma may be the cause of the increase in the coronal column depth. It is straightforward to show that this is a distinct possibility (see the following chapter, Doschek et al., 1984).

For the majority of observed bursts the most likely location of the microwave source is around the top of the loop, as discussed in Section 3.2.2.1. However, the explanation for this observation is different than that presented above for hard X-rays. This is because the microwaves are the result of gyrosynchrotron emission by much higher energy elec-
trons \( (E \approx mc^2) \) which, unlike the lower energy hard X-ray emitting electrons, cannot be confined easily. The confinement to the top of the loop of such high energy electrons by Coulomb collisions would require \( N_p \) well above \( 10^{22} \) cm\(^{-2} \). This is obviously unlikely and trapping in low density \( (n < 10^{10} \text{ cm}^{-3}) \) loops or some other explanation is needed to explain the observed origin of the microwave emission near the top of loops.

Microwave-producing electrons are clearly present throughout the whole loop but the efficiency of gyrosynchrotron emission varies throughout the loop and depends on the direction and strength of the magnetic field. The gyroscopic radiation is strongest in the direction perpendicular to the field line. This factor alone, i.e., assuming everything else is equal throughout the loop, will favor a maximum brightness located between the footpoints. For example, the simple loop of Figure 3.1a, when viewed from above, will be brightest at \( r = 0 \). Note that this aspect is, to the first order, independent of the viewing angle. However, the gyrosynchrotron emission depends also on the number of electrons and the strength of the magnetic field. At a given frequency \( \nu \) the emission from stronger field regions corresponds to lower harmonics of gyromotion \( \nu_b \). The lower harmonics are produced by lower energy electrons, \( E \approx \left( \nu/\nu_b \right)^{1/2} \), which are more numerous. Since the field strength only decreases with the height above the photosphere, emission from higher regions could dominate only if this variation is slow. Finally, the strength of the gyrosynchrotron emission is a function of the pitch angle distribution of the electrons. Only if the pitch angle distribution of the accelerated electrons is broad will the emission from the coronal region be important. Otherwise, if electrons of kinetic energy \( E \approx mc^2 \) are beamed along the field line, they remain so throughout the corona and give little gyrosynchrotron radiation. The pitch angle distribution becomes broader and the radiation efficiency increases only below the transition region, i.e., at the footpoints. Another factor which also aids the relative enhancement of emission from the coronal part of the loop is the fact that gyrosynchrotron radiation is subject to various absorption and suppression effects (cf. Ramaty and Petrosian 1972), all of which are more important in the chromosphere than in the corona.

New observations of flare X-ray polarization by Tramiel et al. (1984), exhibit a low degree of linear polarization \( (< 5\%) \). This result differs from earlier observation by Tindo and his collaborators (see e.g., Tindo et al., 1976). Such low polarization naturally will arise in the so-called thermal models or if the pitch angle distribution of the electrons is nearly isotropic. Earlier studies of nonthermal models (Haug 1972, Langer and Petrosian 1977, Bai and Ramaty 1978), which ignored the scattering of the nonthermal electrons, consequently found a high degree of linear polarization for highly-beamed electrons. Even Brown’s (1972) calculation, which includes the scattering in an approximate fashion but does not include the curved loop geometry, has indicated a higher degree of polarization than observed. As a result it has been assumed that a low observed degree of polarization is a strong evidence against the nonthermal models.

Recent work by Leach, Emslie and Petrosian (1984) has examined the problem more fully, and does not confirm this expectation. The primary reason for a low degree of polarization is that no matter how strong the anisotropy of the injected particles, the average pitch angle distribution in the thick target is very broad (see Section 3.2.2.2). In addition, because the direction of the linear polarization varies along a loop, the integrated X-ray emission from the whole loop will have the characteristic of the emission from a highly broadened distribution and will show lower polarization than one would expect from the distribution of the injected electrons. Table 3.2 summarizes the effect of the spectral index \( \delta \) and the pitch angle distribution dispersion parameter \( \omega_p^2 \) of the injected electrons, assuming the magnetic field to be uniform, and Figure 3.13 compares the observation of Tramiel et al. (1984) with three model calculations.

### Table 3.2 Maximum Percentage Polarization at 16 keV

<table>
<thead>
<tr>
<th>( \delta )</th>
<th>( \omega_p^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \infty )</td>
<td>0.4</td>
</tr>
<tr>
<td>3 ( \leq 5 )</td>
<td>( \leq 5 )</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
</tr>
</tbody>
</table>

*Electrons injected isotropically at the top of the loop.

A low degree of polarization can arise naturally from some nonthermal models. In particular, there is a strong effect of spectral index on polarization; compare tabulated polarizations for various \( \delta \) values in Table 3.2. This is because the degree of polarization increases rapidly as the photon to electron energy ratio approaches unity \( (k/E \approx 1) \). For steep spectra \( (\delta \approx 0) \) the bulk of photons with energy \( k \) are produced by electrons of only slightly larger energies \( (k/E \approx 1) \). However, for harder spectra \( (\delta \approx 5) \) the contribution of higher energy electrons (i.e., those with \( k/E \) considerably less than unity) increases, causing a reduction in the polarization.

Opposite polarization on both sides of the site of maximum brightness (cf. Figure 3.3) is a natural consequence of the simple loop geometry we have been discussing here. The circular polarization is (e.g., Petrosian and McTiernan 1983).

\[
P_c = \cot \theta \left\{ \frac{[3(1+\delta)\nu_b\sin\theta/\nu][1+(\nu_b\sin^2\theta/\nu\cos^2\theta)]^{1/2}}{[1+(\nu_b\sin^2\theta/\nu\cos^2\theta)]^{3/2}} \right\}.
\]
where $\theta$ is the angle between the line of sight and the magnetic field lines. As evident $P_c = 0$ at $\theta = \pi/2$ and increases with opposite sense away from $\theta = \pi/2$ where the total microwave intensity is at its maximum. The simple circular loop, therefore, agrees with the qualitative features of the observations. Alteration of the field geometry changes the variation of $P_c$. In particular, for a highly asymmetric loop (see e.g., Figure 3.11b) it is possible that one sense of polarization becomes dominant, as often observed. Note that this is qualitatively similar to the field geometry proposed by Kundu and Vlahos (1979), though the details of their model and the one presented here are quite different.

### 3.2.3 Combined Soft X-ray and Hard X-ray Spectra

Analysis of the spatially integrated spectra from flares in the range 3 keV to 500 keV can yield much vital information regarding the separation of impulsive and thermal components of the flare. However, the interpretation is complex since several plasma regimes contribute to these spectra, and involve a range of atomic interaction processes. As the lower end, 3 keV to 7 keV, there are dominant line spectra of Fe XXVI, Fe XXV, Fe XXIV, Ca XIX, and in some cases Fe XXIII to Fe XVII. Although these can be excited by thermal or nonthermal electrons, the existence of these species requires that the excitation occur within hot thermal plasma.

Such hot thermal plasma also emits continuum radiation, which may extend up to photon energies of 10 keV, depending on the temperature. Non-thermal electrons, or streams, passing through the hot thermal plasma deposit energy, and contribute an additional component of the continuum spectrum. This process is normally represented by a thick target model. Power law electron distributions give rise to power law spectra. This deposition may take place within the hot plasma itself, or within cold dense plasma outside the hot region, as in the classic foot-point emission models. Although this alternative makes no detectable difference to the continuum radiation emitted, it affects whether this mechanism contributes to the intensities of the above spectral lines. The range of mechanisms is thus large, and is further complicated since coverage of the region normally involves two or three spectrometers of quite different characteristics. However, it is clear that the ambiguity in interpretation is
greatly reduced by attempting a simultaneous analysis over a wide range of photon energies.

3.2.3.1. Line Spectra

The intensities of spectral lines, and their dependence on the plasma parameters, is determined by a range of atomic interaction quantities. For excitation of spectra from the ground state of the ion responsible, excitation and dielectric capture rates are important. For helium and hydrogen-like ions, recent calculations by Bely-Dubau et al. (1982a,b) and Dubau et al. (1981) are widely used. Further work is in progress on the lower ions Fe XXIII to Fe XVII. These spectra lie in the narrow range 1.85 to 1.91 Å but cover a wide range of effective plasma temperatures. They are thus very valuable in extending the differential emission measure analysis downwards below 10^7 K, as well as in studying the transient ionization that occurs during rapid time variations of flare plasma. Moreover, where the ground configuration of the ion is complex (Fe XIX to Fe XXII), the intensity ratio of some of the lines can be used as an electron density diagnostic in the range 10^{12} to 10^{14} cm^{-3}. This could be especially valuable in the early part of the impulsive phase of hot flares.

For many purposes ionization balance calculations are also required for interpreting the line spectra. Many sets are available in the literature. However, some new data are becoming available from the solar flare spectra themselves. Antonucci et al. (1984, and further work in preparation) have used a study of the flare data from the BCS, Hinotori and P78-1 in order to derive ratios of Li-like to He-like ions in Ca and Fe as a function of temperature. The ionization balance calculations of Jacobs et al. (1977, 1980) have been modified to take account of this work, and then used in some of the modeling described later.

3.2.3.2 Transient Ionization

It was earlier hoped that transient ionization effects would form an important diagnostic for the dynamic phases of flares. In this way, departures of the measured ionization ratios from their steady-state values would depend upon local electron density and the rate of change of temperature. It has since emerged that, due to the high electron densities, ionization is very close to steady-state over almost all phases of the flare for which statistically good data are obtainable. It has long been realized that the ratio of Li-like to He-like ions in flares is larger than theory would predict, by about a factor 2. This might be interpreted as a transient ionizing effect, were it not that it continues through all phases, including the flare decay. Koshelev and Kononov (1982) proposed a dynamic oscillating flare model in order to explain this effect by a continuous transient ionization condition. Doschek (1984) has attempted to examine such models on a more quantitative basis. He concluded that it is difficult to find an oscillating model which adequately predicts all the observed spectral intensities. A more likely explanation lies in a factor of 2 error in the ionization balance models derived from theory. This assumption is implicit in the rederivation from observations by Antonucci et al. (1984) mentioned above.

The existence of real departures from steady-state ionization are now being reported by some workers, who examine critically the earliest impulsive phase data (see later in this section). If substantiated, this would provide a valuable measure of the density at this time.

3.2.3.3 Hinotori Analysis

The Hinotori team has attempted to analyze jointly data from two of the instruments; the flare monitor (FLM) and the hard X-ray monitor (HXM). The FLM covers the range 5 to 20 keV with a resolution sufficient to separate line and continuum contributions. The HXM covers the range 18 keV to 400 keV. In an earlier analysis of FLM, Wantanabe (1984) found that while the continuum is best fitted by a power law in the early phase, it rapidly turns into a thermal spectrum. The time histories show that the thermal energy content is adequately represented by the time integral of the hard X-ray signal, indicating that the impulsive phase electrons may carry all the energy required for the flare. He concluded also that there is a strong tendency for all flares to produce a thermal component temperature of 2 \times 10^7 K and density 10^{11} cm^{-3}.

Tanaka et al. (1984), in an attempt to understand the fluorescence origin of the Fe Kα line, have combined the spectra from FLM and HXM. The result, Figure 3.14, shows a smooth blend between the power law behavior above 10 keV with the thermal behavior below 10 keV. One is tempted to conclude that the plasma has a smooth distribution of electron energies merging from streams to thermal around 10 keV. This interpretation would be valid only if both components represented thin target emission. If the streams dissipate in thick target emission, then the transformation from photon spectrum to electron spectrum would be different in the two cases.

3.2.3.4 SMM Analysis

A method has been devised (Gabriel, Sherman, Bely-Dubau, Orwig and Schrijver, in preparation) to model the emitted spectrum based upon input in the form of a differential emission measure distribution plus electron beams. The model computes the predicted fluxes from 8 spectral lines in the Bent Crystal Spectrometer (BCS), the 6 channels of the Hard X-ray Imaging Spectrometer (HXIS) and the 15 channels of the Hard X-ray Burst Spectrometer (HXRBS). This then covers line emission from 10^7 K to 10^8 K and continuum energies from 3 keV to 260 keV. Beam energy distributions can be arbitrary and need not follow power laws.
Thin target emission by the beam within the hot plasma is evaluated as well as the normal thick target emission, but double counting of this radiation is avoided. An inversion technique is not used; it is rather a question of trying various distributions to see which best fit the data. In this way it is possible to take proper account of the varying accuracy of the data. For example, a satellite to resonance line ratio in one BCS channel should be good to \( \approx 10\% \), whereas absolute instrument calibration might be out by up to a factor of 2, although any such systematic errors must then always be present.

The method has been applied to the flare of 1822 UT 29 June 1980 at various times throughout the flare. The results, some of which are shown in Figure 3.15, are preliminary, but the following conclusions emerge:

1. A strong thermal component at \( 10^{7} \) K initially, rising to \( 2 \times 10^{7} \) K later, is always present.
2. This thermal component is not isothermal, but has a real width \( \approx 0.5 \times 10^{7} \) K.
3. At all times evaluated (even past the peak thermal emission) there is an additional high-energy component. If fitted with a thermal distribution, this is at \( T \approx 10^{8} \) K and has an emission measure 3 or more orders of magnitude lower than the primary component. It can also be fitted adequately with a power-law electron beam.
4. Early in the impulsive phase, the first thermal peak in the differential emission measure is smaller and at a lower temperature, while the higher energy component is relatively stronger. At the earliest time tested, it is very difficult to fit a thermal distribution to this feature. It is easier to fit a power law beam with a cut-off around 10 keV, with an electron spectral index of 5-7 on a thick target assumption. At this early stage, a good fit is obtained with the lines only if it is assumed that the ionization balance lags behind its steady-state value.

