Solar flares

JOHN C BROWN† and DEAN F SMITH‡

High Altitude Observatory, National Center for Atmospheric Research§, PO Box 3000, Boulder, Colorado 80307, USA

Abstract

The current observational and theoretical status of solar flares as a typical astrophysical problem is reviewed. To delimit the super-abundant literature in the field, the overriding philosophy is applied that the essential flare problem is that of the intense and complex energy release in large flares. Consequently only those observations and models are discussed which, in the view of the authors, have made or are likely to make a significant contribution to understanding this problem.

Observations and their diagnostic applications are discussed in three broad areas: thermal radiation at temperatures $T \lesssim 10^5$ K; thermal radiation at $T \gtrsim 10^5$ K; and non-thermal radiation and particles. (The first two categories are divided naturally by the thermal stability properties of the solar atmosphere in general while the third is characteristic of the transient nature of flares.) Particular emphasis is given to the most recent observational discoveries such as flare γ-rays, interplanetary Langmuir waves, and the ubiquitous association of soft x-ray loops with flares, and also to progress in the important particle diagnostics of hard x-ray and radio bursts.

Theoretical progress in the problem of primary energy release is discussed first in terms of possible magnetic configurations in which energy can be made available by field annihilation or reconnection and second in terms of the plasma instabilities likely to be involved in this release process. Then the central question of achieving the necessary flash phase power is discussed and it is concluded that the most viable models invoke the tearing mode in a sheared magnetic geometry such as an arch. On the other hand, a classical neutral sheet geometry, especially in the Petschek mode, cannot be entirely ruled out in all cases.

Secondary redistribution through the atmosphere of the primary magnetic energy released has recently been modelled in considerable detail in terms of conduction, convection, radiation and particle transport. Models exist for both the flash and decay phases in the cool and hot regimes of the solar atmosphere and also in interplanetary space. The credibility of these models is discussed in terms of their realism, uniqueness, self-consistency and consistency with observations. It is concluded that this has

been an area of substantial progress in the flare problem but that the atmospheric response is linked to the primary release mechanism rather non-uniquely.

Acceleration of particles in flares is the hardest problem after the primary release process itself. Therefore a detailed critique is made of the interpretation of hard x-ray bursts since these have been thought to place the most stringent demands on acceleration efficiency. The present situation seems to be that the required efficiency is uncertain by several orders of magnitude. It is argued that the most likely mechanism for accelerating non-relativistic electrons, which comprise the bulk of particles in any case, is stochastic acceleration by resonant interaction with Langmuir waves, though the means of initiating the necessary waves is not yet clearly known. In any event a considerable fraction of the energy available always goes into heat rather than acceleration. Given this first phase process as an injection mechanism, the acceleration of relativistic electrons and solar cosmic rays to very high energies by second-stage Fermi acceleration presents no difficulty in principle.

Finally an attempt is made to assess the progress which has been made in the flare problem in the past decade, and to give some possible reasons why no convincing solution has yet been found. With these factors in mind, a prognosis is made of the areas in which future progress may be expected in view of the present direction of flare theory and observation.

This review was received in its present form in June 1979.

Contents

1. Introduction

'Solar flares are complex transient excitations of the solar atmosphere above magnetically active regions of the surface involving enhanced thermal and radio emission, hard x-rays, cosmic rays and plasma ejection. Their origin is not yet understood after more than a century of study since the first recorded observations' (Sweet 1969).

Ten years later we can say little to better this statement either in its concise definition of the solar flare phenomenon or in its expression of the continuing enigma of flare mechanisms. We would only add that, aside from their intrinsic interest, flares can be regarded as occupying a central position in astrophysics. Firstly the study of flare mechanisms involves almost the entire gamut of plasma physical processes which are now recognised as central to many astrophysical situations. Secondly, as an observational diagnostic problem the flare is almost archetypal for, on the one hand, it presents the diverse range of manifestations typical of astrophysics while, on the other hand, the proximity of the Sun permits sufficiently detailed observations to counter the over-simplistic modelling which is not uncommon in astrophysics.

Flares have been observed with more or less increasing frequency ever since their discovery, with optical (specifically Hα) data heavily dominating flare interpretations and classifications until recently (e.g. Smith and Smith 1963). However the advent of space-borne instrumentation in xuv and shorter wavelengths, together with technical developments in radio and other ground-based techniques, has resulted in a veritable explosion of literature on flare observations over the past twenty years. Simultaneously plasma physics developments in the laboratory and in theoretical studies, aided by the advance of numerical simulation techniques, have resulted likewise in a profusion of publications on flare plasma physics. The formidable task of reviewing all this material has been somewhat reduced, in the observational and diagnostic areas at least, by the publication of Švestka's (1976) excellent comprehensive monograph on the subject and, for some of the most recent material, by the reports of the various teams involved in NASA's Skylab Flare Workshop (hereafter referred to as Sturrock *et al* 1978). Nevertheless, in order to make the task tractable and the review readable, it has been necessary to restrict our scope considerably. As far as theory is concerned, we have taken the view that *the* central flare problem is that of explaining the diverse energetic phenomena associated with large flares, and particularly with the (most powerful) flash phase, since these provide the most stringent demands on theories of the flare energy release process. (An alternative view—e.g. Švestka (1976)—is that small 'flares' should be studied first since they are abundant and simple, often not showing (at presently detectable levels) many of the high-energy effects characterising large flares.)