Figure 3.15 (top) shows ratios of calculated to observed intensities for two times during the flare, early in the impulsive phase and at the peak of the thermal phase. The anomalously high measured intensity of the Fe XXV line during the impulsive phase is an indication of the departure from ionization balance mentioned above. Figure 3.15 (bottom) shows the differential emission measures used in calculating these cases.

It could be interpreted that the lower thermal peak represents the main flare loop being filled with evaporated chromospheric plasma, having a temperature gradient, and temperature stabilization produced by conduction and radiation cooling. The higher energy component is produced by either or both of thick-target bremsstrahlung from the primary fast electrons and a high temperature region identifiable with the primary release. We thus have an indication that for this flare, the energy input continues at a reduced level until well after the peak of the thermal phase. Such an interpretation would be complicated, but not invalidated, if the observed emission arises from two independent flaring regions. In any case, the high energy component, at its peak of intensity, is precisely identical to the normally-assumed signature of the impulsive phase, i.e. a signature of nonthermal bremsstrahlung.

### 3.2.4 Iron Kα Emission

There has been great interest in the radiation mechanism of iron Kα, owing to its value as a diagnostic of nonthermal electron beams. The Kα lines at 1.936 Å and 1.940 Å are radiated by inner-shell ls-2p transitions following the removal of a K-shell (ls) electron from iron ions FeIv-FeXI. The inner-shell ionization occurs either by photoionization by X-rays of energy above the 7.11 keV ionization threshold (fluorescence) or by electron impact ionization. Since the chromosphere is transparent for 7 keV photons, irradiation of cool photospheric iron by X-rays of \( E > 7.11 \) keV could cause the emission (Neupert et al., 1967). Electron impact ionization occurs most effectively for electrons with an energy of about 25 keV. It has been suggested that nonthermal electrons associated with the hard X-ray burst excite Kα lines by collisionally ionizing the iron ions that are in lower ionization stages (Acton 1965, Phillips and Neupert 1973).

Observations of Kα lines in the previous cycle, though performed with relatively low spectral and temporal resolu-
Figure 3.15 Top: Comparison of observed spectra and the model fit. Bottom: The differential emission measure used for the model fit shown above.
ion, have suggested fluorescence as a plausible mechanism (Doschek et al., 1971; Tomblin 1972). During the present
maximum more reliable observations of iron Kα have been
made with much increased spectral and temporal resolution
and increased sensitivity. The P78-1 SOLFLEX (Feldman
et al., 1980) and the SMM XRP (Culhane et al., 1981) ob-
tained clearly resolved Kα1 and Kα2 lines using flat Ge and
bent Ge crystals respectively. The Hinotori SOX (Tanaka
et al., 1982) obtained barely resolved Kα lines and the Kβ
line at 1.755Å by a SiO₂ crystal, and the Tansei IV satellite
(Tanaka 1980) obtained Kα by a LiF crystal which possessed
the highest sensitivity.

These observations have provided results which are in
agreement. All flares show Kα emission which favors
fluorescence mechanism almost throughout the flare. The
fluorescence origin has been quantitatively established by
Parmar et al., (1984) from detailed comparison of the SMM
observations of 40 large flares with Bai (1979)'s model of
fluorescence. In Bai's model, which used a Monte-Carlo
 technique to follow iron Kα photons resulting from the
photospheric absorption of X-rays, the Kα flux is evaluated as
the products of three terms: integrated X-ray flux above 7.11
keV, fluorescent efficiency which depends on the X-ray
source height and source temperature, and a term consider-
ing the absorption and scattering of emergent Kα photons,
which depends on the heliocentric angle of the flare. For each
of the flares Parmar et al., could show close similarity be-
 tween the Kα light curve and that of the E > 7.11 keV X-
ray flux. In fact, the Kα light curves were quite different
from the hard X-ray burst profile of E > 50 keV. A remark-
able center-to-limb variation was found for the Kα to con-
tinuum ratio for different flares at various heliocentric angles,
and explained by Bai's model (Figure 3.16). Referring to
Bai's model, the large dispersion in this ratio for flares at
similar heliocentric angles is consistent with the variety of
flare heights ranging from 0 to 0.05 solar radii for an iron
abundance of 5.5 × 10⁻⁵ per H atom. Conversely, from
the restriction for flare heights in short duration compact
flares which were observed in the Skylab X-ray images
(Pallavicini 1977), a plausible iron abundance has been esti-
mated to be in the range from 5 × 10⁻⁵ to 6 × 10⁻⁵ per
atom.

While the Kα light curve shows gradual variations simi-
lar to those of soft X-rays in most flares, transient Kα emis-
sions in the very early phase of flare have also been reported
in limited number of flares. Culhane et al. (1981) reported
a Kα light curve showing impulsive behavior. During the
15 seconds of the impulsive hard X-ray burst, the measured
Kα flux was 2.5 ± 1.4 times that measured during the fol-
lowing 15 seconds (Parmar et al., 1984). However, this case
has been considered rather exceptional in the SMM data. The
LiF crystal on the Tansei IV satellite detected a sudden en-
hancement of Kα emission that preceded an enhancement
of the high temperature iron lines around 1.85Å in three
flares (Tanaka 1980). This enhancement occurred in the ris-
ing phase of the microwave burst at 17 GHz. The Hinotori
observed similar, but less evident, cases in several (7-8) very
impulsive flares. In these flares the Kα to Fe XXV resonance
line ratio, which normally takes a fairly constant value of
about 0.1, showed a very high value from 0.3 to 1.0 only

![Figure 3.16](image_url)

**Figure 3.16** The ratio of Kα flux to the E > 7.11 keV continuum for various flares plotted against the heliocentric angle. The ratios predicted by Bai's model for an iron abundance of 5.5 × 10⁻⁵ per H atom, a temperature of 2.0 keV and three specific flare heights are shown as broken curves (Parmar et al., 1984).
in the initial phase, in close time coincidence with the hard X-ray burst.

These examples may provide candidates for the electron impact Kα emission. However, Tanaka et al. (1984) have suggested that the observed excess emission is mainly explained by fluorescence from the hard X-ray burst extending in energy to 7 keV. The high resolution continuum spectra from the Hinotori have revealed that a single power-law distribution prevails from 70 keV to below 10 keV in the early phase of some impulsive bursts (see Figure 3.14). Because in the very impulsive flares the thermal X-ray flux originating from 20 million degree plasma at 7 keV is less than the flux of the extended power-law distribution, fluorescence Kα shows impulsive behavior correlated with the hard X-ray burst. In the majority of flares, however, the thermal X-ray flux already dominates the 7 keV region when low-sensitivity crystal spectrometers can observe. One example of an impulsive event is shown in Figure 3.17. While the Kα light curve from the middle phase is explained by fluorescence due to thermal X-rays, the Kα flux in the early phase much exceeds the predicted values. The excess fluxes are explained by fluorescence from the power-law X-ray fluxes assuming that the source is located at 0 km. The transition from power-law X-ray fluorescence occurs rather smoothly around the peak of the hard X-ray burst, as expected from the fact that the thermal X-ray flux increases most rapidly at that time, with its flux increasing approximately in proportion to the time integral of the hard X-ray flux (see, e.g., Tanaka et al., 1982). However, the conclusion that all Kα emission is due to fluorescence depends on assigning a height of zero to the power-law source. If this source is located high up in the corona, as has been observed in some gradual bursts observed by Hinotori, then an excess Kα flux remains to be explained, presumably by electron impact. In the impulsive bursts like this case, however, this seems unlikely.

This result may be compared with the thick target theory of the electron impact Kα developed by Emslie, Phillips, and Dennis (1985). The electron impact Kα yields predicted by this theory are always smaller than the Kα yields for fluorescence due to the power-law X-rays extending down to 7 keV, provided that the spectral index is larger than 3.5. Therefore, in most cases the electron impact Kα emission is veiled under the fluorescence Kα, even if the electron beam does exist. In conclusion, the electron impact Kα appears to be not necessary to explain the observations. On the other hand an evidence of low energy (below 10 keV) power-law photons may raise another problem related to the flare energetics because the energy included in the low energy electrons as derived from the thick target hard X-ray model is overwhelmingly large, probably exceeding the energy radiated in the UV and X-rays regions (Tanaka et al., 1984).

3.2.5 Ultraviolet and Hard X-ray Emission

The spatial, temporal and energetic relationships between impulsive ultraviolet emission and hard X-ray emission have attracted considerable interest as tests of flare particle transport models. Kane et al. (1980) reviewed the status of this field in the Skylab era. Two quite different UV observational approaches have been taken. Below, in Section 3.2.5.1, we review and interpret observations in the wavelength band from 10 to 1030 Å. This band contains emission lines and continua that originate at a wide range of temperatures, from 10^5 K. We then turn to recent SMM observations, which are more straightforward to interpret—observations of the transition zone line 0 V 1371 Å, discussed in Section 3.2.5.2, and their interpretation in terms of both thermal and nonthermal energy-transport models.

3.2.5.1 10 – 1030 Å and Hard X-ray Emission

An extensive set of calibrated concurrent measurements of impulsive hard X-ray and integrated 10-1030 Å (hence-
forth EUV) emission has been obtained by Kane and Donnelly (1971), Donnelly and Kane (1978) and Kane, Frost and Donnelly (1979). The EUV observations have no spectral or spatial resolution, but good temporal resolution (~1 s). Although these measurements were made indirectly using the response of the ionosphere, subsequent direct measurements from spacecraft by Horan, Kreplin and Fritz (1982) have fully validated the indirect method. The hard X-ray observations have no spatial resolution, but have sufficient spectral resolution to allow a power-law spectral fit, and also have good temporal resolution. All studies of the correlation between the EUV and hard X-ray bursts have found coincidence of peaks in the EUV flux and X-ray flux to within the time resolution of the instrument. This observation rules out certain types of thermal models (see Section 3.2.5.2b).

The observational datum of primary interest for each impulsive burst is the EUV/HXR ratio at the peak, where HXR is the flux in hard X-rays (>10 keV), both measured at the top of the earth’s atmosphere. Data for many impulsive bursts are shown in Figure 3.18. Donnelly and Kane (1978) and Kane, Frost, and Donnelly (1979) concluded that the peak EUV/HXR ratio is constant, rejecting nonlinearity in the peak EUV/HXR relationship seen in large flares as a selection effect. On the other hand, McLemont, Canfield and Fisher (1984) concluded that the peak EUV/HXR ratio is not constant, but rather the peak EUV ∝ HXR⁴. The issue arises because the largest flares in Figure 3.18, for which the peak EUV/HXR ratio is smallest, tend to have been observed by a single spacecraft, TD1-A.

Until recently there has been general agreement that the observed EUV/HXR ratio is about 30 times smaller than the value predicted by typical electron beam fluxes in the thick-target nonthermal model, in which all the X-ray emission is produced by nonthermal bremsstrahlung and all the EUV emission by thermal emission from the collisionally-heated transition region and chromosphere (Donnelly and Kane (1978), Emslief, Brown and Donnelly (1978), Kane, Frost, and Donnelly (1979)). Donnelly and Kane (1978) suggested that the discrepancy could be explained by the “partial precipitation” model, in which the bulk of the X-ray emission is from electrons trapped in the corona and only a small fraction of the nonthermal electrons reach the chromosphere. However, Emslief, Brown and Donnelly (1978) pointed out that trapping of nonthermal electrons in the corona could not cut down the EUV flux sufficiently because electrons are scattered out of the trap onto the chromosphere after losing only roughly half of their energy (Melrose and Brown, 1976). Emslief, Brown, and Donnelly (1978) suggested that the thermal hard X-ray model (e.g., Brown 1973a) may be more appropriate.

The peak EUV/HXR ratio also shows a strong center-to-limb dependence. Donnelly and Kane (1978) showed that this dependence could be explained if the EUV emission comes from a source embedded in the chromosphere. They explained the dependence satisfactorily by a model in which the EUV emission comes from the bottoms of cylindrical wells of two different sizes. The idea that the EUV flare source is below the normal nonflaring chromosphere is compatible with observations of Hα spectra discussed in §II.G below.

McLemont, Canfield and Fisher (1984) have looked at another aspect of the predictions of the thick-target model. Following previous authors (Brown 1973b, Lin and Hudson 1976, Emslief, Brown and Donnelly 1978), they calculated the depth of the flare transition region as a function of the electron beam flux per unit area. Considering only collisional energy losses, as in previous work, they find that the increase in depth of the transition region with increasing beam flux causes the EUV/HXR ratio to fall as the flux increases,

$$\frac{\text{EUV}}{\text{HXR}} \propto (\text{HXR/} \text{Area})^{-(d-2)/(d+2)}, \quad (3.4)$$

3-21
where the electron spectral index $\delta$ is related to the hard X-ray spectral index $\gamma$, $\delta = \gamma + 1$. The physical reason for the ratio to decrease with increasing flux is that the mass fraction of the beam-heated column that is filled with chromospheric material, which emits much of the EUV, decreases as flux increases. For the most commonly observed spectral indices, $\gamma = 4$, the predicted slope of the peak EUV/HXR ratio as a function of hard X-ray flux is in excellent agreement with the slope that they find in the observational data, if it is assumed that the beam cross-section is constant in all flares, i.e., only the electron energy input per unit area varies. Second, as shown in Figure 3.18, both the slope and the absolute value of the peak EUV/HXR ratio vs. HXR curve are correctly predicted if one assumes that the beam cross-section is of order $3 \times 10^{16}$ cm$^2$, comparable to typically observed chromospheric flare kernel areas.

Were it not for the fact that extremely high beam fluxes are implied, one would conclude that the thick-target electron heating model, with a universal beam cross-section, explains the observations very well. With a 20 keV cutoff energy, the beam fluxes range from about $10^{13}$ ergs cm$^{-2}$ s$^{-1}$ for the smallest flares to $10^{14}$ ergs cm$^{-2}$ s$^{-1}$ for the largest. The majority of the observed flares have beam fluxes greater than $10^{13}$ ergs cm$^{-2}$ s$^{-1}$, the commonly accepted upper limit for return current stability.

How, then, can the peak EUV/HXR ratio be understood, if beam fluxes greater than about $10^{13}$ erg cm$^{-2}$ s$^{-1}$ are not allowed? McClymont, Canfield and Fisher (1984) considered three processes which decrease the EUV/HXR ratio predicted by the thick target model. Firstly, if the preflare loop, in which electrons are accelerated, is hotter (and therefore denser) than a typical preflare loop—a state that might be inferred from the frequent observation of a soft X-ray precursor to the impulsive phase—the bulk of the electron beam energy may be deposited in the corona. They found that the drop in peak EUV/HXR ratio could be explained by this hypothesis, although rather high preflare temperatures ($\geq 10^7$ K), and correspondingly high densities, were necessary to obtain the very low peak EUV/HXR ratio observed in large flares. Secondly, thermal hard X-rays may be important, as investigated previously. Thirdly, the effects of reverse (return) currents have not been considered fully in calculations of this type. Due to reverse currents, concentration of heating at the foot of the transition region may greatly increase the ability of electron beam heating to raise chromospheric plasma to coronal temperatures, compared to pure collisional heating, thus reducing the emission measure in the temperature range where the EUV emission originates. No quantitative studies of this effect have yet been done.

### 3.2.5.2 OV and Hard X-ray Emission

**a. Observations.** Observations of individual EUV lines are potentially more powerful than measurements of the total 10-1030 Å EUV emission, discussed above. One such line, well observed by SMM, is the OV 1371 Å line. The SMM observations have demonstrated a close observational relationship between this line, formed in the transition region at a temperature of about 250,000 K, and hard X-rays. Progress has been made at defining temporal, spatial, energetic, and morphological relationships.

The temporal relationship between HXR and OV was studied in detail by Woodgate *et al.* (1983). In this work they studied the three SMM flares for which UV data were available on a time scale of 1.3 s, and HXR data (from 25—100 keV) were available every 0.128 s. Each flare displayed at least five impulsive bursts in both HXR and OV that were simultaneous to within the 1.3 s time resolution. Measurement of time lags for the peaks ($t_{OV} - t_{HXR}$) yielded an average of $0.3 \pm 0.5$ s. Thus, their results showed that peaks in emission that occurred in both OV and HXR were simultaneous to well within the time resolution.

The Woodgate *et al.* (1983) results also showed that there was a good occurrence correlation between OV and HXR peaks. For the 16 HXR peaks there were 14 OV peaks (80%) and for 19 OV peaks there were 14 HXR peaks (74%). The correspondence is obviously quite good “with the primary difference being due to some small UV peaks late in the impulsive phase that do not appear in the hard X-rays”. These small OV peaks may be associated with HXR bursts that are below the threshold of detectability, but may also indicate additional sources of heating.

Poland *et al.* (1984) have studied the relation between energy emitted in HXR vs. OV as a function of time for several flares. The results for the flare of 2 November 1980, whose light curves are shown in Figure 3.12, are presented in Figure 3.20. In the latter figure the boxes represent data points before flare maximum, the diamonds are within 5 s of flare maximum and the plusses are after maximum. It can be seen that there is a well-defined relation throughout the flare between the HXR and OV emission. Also, the premaximum and postmaximum emissions lie on almost the same line, with the postmaximum OV remaining slightly brighter than the premaximum emission for the same HXR emission. Ten flares were studied in this manner, with all showing similar results. Most of the flares show the rise phase and fall phase occurring on approximately the same line. However, some do show a significant deviation from this relationship in that OV emission is brighter during the decay phase than during the rise phase. There is also a significant variation in OV emission from flare to flare, but, this may in part be due to the variation in total background in OV from one flare to the next. When these data are plotted with the preflare background removed there is still a large variation in the OV intensity from one flare to the next.

The most important point to be made from these observations is that HXR and OV emission rise and fall together,
Figure 3.19 The upper curve shows emission in OV for the entire flare as a function of time (actual values are $10^{23}$ times larger than shown). The lower curve shows emission in HXR above 25 keV for the entire flare as a function of time (actual values are $3.84 \times 10^{20}$ times larger than shown). Noise errors for a single measurement are shown at two different times for each plot.

Figure 3.20 HXR emission as a function of OV emission in ergs s$^{-1}$ for the flare shown in Figure 3.19. Squares represent data obtained before flare maximum, diamonds are within 5 s of maximum, and plusses are after flare maximum. Only every third point has been plotted for clarity of the figure.
with the rise and fall phases usually on approximately the same line, independent of time, during the entire event. How can this be understood? It suggests that the physical conditions in the transition region are directly related to the production of HXR, and independent of those flare quantities that are known to vary on gradual phase timescales, such as the temperature or emission measure of the thermal X-ray plasma. It is straightforward to show that on gradual phase timescales (\(\sim 1\) m) the emission measure at OV temperatures in a conductively dominated transition region is proportional to the transition region pressure, so the observation of relatively little intensity variation during a flare such as the one shown in Figure 3.20 implies that the transition region pressure in the emitting regions varies very little in the course of the observations (a period \(\sim 1\) m). If the model of an electron beam in a single flare loop is to explain both the impulsive and gradual emissions, then apparently pressure equilibrium is either never reached or is substantially unaltered by the impulsive event. On the other hand, it may be that the impulsive phenomena arise in a different volume than the gradual phenomena; we explore the arguments on this issue below, pro and con.

To within SMM’s ability to spatially resolve emitting regions in HXR and UV, the two emissions arise from the same location on the sun. However, one must bear in mind that the UV observations are obtained at only 3”×3” spatial resolution, at best, and more usually 10”×10”. X-rays between 3.5 and 30 keV are observed with a resolution of 7”×7” (van Beek et al., 1980), at best. Harder X-rays, above 30 keV, are observed without spatial resolution. Hence, our knowledge of the cospatiality of these emissions is strongly constrained by observational limitations.

Spatially resolved observational information on the cospatiality of UV and HXR emission is limited to just a few flares. In a study of the June 29, 1980 flare Poland et al. (1982) showed that positions and changes in positions of UV transition region emission corresponded to positions and changes in position in 3.5 to 30 keV HXIS emission. These correspondences, together with the temporal correspondences between OV and higher-energy HXRBS hard X-ray light curves strongly suggest their cospatiality in this flare. Duijveman et al. (1982), in a study of the November 5, 1980 flare, reported cospatial HXR and UV bursts in one of the flare footpoints. Machado et al. (1982) found similar cospatiality in the April 10, 1980 flare. Although these two latter papers did not study the UV/HXR correlation in detail, the HXIS data quality is better than for the June 29 event.

In another SMM flare observation, Cheng et al. (1981) examined the SiIV and OV UV emission with a resolution of 4”×4” and found that each individual burst was seen in both UV and spatially unresolved hard X-rays, and that each component UV spike originated from a separate discrete flaring point. The density at these points was determined to be in the range of 5×10^{12} to 10^{13} cm^{-3}. There is an uncertainty in using this ratio since the two lines are formed at significantly different temperatures; the reliability issue is discussed in Cheng et al. (1982).

The close time-independent functional relationship between OV and HXR, shown in Figure 3.20, motivates us to ask what is the observational evidence that indicates the relationship between the volumes in which impulsive phase emissions and gradual phase emissions arise? Two arguments suggest that impulsive phase emissions and gradual phase emissions arise in separate volumes. Firstly, it is known that in X-rays the emission is impulsive only for energies above approximately 5 keV; for lower X-ray energies the emission does not vary impulsively. For transition region emission, flare enhancements become increasingly impulsive above 10^4 K. This impulsiveness reaches a maximum near 2×10^4 K, decreases rapidly at 3×10^4 K, but remains significant to approximately 2×10^4 K, above which the emission is no longer impulsive (Donnelly and Hall 1973). Secondly, Widing and Hiei (1984) showed that impulsive emission in one flare arose from a single loop, while gradual phase emission arose from a separate location. The impulsive phase in this flare was seen in hard X-rays (20-30 keV) and in transition zone lines between approximately 80,000 K and 2×10^6 K. The gradual phase emission from this flare, seen in soft X-rays and UV lines above 2×10^4 K, was observed to arise from a separate region or loop near the impulsive loop.

On the other hand, several arguments can be made against the assertion that impulsive and gradual emissions occur in separate volumes. Duijveman, Somov, and Spektor (1983) showed that one can see impulsive variations in bands as low as the 3.5-8 keV band of HXIS. Also, there is a well-known correlation between the amount of hard X-ray emission and the slope of the soft X-ray light curve in a typical flare soft X-ray emitting line (see, e.g., Tanaka et al., 1982a, Machado (1983). Finally, in some cases impulsive-phase phenomena occur in loops that already produce considerable thermal emission (see, e.g., Duijveman, Somov, and Spektor 1983); it is quite possible that the impulsive soft X-ray component is masked within the overall emission, as a spectrum such as Figure 3.14 would suggest. Only with substantially better spatial resolution will this issue be decided in a totally satisfactory manner.

In summary, the impulsive phase observations clearly show that there is a strong physical link between the processes that form HXR and transition zone and chromospheric emission. This link was established for flares in general by the relation in energies of peak emission found by Kane and Donnelly (1971), see Section 3.2.5.1. The simultaneity of individual peaks in HXR and OV in several flares found by Woodgate et al. (1983), and the relation in energy emitted during several flares as described by Poland et al. (1984), places observational limit on the possible links. There cannot be a significant time delay (i.e. greater than 1 s) between the different excitations, and the energies exciting the emissions are functionally related. That the emissions occur from
the same location to within 10" supports the concept of their being physically related, but this is not a serious constraint since this large distance could allow quite separate locations in the same loop or even different loops.

b. Interpretation. Emslie and Nagai (1984) have studied the response of a coronal loop to both heating and nonthermal electron acceleration at the loop top, in order to determine the resulting structure of the transition region plasma \( T = 10^{\pm 1} \) K, and so the emission measure in, and intensity of, optically thin emissions formed in this region. The key quantity in the modeling of Emslie and Nagai (1984) is the logarithmic differential emission measure

\[
\xi(T) = n^2 \frac{dz}{d\ln T}
\]

(3.5)

where \( n \) is plasma density, \( z \) vertical distance, and \( T \) electron temperature. Under an optically thin approximation for the radiative losses from the gas, this quantity is directly proportional to the intensity in a given line, the other factors being atomic in nature and therefore independent of the structure of the atmosphere (e.g., Pottasch 1964). Thus determination of the evolution of \( \xi(T) \) with time \( t \) gives a model prediction of the shape of the intensity versus time profile of the line emission being considered.

Figures 3.21 and 3.22 show the form of \( \xi(T) \) for various times since the onset of energy input into the flare loop for the conductively heated and electron-heated models, respectively. It is readily apparent from these figures that the evolution of \( \xi \) is markedly different in the two heating scenarios. Let us consider first the conductively heated model, in which the rate of energy deposition varies linearly with time, up to a maximum of \( 2 \times 10^{10} \) erg cm\(^{-2}\) s\(^{-1}\) after 30 seconds, and in which the energy deposition is spatially distributed over a Gaussian profile, symmetric about the loop apex, with dispersion \( \sigma = 2 \times 10^4 \) cm. In this model, there is virtually no change in \( \xi \) for the first few seconds, reflecting the fact that the thermal energy has not yet had time to propagate to the transition region layers. As soon as the heat front arrives (at \( t = 7 \) s for the parameters considered), a decrease in \( \xi \) results, due to the strong steepening of the temperature gradients produced by the large conductive flux \( F_c \propto T^{3/2} \frac{dT}{dz} \), which evidently overcomes the increase in the \( n^2 \) factor (Equation (5)) in the determination of \( \xi \). A few seconds later, however, the transition region plasma overheats to coronal temperatures, and the gas at the relevant temperatures (= 10\(^3\) K) is now situated much deeper in the atmosphere, with a corresponding larger density. The density scale height is sufficiently small in the preflare chromospheric layers that the conduction front, which moves at speeds of order of the local ion sound speed \( c_s = (kT/m_p)^{1/2} \), where \( k \) is Boltzmann’s constant and \( m_p \) the proton mass, covers the distance to the “new” transition region in a very short time. Thus \( \xi \) abruptly rises by several orders of magnitude in a second or two. This predicted trend in \( \xi(10^3 \) K) (and

![Figure 3.21](image-url)  

Figure 3.21 The logarithmic differential emission measure (DEM) \( \xi \), equation (3.5), for various times (in seconds) in the impulsive phase, for the conductively heated model. The dashed line shows the initial \( (t = 0) \) profile. Note the behavior at transition region temperatures—constant for the first few seconds, followed by a sudden decrease and an equally abrupt increase. The peak at around \( T = 10^7 \) K on the 8 s curve is real; it corresponds to the density enhancement that causes the large increase in \( \xi \) at transition region temperatures in the 10 s curve.

so in the intensity of optically thin EUV emissions) is simply not reflected in the observations above, implying that such an energy mechanism is not a viable possibility.

On the other hand, for the thick target electron-heated model, in which energy is distributed throughout the flare loop with an identical time structure but with the broader vertical spatial structure resulting from consideration of the collisional dynamics of an electron beam in a relatively cool target (see Emslie 1978), the strong steepening of temperature gradients associated with a relatively local energy released region does not happen. Instead, the dominant effect controlling the behavior of \( \xi \) is the progressive overheating of denser and denser layers to coronal temperatures,
resulting in a monotonic rise of the density at $T \approx 10^5$ K and so in $\xi$ ($10^3$ K).