As far as observations are concerned we have therefore been highly selective, presenting in detail mostly only that material which we see as being directly relevant to this central theoretical problem of the energy release process. Clearly there is a strongly subjective element in this approach. For a more comprehensive survey the reader is referred to Švestka's (1976) book, to the literature as reviewed by Sweet (1969) for example, and to the proceedings of the many recent symposia on flares and flare-related solar physics problems. These last include volumes edited by Howard

(1971), Ramaty and Stone (1973), Athay (1974), Newkirk (1974), Nakagawa and Rust (1974), Kane (1975), Massey *et al* (1976), Švestka (1975, 1977), Kuperus (1977) and Sturrock *et al* (1978). There have also been special journal review issues including de Jager (1974) and Kennel *et al* (1978) and a most useful review of reconnection theory by Vasyliunas (1975) and of the plasma physics of the solar atmosphere by Kaplan *et al* (1974).

The real starting point of the modern regime of flare interpretation may be regarded as the introduction of the first magnetic theories of flare energy release (Giovanelli 1946, Dungey 1953) which seem to be the only ones energetically viable (cf §3.1). Subsequent key developments were the analyses of quasi-steady magnetic-field annihilation in a plane current sheet by Sweet (1958) and Parker (1963) and their elaboration by Carmichael (1964) and Sturrock (1968), Petschek's (1964) analysis of quasi-steady field reconnection in standing slow MHD mode shock waves, and Syrovat-skii's (1966) dynamic dissipation model. The probably fundamental role of the tearing mode instability (Furth *et al* 1963, cf §3) was anticipated by Jaggi (1964). Another important line of thought concerned the role of twisted magnetic flux tubes (loops) both in pre-flare energy storage (Gold and Hoyle 1960, cf Anzer 1968) and in flare energy release by onset of instability in the current generating the twist of the loops (cf Alfven and Carlquist 1967). One of the most important recent observations is the Skylab XUV indication of the ubiquitous presence of hot loops in flares which has regenerated much interest in and development of such mechanisms involving loop geometry (cf Spicer 1976, 1977) with some resulting controversy with the protagonists of the classical neutral sheet geometry (see, for example, Sturrock *et al* (1978) for discussion). Other key observations resulting from the steady improvement in space-borne instrumentation for detection of high-energy photons and particles have been the first observation of γ-ray lines (Chupp *et al* 1973) and of Langmuir waves in direct association with interplanetary electron streams and type III bursts (Lin 1974a,b, Gurnett and Anderson 1976).

Our plan here then is firstly to give a rather subjective survey of flare observations with particular emphasis on those recent data such as noted above which seem to give a direct clue to physical processes. While recognising the limitations of this personal approach, we hope that it may provide for the newcomer to the topic a short-circuit past the super-abundance of phenomenological flare data, with no well-established physical significance, which must inevitably repel those invaluable new researchers who are looking for a piece of clean physics to study. Subsequently we survey the present theoretical status of the flare problem under its three main headings of: primary energy release; particle acceleration; relaxation processes. Finally we assess the extent to which, and some reasons why, the flare phenomenon remains 'not yet understood' after a decade of intensive space-borne observation since Sweet's (1969) review.

2. Observational manifestations of flares

2.1. Thermal flare emissions

2.1.1. Introduction. A convenient, though much simplified, view of the *quiet* solar atmosphere is that it consists of a horizontally stratified structure sharply divided into low- and high-temperature components (the chromosphere and corona) by a thin layer (the transition region) (see figure 1). This form may be understood in terms of the
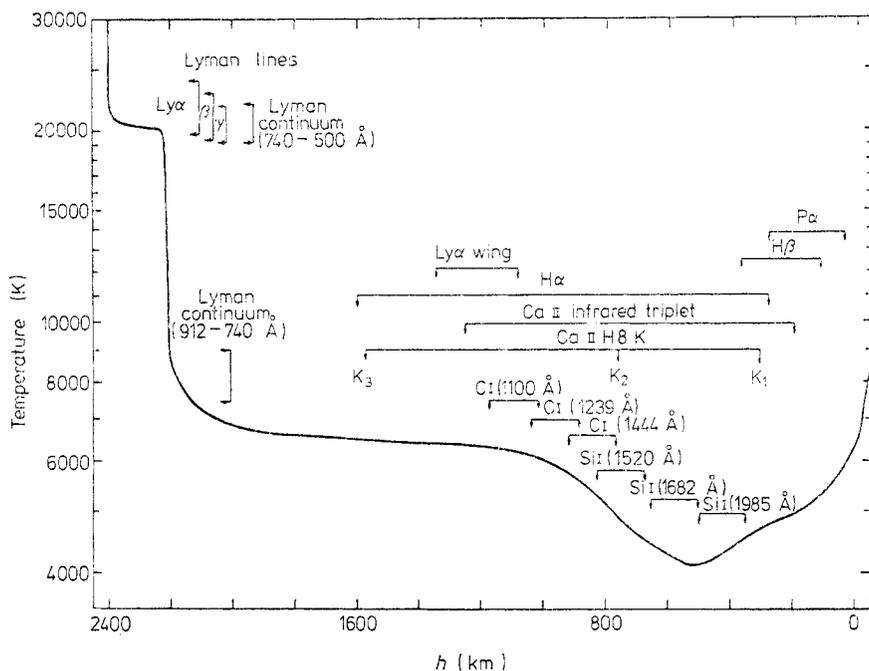
Figure 1. The temperature distribution with height in the quiet solar chromosphere according to the plane-parallel model derived from spectral data by Vernazza *et al* (1973).