This correlation of EUV intensity with energy input (as measured by the instantaneous hard X-ray flux, which is roughly proportional to injected electron energy (Brown 1971)) is in accord with the above observations (see, e.g., Poland et al., 1982, 1984; Woodgate et al., 1983). Indeed, Woodgate et al. (1983) have used the close synchronism of EUV and hard X-ray emissions to rule out conductively heated models such as that discussed above. The results of Emslie and Nagai further support this rejection. However, they also demonstrate that the observed behavior is predicted by the thick target electron heated model. On reflection, it is by no means obvious that this should be so. The observed increase in intensity is only an order of magnitude or so even in large events (Poland et al., 1984), a factor well below that predicted by the $n^2$ factor in Equation (5), since $n$ rises by over two orders of magnitude as a result of flare heating of the overlying layers (e.g., Machado et al., 1980, Emslie and Nagai 1984). Thus the increase in the temperature gradient factor must also play a role, and it is somewhat remarkable that these two large effects in opposite directions combine to produce a relatively small overall increase in $\xi$.

as is observed. The underlying physical reasons for this behavior are not fully clear at present, and merit further investigation.

### 3.2.6 White Light Emission

We define white light flares as those detected in the optical continuum; thus their number is limited by observational considerations. In the last few years patrols at the National Solar Observatory, Sacramento Peak (NSO/SP) and Big Bear Solar Observatory (BBSO) have greatly increased the known number [see catalog of white light flares by Neidig and Cliver (1983)]. It is probable that all major flares have detectable continuum in the blue, although the number detected at longer wavelengths is still small. Thus the term ‘white light flares’ simply means highly energetic flares. Our general motivation in studying white light flares is based on the great energy involved, the problem in transporting energy to the low atmospheric heights of significant continuum opacity, and a general fascination with intense flares. An important special interest here is in the nature of the continuum emission and how it may be produced.

#### 3.2.6.1 Morphology

Our knowledge of the morphology of white light flares is limited by the small number of events for which good data are available. Nevertheless, four kinds of white light flare have been recorded: (1) Bright footpoints, usually at the edge of the sunspot umbra (Zirin and Neidig 1981); (2) Short-lived flashes (Zirin and Tanaka 1981) associated with impulsive X-ray spikes, up to a few arcsec diameter and 20 s lifetime, usually located along the neutral line of a delta configuration; (3) Moving fronts (Machado and Rust 1974; Zirin and Tanaka 1973) occurring after the impulsive phase and possibly associated with the hot thermal plasma; (4) Fixed bright points or ribbons (Zirin and Neidig 1981; Neidig and Cliver 1983), possibly associated with the hot thermal phase.

Except for the last, all the manifestations are moving, sometimes very fast, or jumping from point to point. The white light flare footpoints commonly appear at a number of locations and times throughout the whole event; presumably they are associated with the main bundle of field lines on which flare energy input occurs. Almost always at least one footpoint occurs in the penumbra, usually at the edge of the umbra, but other bright points may occur outside the sunspot, as do white light flare ribbons. The area of white light flare emission averages $\sim 6 \times 10^{17}$ cm$^2$ at flare maximum, with a duration ordinarily less than 10 minutes (Neidig and Cliver 1983).

The white-light kernels and fronts roughly match the emission observed in He D3 and the wings of Hα. The He D3 emission, which may cover a somewhat larger area than the associated white-light emission, becomes especially bright in the post-impulsive phase (for example, the July 1 flare...
(Zirin and Neidig 1981)). Probably all white light flares are accompanied by D3 emission.

Very little information is available regarding the white light flare relationship with spatially-resolved, high-energy emissions. In the June 6, 1982 flare (Ohki et al., in preparation) the Hinotori hard X-ray and soft X-ray images generally show spatial and temporal agreement with kernels of emission at \( \lambda 3862 \, \text{Å} \). Videotapes show small kernels a few arcsec across, with changes occurring on time scales of a few seconds.

3.2.6.2 The Optical Spectrum

**a. Observations.** Spectral data on white light flares are usually obtained either from a series of broad band filtergrams at several widely spaced points in the spectrum (e.g., Zirin and Neidig 1981) or from spectrograms with dispersions of several Å/mm (e.g., Neidig 1983). In filtergrams, the variation of flare intensity with wavelength is affected by changes in seeing between the successive images, or even by temporal changes in the flare itself. Fast switching with video measurements removes some of these problems, and the seeing may be evaluated.

Another problem in interpreting the filtergram measurements is the possible effect of lines in the blue. Although the Balmer line emission can be minimized by using, for example, a narrow (15 Å) filter at 3862 Å (midway between \( \text{H}_\alpha \) and \( \text{H}_\beta \)), there still remains the possibility of emission by numerous photospheric lines (Hiei 1982).

The spectrograms, although able to record a large expanse of the spectrum in a single exposure, are typically obtained with lower spatial resolution, with the result that the flare emission may be highly diluted or the most intense kernels may not fall on the slit. Judging by the intensities observed, the spectra obtained so far have not recorded the most intense kernels. Intensities up to twice the continuum at visible wavelengths have been recorded with filters, while the spectral data have only revealed enhancements of 10-20% above the continuum. In addition, the variation of intensity with wavelength, as determined from spectrograms, may suffer from systematic errors due to chromatic aberration of atmospheric or instrumental origin. Thus while the gross features of the white light flare spectrum, the line intensities, and the Balmer jump can reliably be obtained from spectra, the slope of the continuum thus obtained, which might be used to derive the flare temperature or to identify the emission mechanism, should be treated with caution.

The filter measurements, as well as some of the spectrograms, show a fairly flat continuum in the visible with an increase below 4000 Å. Superposed on this there is often a relatively strong Balmer continuum. Figure 3.23 combines filter data on several flares; the data for July 1 and April 24 are from photographic measurements; that for May 9 is from high speed video measurements. Figure 3.24 shows spectrographic data in the vicinity of the Balmer jump for three representative flares.

The Balmer continuum has been identified in spectrograms of six events (Hiei 1982; Hiei et al., 1982; Neidig and Wiborg 1984; Donati-Falchi et al., 1984), although several cases are known (e.g., Machado and Rust 1984; Boyer et al., 1985) in which a short wavelength continuum was present, but without a measurable Balmer jump. A Paschen jump was detected in the spectrum of the 24 April white light flare (Neidig and Wiborg 1984); unfortunately, the latter flare is the only known event with an infrared continuum sufficiently strong to permit a conclusive search for the Paschen jump.

In addition to the large intensities observed in the Balmer continuum, the increase in the flare contrast in the blue, especially below 4000 Å, can be quite striking in some cases (Zirin 1980, Zirin and Neidig 1981). This “blue continuum” is probably related to the bluish color noted in a number of visual observations of white light flares, although the cause for it is not understood. An increase in the flare contrast could be expected simply as a result in the decline in intensity of the solar background at short wavelengths (this would be particularly evident in the broad band filtergrams where the solar background is blanketed by lines). On the
Figure 3.24 Spectrographic measurements for three white light flares: Apr. 24, 1981 (Neidig 1983); Sept. 10, 1974 (Hiei 1982); Aug. 7, 1972 (Machado and Rust 1974). I_o refers to the intensity in the bright windows of the spectrum adjacent to the flare.

On the other hand, intense brightenings ($\Delta I/I_o = 1$) have been observed in filtergrams at 3835 or 3862 Å when the corresponding emission at longer wavelengths was weak or absent (Zirin 1980). The later observations indicate differences in contrast too large to be attributed entirely to differences in the intensity of the solar background. The question as to whether the blue continuum is truly anomalous, arising, perhaps, from some unidentified source of opacity or from photospheric lines, cannot be answered until more precise observations of the spectrum become available.

b. Interpretation. The principal task in interpreting the spectral observations is to identify the emission mechanism; this, in turn, determines the general atmospheric regime of the white light flare phenomena. The spectra in Figure 3.23 suggest a strong variation with wavelength in the opacity of the white light flare source, indicating an origin in layers that are not optically deep. The Balmer continuum (Figure 3.24), as well as the bright He D3 emission associated with white light flares, is characteristic of chromospheric densities (electron density $1 - 5 \times 10^{13}$ cm$^{-3}$; see Machado and Rust 1974; Neidig and Wiborg 1984; Zirin 1983). Nevertheless, it is by no means certain that the emission underlying the Balmer continuum and extending to longer wavelengths is due to H$_{\beta}$ emission in general (the 24 April flare, which shows a Paschen jump, is a good candidate for an event dominated in the visible and infrared by H$_{\beta}$ emission). Hiei (1982) has pointed out that the continuum extending to longer wavelengths might be attributable to H$^-$ emission, possibly originating in the photosphere; this hypothesis remains untested because information on the Paschen jump is not available for the flares studied by Hiei.

White light flares in which the Balmer jump is absent (e.g., August 7, in Figure 3.24) might suggest either an optically deep source with small temperature enhancement or an optically thin, very hot source with opacity dominated by free-free absorption. The latter possibility, however, might require a prohibitively large emission measure at high temperature ($n_e^2 \Delta z \geq 10^{35}$ cm$^{-5}$ at $10^5$ K, in order to produce a 10% enhancement at 4000 Å). The spectrum of the August 7 flare was studied by Machado and Rust (1974) who concluded that the emission was Paschen continuum originating in the flare chromosphere. The latter interpretation, of course, would require Balmer continuum to be present in this flare, and therefore, unless the absence of the Balmer jump can be explained, a deeper, low-temperature source with H$^-$ opacity might be indicated. In any case the height of the source in the August 7 flare would have to be greater than 300 km above $\tau_{5000} = 1$, as Machado and Rust found a probable absence of emission in lines formed below that level. With regard to a deep photospheric origin for white light flares in general, it is worth noting that strong reversals in the cores of photospheric absorption lines have not been observed in white light flare spectra.

Recent model calculations by Aboudaram et al. (1984) show that the white light flare does not arise from heating at $\tau_{5000} > 1$, because the predicted $\Delta I/I_o$ would then vary as $1/\lambda$, in contrast to the observations in Figure 3.23. Aboudaram et al., also show that the continuum cannot be explained by heating of the upper photosphere and temperature minimum region; such models produce $\Delta I/I_o$ increas-
ing with wavelength in the visible and, in particular, they do not reproduce the observed increase in contrast below 4000 Å.

Although the spectral data described here are fragmented and sometimes ambiguous, we note that the brightest white light flares observed with spectrographs (e.g., Donati-Falchi et al., 1984; Neidig and Wiborg 1984), as well as as the events observed with filters at several wavelengths, show strong Balmer continuum which must arise in a chromospheric-like regime.

In order to estimate the temperature in white light flares it is first required to know how the opacity varies with wavelength. But even if the latter is known, the interpretation will be limited by our incomplete knowledge of the optical thickness, departures from LTE, and by possible systematic errors in the data. If we assume Hβ emission, where the opacity varies as λ3, the observed slope and intensity of the continuum in the visible region can satisfactorily be fitted to chromospheric temperatures and densities by making appropriate (although arbitrary) choices for geometric thickness and departures from LTE. In this way the white light flare continuum observed by Machado and Rust (1974) yielded temperatures of 8500-20,000 K. Similarly, the April 24 flare (well-observed with broad-band filters) can be fitted to 10,000-20,000 K. On the other hand, if H− contributes to the white light flare continuum, then we might expect T_e ≤ 8000 K for the layers in which this emission arises.

The temperature estimates discussed here do not allow for effects due to possible unknown sources of opacity in the blue. If such opacity is present, no further progress can be made until these sources are identified. Finally, we again remind the reader that the spectrograms obtained thus far may not accurately represent the brightest kernels of white light flares; hence a precise description of the brightest white light flare spectra, including possible contribution by line emission, remains in question.

3.2.6.3 Timing Relationships

In Figure 3.25 we compare the hard X-ray (HXR) burst profiles with the total power radiated in the white-light continuum for six white light flares. One event (Feb. 11, 1980) was obtained from Big Bear Solar Observatory (BBSO) at a single wavelength (4306 Å); one event (July 1, 1980) is based on data from both BBSO (λ 3862 Å) and NSO/SP (λλ 3610, 4275, 4957, 5650 and 6203 Å); the remaining four events use NSO/SP data at five wavelengths. In computing the total power it is assumed that the white light flare emission extends over the range 2500-10,000 Å. Time resolution varies from 15 s to 30 s except for the July 1 event (Zirin and Neidig 1981) where the resolution was 90 s in the five wavelengths of the NSO/SP data and 15 s in the BBSO data. In the case of the February 11 event we assume that the white light flare spectrum is qualitatively similar to the spectra in Figure 3.23.

The most striking feature of Figure 3.25 is the overall similarity between the white light and HXR. With the exception of the late phase of the July 1 flare, the correlations suggest that the HXR and white-light sources may be closely related, especially in the impulsive phase. This result confirms the conclusions of an earlier study (Rust and Hegwer 1975) which compared HXR and white-light intensity data at two wavelengths (4950 and 5900 Å).

The most unfavorable of the six comparisons in Figure 3.25 is the July 1 flare. In this case there is good agreement between hard X-rays and white light in the impulsive phase, but the brightest white light flare point occurred later and approximately coincided with the peak thermal energy content (Antonucci 1981) and Fe XXV emission (Phillips 1981). This suggests an association with the hot (T ≈ 2×10^6 K) thermal plasma in the flare. Similarly, Zirin and Tanaka (1973) observed a white light flare wave event in the hot thermal phase of the August 2, 1972 flare.

In summary, the HXR and white light are well-correlated in most cases, while in others the white light rises with the HXR and declines with the hot thermal plasma. In this sense the behavior of the white light in the impulsive and thermal phases bears some similarity to the rise and fall of Hα emission relative to hard and soft X-rays in non-white light flares (Zirin 1978).

3.2.6.4 Energetics

The principal datum that motivates our study of white light flares is the large observed continuum flux. For nine flares the mean peak flux is estimated to be 1.5×10^16 ergs cm^{-2} s^{-1} between 2500 and 10,000 Å (Neidig and Cliver 1983), with the largest reported flux for a single event being 5×10^16 erg s^{-1} cm^{-2} (Slonim and Korobova 1975). Allowing for the low spatial resolution of these observations, we suggest, for purposes of model calculations, that a reasonable upper limit on the flux is 10^17 ergs cm^{-2} s^{-1}. The peak continuum flux far exceeds the combined flux of all the Balmer lines at the same location in the flare, although the line emission, being more widespread, may be comparable in total energy when integrated over time and space. The total peak continuum power, obtained by integrating over the white light flare area, is less sensitive to resolution, and is approximately 10^{28} erg s^{-1} for the largest white light flares.

These numbers are quite impressive, and they place severe constraints on the energy transport mechanisms in white light flares. A number of transport processes have been suggested in the literature, although none can be claimed to be satisfactory in general. Chromospheric heating by fast electrons has been a long-time favorite mechanism of modelers, and we shall demonstrate here that this mechanism might
work in the impulsive phases of two well-observed white light flares, but fails in the post-impulsive (hot thermal) phase of one of these flares.

For both the impulsive and post-impulsive phases of the July 1 and April 24 flares, we list in Table 3.3: (1) the observed peak power in white light, (2) the electron energy $E_0$ above which the power in the nonthermal electron spectrum (in a thick target approximation) is equal to the white-light power, (3) the stopping height $h(E_0)$ of an electron with energy $E_0$ in the non-flaring atmosphere, (4) the column density $n(z)$ of ionized gas that can be penetrated by an electron with energy $E_0$, and (5) the estimated column density $n(z)$ required to produce the observed white-light intensity via $H_\alpha$ emission at an assumed temperature of $10^4$ K and a density equal to the density in the non-flaring atmosphere at the electron stopping height. The nonthermal electron data are derived from Dennis (1981), Batchelor (1984), Kane (1983), and Kane et al. (1985). We see that in the impulsive phase of both flares sufficient power in nonthermal electrons might, within limits of observational uncertainties, be transported through a column comparable in extent to that required to produce the amount of white-light emission. The only problem is producing the observed white light flare spectrum. Sufficient power is marginally available also in the post-impulsive phase of the 24 April flare, although for the post-impulsive phase of the July 1 flare the nonthermal electron model seems to fail completely.

Heating by $E \geq 1$ Mev protons in the impulsive phase white light flare was suggested by Najita and Orrall (1970) and Svestka (1970). The protons certainly have enough range, and the calculations of Lin and Hudson (1976) show that, at least in the August 4, 1972, flare, the proton flux alone was probably adequate to power a typical white light flare at column depths up to $10^{21}$ cm$^{-2}$. More recent work...
Table 3.3 White Light Flare and Energetic Electron Transport Parameters

<table>
<thead>
<tr>
<th>Date</th>
<th>WL Power</th>
<th>Impulsive Phase</th>
<th>Post-Impulsive (Thermal Phase)</th>
</tr>
</thead>
<tbody>
<tr>
<td>July 1, 1980</td>
<td>WL Power</td>
<td>2.3 × 10^{27} ergs s^{-1}</td>
<td>4.5 × 10^{27} erg s^{-1}</td>
</tr>
<tr>
<td></td>
<td>E₀</td>
<td>53 keV</td>
<td>35 keV</td>
</tr>
<tr>
<td></td>
<td>h(E₀)</td>
<td>880 km</td>
<td>1020 km</td>
</tr>
<tr>
<td></td>
<td>nz(E₀)</td>
<td>1.2 × 10^{21} cm^{-2}</td>
<td>5.2 × 10^{20} cm^{-2}</td>
</tr>
<tr>
<td></td>
<td>nz(WL)</td>
<td>3.9 × 10^{21} cm^{-2}</td>
<td>2.1 × 10^{22} cm^{-2}</td>
</tr>
<tr>
<td>Apr. 24, 1981</td>
<td>WL Power</td>
<td>1.4 × 10^{27} ergs s^{-1}</td>
<td>7.0 × 10^{27} s^{-1}</td>
</tr>
<tr>
<td></td>
<td>E₀</td>
<td>119 keV</td>
<td>85 keV</td>
</tr>
<tr>
<td></td>
<td>h(E₀)</td>
<td>650 km</td>
<td>730 km</td>
</tr>
<tr>
<td></td>
<td>nz(E₀)</td>
<td>6.0 × 10^{21} cm^{-2}</td>
<td>3.1 × 10^{21} cm^{-2}</td>
</tr>
<tr>
<td></td>
<td>nz(WL)</td>
<td>3.3 × 10^{20} cm^{-2}</td>
<td>3.3 × 10^{21} cm^{-2}</td>
</tr>
</tbody>
</table>

by Hudson and Dwivedi (1982) has shown that sufficient heating by protons may be possible even in the upper photosphere (column depth 5×10^{22} cm^{-2}), but that the heating at depths near \( \tau_{2000} = 1 \) (column depth 2×10^{24} cm^{-2}) is negligible. Lin and Hudson originally dismissed the proton heating mechanism because of an apparent delay in the 2.2 MeV \( \gamma \)-ray lines relative to the flare impulsive phase. But recent SMM data shows no delay in the energetic protons that produce (via collisions) the free neutrons responsible for the 2.2 MeV \( \gamma \)-ray line, so the delay is not a problem. This conclusion is corroborated by Ryan et al. (1983) who show, in addition, that while proton heating in the impulsive phase of the July 1 flare might have been possible, the proton event had ended before the post-impulsive white-light emission reached maximum. Thus, we conclude that heating by non-thermal electrons or protons cannot account for the post-impulsive phase white light flare.

A more recently investigated mechanism for the white light flare is a two-stage transport process. We know that sufficient energy is available in E \( \geq 35 \) keV electrons even in the post impulsive phase. These electrons are stopped near the 1000 km level, and the problem then is to efficiently transport the energy to regions of higher density to produce the optical continuum. Livshitz et al. (1981) have concluded that a shock wave may propagate downward from the heated chromospheric region with velocity 1000 km s^{-1}, and that this could produce heating at lower levels where the white light flare is seen. The principal difficulty here is the relatively short duration of the shock-induced effect (~ 10s), which is in contrast to observed white light flare durations as long as 10 minutes for a single bright kernel [see Figure 3.25, and Kane et al. (1985)].

The post-impulsive white light flare emission in the July 1 flare seems to be associated with the hot thermal phase of the event. We might be led, therefore, to consider heat conduction as a transport mechanism, were it not for the fact that sufficient conductive flux can be obtained only at temperatures of millions of degrees in the white light flare source. Clearly, this is incompatible with a chromospheric origin for the white light flare; moreover, it requires an emission measure several orders of magnitude larger than is observed from soft X-ray data. Thus we are unable to identify a satisfactory mechanism for the energy transport in post-impulsive phase white light flares.

3.2.7 \( H\alpha \) Emission

Study of the morphology, timing and spectrum of \( H\alpha \) (\( T = 10^4 \) K) emission during the impulsive phase, relative to hard X-rays and microwaves, bears on several basic questions of flare energy transport. Do flare heating effects occur nearly simultaneously (say within 1 s) throughout loops, down to \( H\alpha \) formation depths? Do fast electrons penetrate as far as the chromosphere? Is the observed depth dependence of flare heating in the chromosphere the dependence predicted by stopping of nonthermal electrons? Since the Skylab era much progress has been made at both observational and theoretical approaches to these questions.

3.2.7.1 Observations

During the recent solar maximum the usefulness of \( H\alpha \) observations was increased markedly by improvements in observing techniques that ensured, for the first time, during the impulsive phase, observations with high temporal, spatial and spectral resolution, coordinated with temporally, spatially, and spectrally resolved X-ray and microwave data.

a. Timing. The temporal resolution and precision of measurement in \( H\alpha \) and either hard X-rays or microwaves required to discriminate between conduction front and energetic particle models of the transport of energy from the corona to the chromosphere during the impulsive phase is...
approximately 1 s. It is also necessary to have spatial resolution, since not all parts of the chromospheric flare necessarily reflect the same transport process. In the past, the lack of adequate temporal resolution led to much confusion about the time delays between hard X-rays (or microwaves) and Hα (cf. Vorpahl 1972; Zirin 1978).

The first observations with sufficient resolution to be compelling have been obtained recently by Kaempfer and Schoechlin (1982) and Kaempfer and Magun (1983). In a study of one flare with 1.4 s and 100 ms temporal resolution in Hα and microwaves respectively, Kaempfer and Magun found evidence for both fast electron transport, at one site of a flare, and hydrodynamic or nonclassical conductive transport, at other sites of the same flare. At the former site, they observed Hα and microwave synchronization to within two seconds; the lack of delays at different microwave frequencies also supported an energetic electron interpretation. At the latter sites, delays of about 10 s were observed. The authors show that these delays are consistent with the propagation of disturbances at about 2000 km s⁻¹, i.e. roughly the velocities expected for collisionless conduction fronts (Brown, Melrose and Spicer 1979).

b. Morphology. The spatial coincidence of Hα and hard X-ray emission during the impulsive phase of Type B flares supports the idea that both emissions often arise as a consequence of the transport of coronally accelerated nonthermal electrons to the footpoints of flare loops (see Section 3.2.2). Although this conclusion applies only to impulsive hard X-ray flares, not gradual ones, there is no controversy over this spatial relationship in the horizontal plane.

On the other hand, there has been a continuing controversy over the height of origin of impulsive chromospheric flare emission. On one side, Zirin (1978) has argued that impulsive Hα emission is produced by direct collisional ionization and recombination in an elevated source—prominence material—well above the quiet chromosphere. This belief is based on line center Hα filtergrams at and near the limb, during hard X-ray bursts. Zirin's observations show several events in which impulsive phase Hα brightening is seen in structures that typically extend 10⁴ km above the limb. On the other side, theoretical models all show that the Hα emission of flares comes about primarily thermally, from deeper than the quiet chromosphere (see, e.g., Brown, Canfield and Robertson 1978 or Canfield, Gunkler and Ricchiazzi 1984). The latter picture is also what is inferred for the primarily chromospheric source of impulsive 10-1030 Å EUV emission (Donnelly and Kane 1978), and is discussed in Section 3.2.5).

A recent observation of a limb flare, by Kurokawa (1983), bears on the controversy. It shows that flare emission sites observed at Hα line center ± 2.4 Å in a limb flare show abrupt intensity changes directly correlated with microwave fluctuations. It also shows that these emission sites are brightest at Hα ± 1.2 Å, where they appear about 1800 km above the photosphere, but are obscured by the quiet chromosphere at Hα line center. Kurokawa (1983) concluded that the main part of the flare chromosphere is confined to the low chromosphere, much lower than the surrounding undisturbed structures.

c. Spectra. Whereas before the recent solar maximum there were observations of spectral line profiles from flare kernels during the impulsive phase of perhaps one flare (Zirin and Tanaka 1973), there are now observations of hundreds. Many of these are accompanied by X-ray and microwave observations. Such observations of the Hα line (Acton et al., 1982, Gunkler et al., 1984, Canfield and Gunkler 1984) provide strong evidence for the propagation of nonthermal electrons down into the chromosphere during the impulsive phase.

Figure 3.26, from Canfield and Gunkler (1984), shows Hα line profile evidence for chromospheric heating by nonthermal electrons during the flare of 7 May 1980. Two of the times shown, 14:56:16 and 14:56:42 UT, are during the impulsive phase, and the hard X-ray emission at all energies above 30 keV was 3-4 times stronger at the second time. There were two distinct kernels in this flare, shown by the cross-hatching in Figure 3.26; 2-3 times more 16-30 keV emission was observed to come from the south (lower) kernel. During the time of impulsive hard X-ray emission, strong Hα wings develop. They disappear very quickly after the impulsive emission ends. The number of pixels that show such behavior rises and falls in direct proportion to the intensity of the hard X-ray emission.

3.2.7.2 Interpretation of Spectra

Since fast electrons lose several orders of magnitude more energy to heating than to bremsstrahlung X-ray production, one would expect to be able to see clear effects in thermally generated chromospheric emission. Ricchiazzi and Canfield (1983) recently developed static models of the effects of enhanced fluxes of nonthermal electrons, treated strictly collisionally, as well as thermal conduction and enhanced pressure from the flare corona, on the temperature, density, and ionization structure of the chromosphere. Their work assumed that the electrons last long enough to establish a new hydrostatic equilibrium in a closed loop (at least tens of seconds). Subsequent work, in an impulsive approximation, by Canfield, Gunkler and Ricchiazzi (1984), henceforth CGR, assumed that very little time (at most a few seconds) has passed since the start of electron heating, so no mass motions have had time to develop. These models superseded work carried out in the Skylab era by Brown, Canfield and Robertson (1978).

By using these model flare chromospheres to compute profiles of optically thick spectral lines, which are formed
Figure 3.26 Impulsive-phase Hα profile observations. The two large contiguous squares framing each of the 6 panels indicate two 8" × 8" HXIS pixels; the smaller rectangles indicate 2" × 2.67" Hα pixels. The shaded Hα pixels indicate the north and south kernels. In each Hα pixel its spectrum at the indicated time I(Δλ) is plotted in units of $I_0$, the observed quiet sun continuum intensity near the flare site. Each pixel's preimpulsive phase spectrum (145525 UT) is shown dotted, for comparison. From Canfield and Gunkler (1984).
over a range of depths along the direction of propagation of the electrons, one can predict theoretical line profile signatures of the distribution of heating appropriate to fast electrons. CGR calculated Hα line profiles based on both hydrostatic and impulsive model chromospheres. They showed that both hydrostatic and impulsive model atmospheres heated by nonthermal (power-law) electrons characteristically produce wide and bright Hα profiles with a central reversal. Enhanced thermal conduction reduces the width and total intensity of the profiles. High thermal conduction alone cannot account for flare Hα enhancements. High coronal pressure dramatically increases the width and total intensity of the Hα profiles, while reducing the central reversal.

CGR found that only high values of the flux of nonthermal electrons produced Hα profiles with obvious broad (Stark) wings. Their results are shown in Figure 3.27. In the upper panel, Hα profiles are shown for the impulsive model chromospheres; those in the lower panel are for hydrostatic models. In both panels it can be seen that pronounced wings develop only for values of electron energy flux (above 20 keV) above about 10^{10} ergs cm^{-2} s^{-1}. The calculations indicate that no combination of values of thermal conduction and coronal pressure can give the same strong Stark wings, accompanied by central reversal, as those associated with nonthermal electron heating.

The observations cited above show that Hα profiles develop obvious broad wings only in spatial and temporal coincidence with hard X-ray emission. Gunkler et al. (1984) and Canfield and Gunkler (1984) found, by quantitatively comparing the extent of observed Hα wings to the theoretical profiles of Figure 3.27, that values of F_{20} \approx 10^{11} ergs cm^{-2} s^{-1} were implied. They showed that this same value was implied by combining the power of electrons above 20 keV inferred from the HXRBs observations with the observed H\alpha stark-wing area. From the point of view of transport theory, the implication of this result is that the electrons that produce impulsive hard X-rays indeed have a range that is in at least order-of-magnitude agreement with the predictions of a strictly collisionless thick-target model of electron propagation. The fact that this inferred value of F_{20} is so large (approaching the return-current stability limit) is interesting. This is not the first time such large electron fluxes have been found from chromospheric spectra; using an inversion method, F_{20} \approx 10^{12} ergs cm^{-2} s^{-1} was inferred from semiempirical flare chromospheric models (Machado et al., 1980) by Emslie, Brown and Machado (1981).

3.3 THEORETICAL STUDIES OF TRANSPORT PROCESSES

3.3.1 Electron Beams and Reverse Currents

The existence of large numbers of nonthermal electrons in solar flares has been supposed for some time now as means of explaining, through the mechanism of electron-proton bremsstrahlung, the high flux of hard X-rays observed in large events (see e.g., Brown 1971). With observations (Tanaka, Nitta, and Watanabe 1984) now showing the hard X-ray continuum (i.e. line-subtracted) spectrum extending in an unbroken power-law down to energies \leq 7 keV, it appears that, at least in some events, suprathermal bremsstrahlung-producing electrons exist at all energies above this value. This implies that an extremely large number of these electrons must be accelerated in the primary release process, whether through directed acceleration ("nonthermal" or "beamed" model) or bulk energization ("thermal" model). In particular, in the nonthermal model such hard X-ray observations lead to nonthermal electron injection rates \geq 10^{36} s^{-1} above 25 keV (Hoyng, Brown and van Beek 1976) and so, by extrapolation of typically observed power-laws, \geq 10^{38} above \sim 7 keV.

One of the major theoretical problems of the last decade has been to describe quantitatively the energy, momentum and charge transport effected by such strong beams. Early treatments (e.g., Brown 1972, 1973; Emslie 1978) neglected collective effects within the electrons comprising the beam, and instead concentrated on the mean behavior of individual electrons through Coulomb interactions with the ambient plasma particles. It was soon recognized, however, (Hoyng, Brown and van Beek 1976; Knight and Sturrock 1977; Hoyng, Knight and Spicer 1978) that the passage of such a large number of electrons through the stationary solar atmosphere constituted an extremely large current. This required that a reverse (or return) current be set up in the ambient plasma, in order to achieve both charge and current neutrality, and to provide fresh electrons to the acceleration site. We refer the reader to Knight and Sturrock (1977) for a discussion of the unacceptably large electrostatic and magnetic fields produced in the absence of a reverse current, and to Hoyng, Brown and van Beek (1976) for a discussion of the "electron number" problem associated with the depletion of electrons from the coronal acceleration site. This return current must be driven through a medium of finite resistivity, modifying (by an as yet unresolved amount—see below) the energetics of the primary electron beam. In addition, other collective effects, such as two-stream instabilities and collisionless pitch-angle dispersion, modify the dynamics of such beams, with implications for both their hard X-ray production and atmospheric heating. Finally, the beam may have yet another significant consequence beyond those of photon production and heating; the stopping of the beam may effect significant acceleration of the coronal and chromospheric plasma through which it passes. This section discusses the following aspects of nonthermal electron beams: the energetics of, and driving mechanism for, the return current, and collective instabilities driven by the beam itself. The motivation for such in-depth study of nonthermal electron beams is provided by recent observations from SMM and other spacecraft, which indicate a substantial nonther-
Figure 3.27 The effect on Hα line profiles of varying the energy flux of nonthermal electrons above 20 keV, $F_{20}$. The electron beam is assumed to be vertical, and the spectral index of the power-law electron energy distribution function is $\delta = 5$. Intensity is measured in units of the quiet sun preflare continuum near Hα. $F_s$ is the conductive flux from the corona, measured at $T = 10^6$ K. $P_0$ is the gas pressure at the top of the coronal loop, the assumed acceleration site. Top: impulsive model. Bottom: hydrostatic model. From Canfield, Gunkler and Ricchiazzi (1984).

3.3.1.1 Energetics of the Reverse Current

Emslie (1980) showed that the ohmic energy losses sustained in driving the reverse current through the finite resistivity of the ambient plasma could, in some cases, exceed the losses suffered through direct collisions of the beam electrons with ambient ones. For classical resistivity, the ratio of ohmic to collisional heating scales like $F \propto n^{-1} T_e^{-3/2}$, where $F$ is the injected electron flux, $n$ the ambient density and $T_e$ the electron temperature (Emslie 1980), and hence only for relatively large fluxes of electrons passing through relatively cool coronae could this ohmic heating dominate. This raises the question of the self-consistency of Emslie’s (1980) treatment through considerations of energy balance and, perhaps more fundamentally, through the requirement that the electron drift velocity associated with the reverse current be below the threshold for plasma instabilities to develop, typically around the ion-sound speed $c_s = (k_B T_e/m_p)^{1/2}$, where $k_B$ is Boltzmann’s constant and $m_p$ the proton mass (Fried and Gould 1961). These matters were addressed independently by Emslie (1981) and by Brown and Hayward (1982). Emslie (1981) considered the flux limitation imposed by the stability requirement on the reverse current drift velocity. He showed that this could reduce the importance of reverse current ohmic heating, except during the very early stages of the event, when low preflare coronal temperatures result in a very large plasma resistivity (see also Knight and Sturrock 1977), and also during the late phase of long events ($t \gtrsim 10 \, s$), by which time the ambient protons could be significantly heated by the hot electrons, resulting in an equilibration of ion and electron temperatures, an increased threshold drift velocity (Fried and Gould 1961) and so a larger possible ohmic to collisional energy loss ratio. Brown and Hayward (1982), see also Hayward (1984) considered the other aspect of the self-consistency problem, namely the determination of the ambient plasma electron temperature $T_e$ through consideration of the energy equation. Including only dominant terms, this can be expressed as

$$\eta j^2 + \nabla \cdot F_c = 0,$$

where $\eta$ is the plasma resistivity, $j$ the beam (or reverse) current density, and $F_c$ the thermal conductive flux. Equation (3.6) assumes that direct collisional losses are unimportant, an assumption that can be tested a posteriori by comparing the ohmic losses in a plasma whose temperature follows from Equation (3.6) with the collisional losses in the same plasma. The collisional losses in turn depend on the density $n$ and the average individual electron energy $E$. Thus the ratio of ohmic energy losses to collisional ones depends on $F$, $E$, $T_e$, and $n$. In addition, from the values of $n$ and $T_e$ [found from solution of Equation (3.6)], we can test whether the beam flux exceeds the stability threshold $n_{\text{crit}}$, where $v_{\text{crit}}$ depends solely on $T_e$. From these considerations, Brown and Hayward (1982) deduced that when the reverse current is stable to the generation of plasma turbulence it is unimportant energetically; the effects of an unstable reverse current are yet to be determined.

Emslie (1985) has recently contested the Brown and Hayward (1982) results, on the basis of the inadequacy of their treatment of the conduction term in (3.6) as $x_o T_e^{-5/2}/I_{\text{es}}$, where $x_o$ is the Spitzer (1962) coefficient of thermal conductivity and $I_{\text{es}}$ the stopping distance of the beam electrons under the decelerating force produced by the charge separation electric field $E = n j$. He argued that the boundary condition determining the point at which the electrons stop is not in fact $I_{\text{es}}$, but instead the half-length $L$ of the flare loop (beyond which Coulomb collisions in the high density chromosphere effectively stop the beam in a very short distance), and therefore recast Equation (3.6) in the form

$$\eta_0 T_e^{-3/2} j^2 + d/dz \left( x_o T_e^{5/2} dT_e/dz \right) = 0,$$

imposing a boundary condition on $T_e$ at a fixed distance $L$ corresponding to the half-length of the flare loop. Using this revised version of the energy equation, and also a more realistic power-law-type injected electron energy distribution, Emslie (1985) found, in contrast to the results of Brown and Hayward (1982), that a regime (albeit a small one in the available parameter space) where a stable reverse current dominates the energetics of the upper coronal plasma can exist. Physically, this is because the shorter and more physical boundary values of $L$ imposed by Emslie result in a much lower apex temperature; since the resistive losses for a stable reverse current scale as $T_e^{-3/2}$ and the drift velocity stability threshold scales as $T_e^{1/2}$, a reduction in this peak temperature can have relatively large effect on the importance of ohmic heating, while having a relatively small effect on the maximum beam flux allowed to pass stably. For low densities, this stable reverse current heating is dominant, while for high densities collisional heating dominates. Thus there is a minimum energy input $\epsilon = \max (\epsilon_{\text{collisions}}, \epsilon_{\text{reverse current}})$ into the coronal plasma, and, in turn, a minimum peak coronal temperature for a given hard X-ray flux, independent of coronal density. This relationship between hard X-ray burst intensity and peak coronal temperature (deduced from soft X-ray observations) in different events may afford observational tests of (a) the thick-target model and (b) the relative role of reverse current and collisional heating in such a model.

3.3.1.2 Driving Mechanism

A fundamental issue for the physics of nonthermal electron beams is whether the reverse current is established elec-
trostatically or inductively. Spicer (1982) and Spicer and Sudan (1984) have claimed that an assumption made by all previous authors, namely that the reverse current is in steady-state balance with the injected beam flux, and that it is established by the electrostatic field associated with the charge separation effect by the beam, is incorrect. Rather, they argued that the unneutralized current of the beam sets up, through Ampere’s law, a self-magnetic field \( B \) which then establishes a return current inductively through Lenz’s law. In addition, they argued that this return current decays resistively in the ambient plasma, so that eventually it can no longer effectively neutralize the beam; the resulting “bare beam” cannot propagate (for the reasons discussed by Knight and Sturrock 1977) and so turns off. They suggested that this may be the cause of the extremely rapid decays found in some hard X-rays events (Kiplinger et al., 1983). Brown and Bingham (1984) have criticized Spicer and Sudan’s work on the grounds that they neglected the displacement current term in Ampere’s law at the head of the beam. The inclusion of this term, according to Brown and Bingham, returns us to a scenario where the reverse current is indeed set up electrostatically, over a Debye length scale at the head of the beam; since the time required for the beam to traverse such a small distance is much smaller than timescales for hard X-ray bursts, they argued that a steady-state description (Knight and Sturrock 1977; Emslie 1980) is indeed valid. With regard to the resistive decay of the current, they agreed with Spicer and Sudan that the finite curl of the \( E \) field (due to edge effects at the outside of the beam) leads to the establishment of an azimuthal magnetic field \( B_\perp \). While this field would indeed choke off the beam over a short timescale (Alfvén 1939; Lawson 1957) in the absence of a strong guide field, Brown and Bingham (1984) pointed out that the presence of such a guide field, as exists in flare loops, results only in a increased helicity in the guiding field lines; as long as the azimuthal \( B \) field remains small or comparable to the guiding longitudinal field \( B_\parallel \), then little modification to the beam energetics results. The consensus reached is that, on the one hand, for beams spread out over large areas, edge effects are of minor importance, and so the predominant mechanism controlling the reverse current is the electrostatic field established by the displacement current at the head of the beam. On the other hand, for beams of small area, e.g., if the injection occurred over many sub-resolution elements (a “pepper-pot” scenario), the inductive edge effects would predominate. In such a case, “choke-off” of the beam (Spicer and Sudan 1984) may occur on timescales relevant to observations, but for this effect to have an appreciable hard X-ray signature, many elements of the “pepper-pot” would have to act coherently, and it is far from clear why this should occur. However, in both these cases it appears that the earlier steady-state electrostatic treatments of the reverse current energetics give valid results.

### 3.3.1.3 Collective Instabilities

Emslie and Smith (1984) showed how, due to the inverse square energy dependence of the Coulomb collision cross-section, the velocity distribution for a nonthermal electron beam injected into a cold plasma is preferentially depleted at low energies, resulting in the formation of a “hump” in the combined (background plus beam) distribution function. This type of distribution is well-known to be unstable to the growth of Langmuir waves; Emslie and Smith calculated the resulting wave level and discussed nonlinear interactions of these waves, through the process \( \ell + \rightarrow t \), (where \( \ell \) is a Langmuir wave and \( t \) a transverse wave with angular frequency \( \approx 2\omega_{\text{pe}} \), where \( \omega_{\text{pe}} \) is the electron plasma frequency), to produce microwave radiation. They showed that the level of microwave radiation produced by this process is, for typical thick target beam parameters, much larger than observational upper limits on microwave fluxes from flares. This implies that either the nonthermal microwave radiation produced is strongly absorbed by gyroresonance absorption in the surrounding plasma, with corresponding strong implications for the energetics of the flare as a whole (Melrose and Dulk 1982b), or that the thick target electron heated model in its present form is incorrect.

Mok (1985) has examined another instability connected with the scattering of electron beams in the solar atmosphere — the anomalous Doppler resonance instability (Kadomstev and Pogutse 1968). This instability is driven by anisotropy in the electron velocity distribution function and operates through “beat” resonances of the form \( k \cdot v = \omega + n \Omega \), where \( k \) is the wavenumber, \( v \) the electron velocity, \( \Omega \) the electron plasma frequency, \( \Omega \) the gyrofrequency and \( n \) an integer. Using a power-law pitch angle distribution centered in the direction of the guiding magnetic field lines, with a moderate degree of anisotropy (\( \theta = 0.25 \)), Mok found that for low \( \omega_{\text{pe}}/\Omega \) (e.g., in the corona), no resonant amplification of waves through this instability results. However, with the increase in \( \omega_{\text{pe}} \) caused by increasing density in the chromospheric regions of the loop, the ratio \( \omega_{\text{pe}}/\Omega \) becomes large enough so that the instability threshold is exceeded. This results in a growth of lower hybrid waves (Krall and Trivelpiece 1973), leading to fast collisionless scattering of the beam electrons into a more isotropic distribution. This has several effects. First, it causes the beam to suddenly slow down. Second, there is strong local heating produced near the region where the onset of instability occurs, possibly affecting the ratio of hard X-ray and EUV fluxes (cf. Section 3.2.5). Third, some microwaves are emitted in this region, although they may not represent a dominant contribution to the microwave emission during the impulsive phase.

### 3.3.1.4 Electron Beam Momentum

The strong observational evidence for the applicability of the thick target electron beam model to the transport of
impulsively accelerated electrons has led many authors to carry out studies of its hydrodynamic and energetic consequences on the structure and dynamics of the target atmosphere (see, e.g., Sections 3.2.5 and 3.2.7). However, these studies have all focused their attention on the heating aspects of the electron beam; they have neglected the effect of direct momentum deposition in the equation of motion. As pointed out by Brown and Craig (1984), beam momentum is potentially of great interest to impulsive phase transport because it is transferred directly to the atmosphere, and hence produces an immediate acceleration \( \mathbf{a}_B \), whereas acceleration by nonhydrostatic thermal pressure gradients, \( \mathbf{a}_p \), takes time to build up.

Brown and Craig found that an electron beam, for high values of the energy flux, produces a beam acceleration that is much greater than solar gravity (\( \mathbf{a}_B \approx 100 \, g_o \), for beam energy flux \( 10^{14} \, \text{ergs cm}^{-2} \, \text{s}^{-1} \)). They compared \( \mathbf{a}_B \) to \( \mathbf{a}_p \) on the assumption that beam heating is balanced by conduction. Their comparison implied that \( \mathbf{a}_p \) was less than \( \mathbf{a}_B \) for classical conduction and of the same order as \( \mathbf{a}_B \) for saturated heat flux. Hence, they concluded that beam momentum deposition is important even after nonhydrostatic thermal pressure gradients have developed.

McClymont and Canfield (1984) came to quite a different conclusion. They estimated the time scale for \( \mathbf{a}_p \) to exceed \( \mathbf{a}_B \), arguing that conduction and radiation can be neglected on the short time scales of interest, and so considered beam heating alone. They found that in the corona, \( \mathbf{a}_B \) dominates \( \mathbf{a}_p \) for at most the first few seconds; in the chromosphere it will do so even more briefly (\( \lesssim 10^{-2} \, \text{s} \)).

The basic reason for this is the large density gradient across the transition region in the preflare atmosphere. Since coronal heating gives rise to a temperature increase rate which is (locally) independent of density (Emslie 1978), the rapid heating of the preflare transition region layers to coronal temperatures gives rise to a large pressure gradient across this region. Hence they concluded that electron beam momentum deposition is not important to the global evolution of the flare atmosphere.

Both Brown and Craig (1984) and McClymont and Canfield (1984) also considered the hydrostatic case, i.e. an electron beam that has been on for long enough to reach a new pressure equilibrium. Because \( \mathbf{a}_B \) can so greatly exceed gravity, Brown and Craig concluded that steady-state hydrostatic models require revision. However, McClymont and Canfield called attention to the fact that flare electrons are accelerated in closed loops. Due to containment of the heated plasma, the thermal coronal pressure is much greater than that due to gravity alone. As a result, the addition of beam pressure to gravitational and containment pressure has a negligible effect.

At this time no study has yet been carried out that examines the hydrodynamic importance of beam momentum deposition, taking fully consistent account of beam heating, conduction, radiation, and mass motion, in order to unambiguously demonstrate which of these conflicting views is correct. For hydrostatic models, on the other hand, it seems clear that the electron beam pressure can safely be neglected. However, under rather specialized conditions the momentum imparted to the atmosphere by proton beams may be significant (Brown and Craig 1984, Tamres, Canfield and McClymont 1985).

### 3.3.2 Proton Transport

The timing, spectrum and morphology of flare protons, when compared to electrons, provide important clues on the nature of flare acceleration and transport. Here we concentrate our attention on specific recent advances in understanding proton transport processes between the acceleration site and the lower solar atmosphere; the reader interested in interplanetary proton phenomena is referred to Forman, Ramaty and Zweibel (1985).

#### 3.3.2.1 Lyman-\( \alpha \) Charge-Exchange Emission

The broad-band X-ray and microwave spectrum of the impulsive phase of solar flares has been observed well enough so that we can confidently state that the most commonly accelerated nonthermal electrons are those in the deka-keV range. Are the most commonly accelerated protons also in this energy range? Or does the fact that gamma rays are the only observed emission due to protons mean that flares accelerate protons primarily to MeV energies? Suppose that flares typically accelerate protons, in numbers comparable to electrons, in the deka-keV range. What are the observable manifestations?

It is known that if energetic protons are injected into the largely un-ionized chromosphere, they will efficiently pick up electrons from ambient neutral hydrogen atoms (Orrall and Zirker 1976). The nonthermal protons thereby become nonthermal hydrogen atoms, and radiate the usual hydrogen emission spectrum, redshifted by the Doppler effect. Orrall and Zirker carried out an equilibrium calculation of charge exchange in an empirical model chromosphere that showed that the nonthermal \( L\alpha \) radiation would be readily observable above the quiet sun Lyman-\( \alpha \) spectrum, for an input energy flux of \( 2 \times 10^7 \, \text{ergs cm}^{-2} \, \text{s}^{-1} \), assuming a power-law proton energy distribution with a 10 keV low-energy cutoff and exponent \( \delta = 2.5, 3, \) and 4. Recognizing that the ionization structure of the atmosphere would not long remain undisturbed by the energetic protons, they showed that for the beam parameters considered, the chromospheric ionization time was in the range \( 10^3 \sim 10^5 \, \text{s} \).

Recently Canfield and Chang (1984) examined the sensitivity of the Lyman-\( \alpha \) photon emission spectrum to proton energy. Since the largest flares are thought to inject about \( 10^{11} \, \text{ergs cm}^{-2} \, \text{s}^{-1} \) into the lower atmosphere, they investigated the Lyman-\( \alpha \) radiation that would be generated by this same flux, in the form of monoenergetic proton beams. Since the temperature and density structure of the atmosphere
is of secondary relevance compared to its ionization structure, they assumed a uniform atmosphere. They adopted a 10% hydrogen ionized fraction, and assumed an equilibrium between the fast protons and the ambient chromosphere. Their result, shown in Figure 3.28, demonstrates a high level of emission over a wide range of proton energies.

It remains to be demonstrated that the radiation generated by charge exchange is produced when the atmosphere is heated (and perhaps largely ionized) by the proton beam. In particular, for what range of proton beam parameters will enough of the atmosphere remain sufficiently neutral to make charge exchange emission comparable to that calculated? Canfield and Chang (1984) showed that at high input beam fluxes the ionization time is much shorter than the charge exchange time. At high values of incident proton energy the ionization time tends to decrease more rapidly than the charge exchange time. At low values a proton accelerated at the top of a typical coronal loop does not reach the chromosphere. At low values of the incident energy flux $F(0)$ the emission predicted by the steady-state calculation is too weak to be detectable above a typical preflare (active region) background. Finally, at high values the ionization time becomes short, making the emission difficult to detect. Notwithstanding, Canfield and Chang showed that some cases of potentially great interest (e.g., a beam of 300 keV protons and energy flux $10^6$ ergs cm$^{-2}$ s$^{-1}$) are well approximated by an equilibrium calculation for tens of seconds, and easily detected above the active-region Lyman-$\alpha$ background. Future calculations should examine time-dependence, a realistic target atmosphere, and spectra including Bessel functions (Forman, Ramaty, and Zweibel 1984).

### 3.3.2.2 Heating of a Thick-Target Atmosphere

With the simultaneous operation of the Hard X-ray Burst Spectrometer (HXRBS) and Gamma Ray Experiment (GRE) on SMM, a few flares exist for which there are good data in both hard X-rays and $\gamma$-rays. These data provide information of the populations of energetic deka-keV electrons and Mev protons, respectively. Emslie (1983a) has used the available data from both experiments for the 1980 June 7 flare (see Kiplinger et al., 1983) to infer the energies in both electrons and protons for this event. Through thick target modeling of these emissions (Brown 1971, Ramaty 1984),

---

**NONTHERMAL L$_\alpha$ SPECTRA FOR MONOENERGETIC PROTONS**

![Graph showing nonthermal Lyman-\(\alpha\) spectra for monoenergetic vertical proton beams of energy flux \(10^{11}\) ergs cm$^{-2}$ s$^{-1}$. The long-dashed curve, marked F, is the average of two flares at Lyman-\(\alpha\) flare maximum (Canfield and VanHoosier, 1980). The short-dashed curve is an active-region spectrum (Cohen 1982).](image)

**Figure 3.28** Solid curves: theoretical nonthermal Lyman-\(\alpha\) spectra for monoenergetic vertical proton beams of energy flux \(10^{11}\) ergs cm$^{-2}$ s$^{-1}$. The short-dashed curve, marked F, is the average of two flares at Lyman-\(\alpha\) flare maximum (Canfield and VanHoosier, 1980). The long-dashed curve is an active-region spectrum (Cohen 1982).

---

3-39

---

*Provided by the NASA Astrophysics Data System*
Emslie has calculated the energy deposited per unit height by both electron and proton bombardment, as a function of depth in the atmosphere. Apart from the (relatively) unknown area factor, determination of the instantaneous energy input rate is possible only for electron heating, since hard X-rays are “prompt” data, providing information on the instantaneous target-averaged electron flux. On the other hand, some γ-rays (e.g., the deuterium formation 2.223 Mev line) are produced over a considerable period (compared to observational time scales) after the injection of the protons which lead ultimately to the γ-ray emission. We can calculate only the event-integrated energy deposition with reliability.

Proton spectra are in general harder (shallower) than electron spectra. Since Mev protons have a greater individual energy and stopping depth (Emslie 1978) than deka-keV electrons, it follows that in general proton heating dominates electron heating at large depths, while high in the atmosphere electron heating dominates. The “cross-over” point depends on the actual shape of the electron and proton spectra, and on the relative energy fluxes in the two sets of particles. Figure 3.29 shows, for the 1980 June 7 flare, the energy deposited per unit height interval by both electron and proton bombardment. It can be seen that for this event, electron heating dominates all the way down to a particle column density $N \approx 3 \times 10^{22}$ cm$^{-2}$. Since this level is below any reasonable location of the chromospheric flare (e.g., Machado et al., 1980; Ricchiazzi and Canfield 1983), it follows that electron heating dominates throughout most of the region that gives rise to observed flare emissions (with the exception of emissions from around the temperature minimum [see Machado, Emslie, and Brown 1978]). Since the 1980 June 7 event had a particularly high γ-ray to hard X-ray flux ratio (Kiplinger 1983), it follows that the latter conclusion is valid for all flare events (see also Lin and Hudson 1976). However, Emslie (1983a) pointed out that this conclusion is critically dependent on the assumption of thick target emission for both the hard X-rays and the γ-rays. While Ramaty (1984) has conclusively demonstrated that γ-ray emission in solar flares is a thick target process, there is still considerable speculation (see, e.g., Brown and Smith 1980; Emslie 1983b) on whether hard X-rays are produced principally by thick target bremsstrahlung or by thermal bremsstrahlung from a confined ensemble of very hot ($\lesssim 10^8$ K) electrons. In the latter case, the number of precipitating electrons is reduced relative to the hard X-ray photon flux, and less chromospheric heating results (see Emslie and Vlahos 1980). Thus, if hard X-rays indeed turn out to be emitted principally by a hot thermal source, then the importance of proton heating in the chromospheric energy budget would merit closer examination.

![Figure 3.29](image_url) Figure 3.29 The quantity $I_B/n = \int \{ Q(n,\gamma,\nu)/n \} \mathrm{dx} \, \mathrm{dy} \, \mathrm{dt}$ (units ergs per cm of vertical height per ambient electron), representing the temporal and areal integral of the specific energy deposition rate $Q$ as a function of column depth $N = \int n(z) \, dz$, for the 1980 June 7 event. From Emslie (1983a).
3.3.3 Radiative Energy Transport by Amplified Decimetric Waves

Following the initial work on the application of maser theory to radio bursts from the Sun and stars (Holman et al., 1980; Melrose and Dulk 1982a) and a discussion of the energetic implications (Melrose and Dulk 1982b), Melrose and Dulk (1984) have carried out new work on the transport of energy by amplified radio radiation. These transport aspects are summarized below; the role of masers in particle acceleration is discussed in the previous chapter.

Strong theoretical arguments suggest that up to 50% of the energy in a flare should go into radiation at the cyclotron frequency \( \Omega_c \), the exact percentage depending on the degree of field line convergence toward the footpoints of the flux tubes, and the amount of fast electron energy lost through collisions. These arguments (based on an extension of then-existing theory by Wu and Lee in 1979) are reinforced by consideration of how the electrons accelerated in a flare lose the component of their energy perpendicular to the magnetic field in order to precipitate into the footpoints of magnetic flux tubes and hence heat the chromosphere and generate hard X-ray bursts. The perpendicular energy is radiated away through amplification of decimetric radio waves. The radiation travels from the flux tubes of energy release, across field lines, and is then reabsorbed at the second harmonic of \( \Omega_c \), i.e. at points where the field strength \( B \) is half its value in the source. From energetic considerations there seems to be enough energy in the radio-frequency (RF) radiation to heat the plasma of the absorbing region so that it produces the soft X-rays. In addition it is likely that this RF heating process can account for other properties of soft X-rays in the impulsive phases of flares: large source sizes, excess line widths, characteristic temperatures, and their temporal development.

3.3.3.1 Mechanism for Emission and Amplification

The generation of the high levels of RF radiation is thought to occur in an impulsive flare as follows.

1. It is assumed that energy release occurs mainly near the tops of one or more magnetic flux tubes of the kind illustrated in Figure 3.30. The accelerated (or heated) electrons are assumed to have energies \( E \sim 10 \) to 100 keV (or \( T \sim 10^8 \) to \( 10^9 \) K) and to have a nearly isotropic pitch angle distribution, i.e., \( f(E) = f(E_1) \).

2. The fast electrons, with \( v \geq 0.1 \ c \), travel down the legs of the flux tubes. Those with moderate to large pitch angles \( \alpha \) reflecting in the converging fields, and those with small \( \alpha \) precipitating. In a typical travel time of \( \sim 1 \) s the electron pitch angle distribution \( f(\alpha) \) in the legs is anisotropic, with no upgoing electrons with small \( \alpha \). The anisotropy represents a source of free energy because it creates a non-equilibrium distribution with an inverted population, i.e., there is a region of velocity space where \( \delta f/\delta v > 0 \).

3. Provided that the plasma frequency \( \omega_p \) is \( \leq \Omega_c \), there can be a resonant transfer of energy from electrons to waves: in particular, electromagnetic waves at the cyclotron frequency can be amplified. The growth rate of this "cyclotron maser" is very large, \( \sim 10^7 \) s\(^{-1} \), so in a few microseconds the waves can be amplified by a factor of \( \sim e^{10} \) or so, at which time they saturate, having extracted most of the free energy. With \( B \) in the expected range of 100 to 1000 gauss, the amplified radiation is in the range 0.3 to 3 GHz, i.e., at decimetric wavelengths.

4. Saturation of the maser is expected to occur by quasilinear relaxation: the maser converts the perpendicular energy of certain electrons into radiation and leaves them with mainly parallel energy, i.e. it diffuses them into the loss cone at the maximum possible rate. Provided that the density is low and/or the electron energy is large, so that collisions are infrequent, this is the only known process which can prevent most of the electrons from remaining trapped in converging, coronal flux tubes for many bounce periods, i.e. for several to many seconds. For example, 30 keV electrons trapped between mirror points \( 3 \times 10^8 \) cm apart would bounce \( \sim 10^2 \) times (\( \sim 10^2 \) s) in a plasma of \( n_e = 10^6 \) cm\(^{-3} \) before collisions would diffuse them into the loss cone (Section 3.2.2).

3.3.3.2 Energy Content

From studies of hard X-ray bursts it is known that the energy going into fast electrons in the impulsive phase is \( E \sim 10^{37} \) to \( 10^{38} \) ergs s\(^{-1} \) (e.g., Brown and Smith 1980). Provided that the electrons are not highly collimated, about half of this is in the parallel component and, again assuming

Figure 3.30 Schematic drawing of a sequence of loops showing an energy release region, the maser source region in one leg where the field strength is \( B \), the cone of maser radiation with reflection of radiation directed toward higher field strengths, and the absorption regions where the field strength is \( B/2 \). From Melrose and Dulk (1984).
a low collision frequency in the coronal portion of the flux tubes, it is deposited in the low atmosphere at the footpoints of the loops of energy release. Most of the other half goes into the maser radiation, and then (as described below) into reabsorption and heating of the coronal plasma in a large volume, \( V \sim 10^{24} \) to \( 10^{28} \) cm\(^3\), surrounding the loops of energy release. Thus the average heating rate is \( E/V \sim 1 \) to \( 10 \) ergs cm\(^{-3}\) s\(^{-1}\), assuming that the larger values of \( E \) and \( V \) go together. On average this can increase the internal energy \( n k T \) of the preflare corona by a factor of 2 to 20 each second.

### 3.3.3.3 Reabsorption at Distant Locations

The RF radiation travels outward from the maser generation regions in the manner sketched in Figure 3.30. Only one maser emission region is shown, but there are expected to be \( \geq 10^7 \) maser pulses per sec, each lasting about 1 ms, coming from various points in the legs of the flux tube, and from all flux tubes of energy release. Each pulse is narrowband and at a frequency \( \omega \) slightly higher than \( \Omega_e \), but in total the pulses should cover a frequency range of an octave or so because \( B \), hence \( \Omega_e \) and \( \omega \), varies with position in the legs of the loops.

The maser pulses travel across field lines at nearly the speed of light until they encounter a region where the field strength is half its value at the source: this may be \( 10^3 \) to \( 10^4 \) km distant, depending on the gradient of \( B \). Then gyroresonance absorption by the ambient electrons occurs at the second harmonic, very efficiently, with an optical depth \( \tau \sim 10^4 \). A series of pulses can heat a small volume of plasma, \( \leq 10^{20} \) cm\(^3\), to a few \( \times 10^7 \) K in \( \sim 0.1 \) s. Temperatures above about \( 3 \times 10^7 \) K energy are difficult to achieve by maser heating because of the temperature dependence of the collision frequency; since the mean free path is \( \propto T_e^{-1.5} \), further energy deposition leads to heating of a larger volume of plasma rather producing a higher temperature.

Turbulent plasma motions should occur because the maser pulses heat the ambient electrons in many, small, localized volumes, a result of the fine beaming and small absorption distance. The hot electrons tend to propagate out of the pockets of high pressure so created and thus provide an explanation for the highly broadened soft X-ray lines observed during impulsive flares.

The overall volume heated depends mainly on the field gradient in the vicinity of the loops of energy release. The model suggests that in a small flare a volume \( \sim 10^{26} \) cm\(^3\), some 10 times larger than the volume of energy release, would be heated to \( \geq 10^7 \) K in 10 seconds. In a large flare it is expected that the heated volume would be much greater than \( 10^{26} \) cm\(^3\), as would the volume in which energy release occurs. Because the maser radiation is initially directed at a large angle (\( \approx 70^\circ \)) to the field lines, most of the energy transport is across the field.

### 3.3.3.4 Secondary Effects

According to the RF heating model, the time profile of maser emission in simple impulsive flares should mimic the hard X-ray profile, and this should follow closely (within \( \sim 1 \) s) the profile of energy release. Much of the RF energy is deposited in the soft X-ray plasma and is radiated or conducted away, relatively slowly. Hence the slope of the soft X-ray intensity profile is expected to be approximately proportional to the hard X-ray intensity as observed. While this property is not unique to the RF heating model, it applies here to a soft X-ray emitting volume much larger than that of initial deka-keV energy release.

Densification of heated loops and the origin of the blue-shifted component of soft X-ray lines can occur in two ways. First, energy in the parallel component can "boil off" chromospheric material into the loops of energy release. Second, as other loops become heated by the maser radiation, enhanced conduction to the footpoints would boil off material there, though presumably not as fast.

The RF heating model has implications for enhancement of chromospheric evaporation, a major topic of the following chapter. As pointed out there, chromospheric evaporation can take place by both direct Coulomb heating from the parallel component of the nonthermal electron energy and by thermal conduction from energy deposited in the corona (possibly a slower process). Hence, if the RF heating mechanism heats loops other than those in which direct energy release has taken place, chromospheric evaporation may take place over a larger area of the chromosphere.

If the density in a flaring loop increases to the point where \( \omega_p \geq \Omega_e \), e.g. \( n_e \geq 10^{10} \) cm\(^{-3}\) if \( B = 300 \) gauss, the maser would stop. Or if the initial density in the flaring loop were such that \( \omega_p \geq \Omega_e \), there would be no maser, and probably only a gradual phase. However, in this high-density case, Coulomb collisions would scatter and degrade the energy of all but the highest energy electrons. Thus the conditions would be appropriate for a "high coronal flare" without an impulsive hard X-ray burst from footpoints.

### 3.3.3.5 Conclusions

As the RF heating radiation is entirely reabsorbed in the corona, it is not possible to observe it directly. However, there is strong evidence that this maser mechanism is the explanation for the very intense auroral kilometric radiation of Earth and for similar radiation from Jupiter and Saturn. From the Sun and stars, direct maser action (or a related process involving conversion of amplified waves) producing radiation just above the second harmonic is the favored explanation for "microwave spike bursts" of very high brightness and circular polarization; Enome (1983) has recently examined some of the properties of such bursts, and earlier studies have been made by Droge (1977), Slottje (1978) and Zhao (1982).
3.4 SUMMARY

In this chapter our main interest has been the transport of energy, momentum and charge during the dramatically nonthermal and energetic impulsive phase. Not surprisingly, the technological and operational advantages embodied in the SMM and Hinotori instruments, and their collective and coordinated use with forefront ground-based optical and radio instruments, have brought about many advances in our understanding of impulsive phase flare physics.

The topic most thoroughly explored in this chapter is that of the transport of nonthermal electrons. The thick-target electron beam model, in which electrons are presumed to be accelerated in the corona and typically thermalized primarily in the chromosphere and photosphere, is supported by observations throughout the electromagnetic spectrum. At the highest energies, the anisotropy of γ-ray emission above 10 MeV clearly indicates that these photons are emitted by anisotropically-directed particles. The timing of this high-energy γ-radiation with respect to lower-energy hard X-radiation implies that the energetic particles have short lifetimes. For collisional energy loss, this means that they are stopped in the chromosphere or below. Stereoscopic (two-spacecraft) observations at hard X-ray energies (up to 350 keV) imply that these lower-energy (but certainly nonthermal) electrons are also stopped deep in the chromosphere. Hard X-ray images show that, in spatially resolved flares whose radiation consists of impulsive bursts, the impulsive phase starts with X-radiation that comes mostly from the footpoints of coronal loops whose coronal component is outlined by microwaves. Recent X-ray polarization measurements at the low end of the deka-keV range show small values, but these are still large enough to be compatible with highly collimated particle injection somewhere in the coronal part of a loop. Preliminary analyses of combined hard and soft X-ray spectra appear to demand the presence of nonthermal electrons early in the impulsive phase. The thick-target electron-beam model accounts for both the close temporal coincidence of UV and hard X-ray bursts and the relative decrease of UV emission with increasing hard X-ray intensity. White-light emission from the largest flares is closely related temporally and energetically to the nonthermal electrons that produce deka-keV X-rays. Moreover, white-light spectra imply that commonly such emission originates in the chromosphere, where deka-keV electrons are stopped. Both timing and spectra of Hα emission imply that energetic nonthermal electrons penetratively heat the chromosphere during the impulsive phase. Finally, we now suspect that microwave spike bursts may require a level of microwave maser emission that dramatically enhances the precipitation of thick-target electrons.

Although the thick-target nonthermal-electron model meets with some success, it describes only a fraction of observed flares, and shows some puzzling and challenging failures. The limited amount of stereoscopic hard X-ray data at 350 keV shows little directivity, compared to the predictions of models of collisional electron transport in loops. If this trend is borne out by more comprehensive observations, it could provide a strong limitation on the thick-target model. The height dependence of the hard X-ray spectrum is systematically different from that predicted. The majority of images at the low end of the deka-keV energy range show single X-ray sources; in many flares, the X-ray producing electrons in this energy range are obviously confined to the corona. It now seems clear that direct collisions by nonthermal electrons are not necessary to understand the production of inner-shell Fe Kα emission; it appears that impulsive variability of these lines comes about because of irradiation of the chromosphere and photosphere by an impulsively varying spectrum that extends down to at least 7 keV. The thick-target model appears to predict much more UV emission, in proportion to hard X-ray emission, than is observed. In some flares, white-light emission is observed well beyond the end of the impulsive phase, which appears utterly inexplicable in the thick-target model; such emission appears to correlate more strongly with the thermal X-ray emission of the hot (T > 2 x 10^7 K) thermal plasma. Furthermore, some flares show spectra that imply that the white-light emission comes from deeper than the chromosphere, in regions that are collisionally inaccessible to deka-keV electrons. Finally, it is not well understood whether electron beams of the required flux can be maintained, in the presence of return-current instabilities known to exist in laboratory plasmas, and in view of the observational uncertainty in beam area.

Much of the work that we have described above leaves basic questions unanswered, requiring future observational and theoretical work. First, high spatial resolution, from microwaves through γ-rays, is the observational key to understanding impulsive-phase transport. What is the spatial scale of particle beams? Return currents? To what extent are they filamentary? What processes dominate the physics of return currents? What physical processes could disrupt beams? What role do they play in heating the ambient plasma? Does the relative size of hard and soft X-ray sources suggest an important energy transport role for microwave maser radiation? Second, what is the mass and charge composition of flare-accelerated particles? Do proton beams exist in the deka-keV range? Are energetic protons important to impulsive-phase momentum transport? Do an energetically significant number go undetected because of their inefficiency of bremsstrahlung production? Third, where are nonthermal particles accelerated within flare loops? What is their degree of anisotropy and their energy spectrum? How are they confined to the corona in those flares that do not show footprint brightening in hard X-rays? How are they thermalized as they travel along flare loops? Is their transport dominated by particle collisions or by wave generation or by the emission or absorption of radiation? Are regions of particle precipitation identical to regions of explosive coronal and chromospheric hydrodynamic motion? Will the commonly
accepted model of coronal loops as the basic building block of flare geometry withstand close scrutiny? Not until the answers to these questions are better understood can we claim to understand the basic physics of impulsive-phase transport.

3.5 REFERENCES


Hayward, J. 1984, unpublished.