energy balance of the atmosphere and, in particular, by the occurrence of a maximum in the radiative loss function of an optically thin plasma with cosmic or solar abundances, the maximum occurring around temperature T of 6×10^4 K for constant pressure conditions (Pottasch 1964, Cox and Tucker 1969, McWhirter *et al* 1975) (figure 2).

Above the photosphere ($T \simeq 6000$ K), the temperature declines in the low chromosphere due to the decreasing opacity of material. At the height where the temperature has dropped to around $T = 4200$ K increasing importance of acoustic energy deposition (from waves generated in the convection zone of the solar interior) causes the temperature to rise again (e.g. Ulmschneider 1977). The decline of density with height (required for hydrostatic equilibrium) demands higher temperatures for radiative losses to balance the input. Thus the temperature progresses upward through the low chromosphere from $T_{min} \simeq 4200$ K (where H$^-$ ion radiative losses are dominant) through 10^4–3×10^4 K (where hydrogen Balmer lines and continuum then Lyman emissions dominate) up to around 6×10^4 K (upper chromosphere) by which time collisionally excited lines of heavy elements have become the major loss. As may be seen from figure 2, no stable equilibrium is possible between an input and radiative losses at $T \gtrsim 6 \times 10^4$ K due to the monotonic decline of the radiative loss function under constant pressure conditions. Such material is radiatively unstable in the sense that the smallest increase in temperature from an equilibrium value with some input results in decreased radiative losses and thermal runaway to high temperatures (cf Cox and Tucker 1969). In practice this runaway is limited by the onset of thermal conduction back across the steep temperature step (transition region) formed between the cool (chromospheric) and hot (coronal) regions. The final temperature of the corona $T \simeq 10^6$ K is thus a result of balance between conductive losses and input processes—acoustic and magnetic-field dissipation. Densities in the atmosphere range
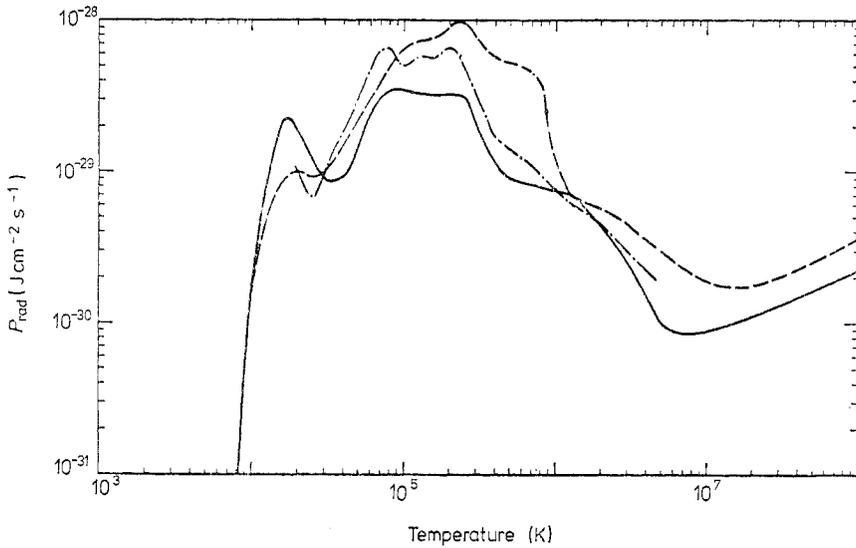
Figure 2. The radiative power loss of an optically thin plasma of solar abundances as a function of temperature (at constant density n) as calculated by McWhirter *et al* (1975) (full curve) compared to earlier results. Such results are based on detailed computation of collisional transition equilibrium and so are scaled to the square of the plasma density. For our purpose the essential feature is the existence of the maximum at $T \simeq 10^5 \text{K}$. Under constant pressure conditions this maximum moves down to $T \simeq 5 \times 10^4 \text{K}$ and the rise of P at $T \geqslant 10^7 \text{K}$ due to free–free emission is eliminated.

from around 10^{17} cm^{-3} hydrogen particles (neutrals plus ions) in the photosphere down to around 10^{11} cm^{-3} just below the transition and about 5×10^9 cm^{-3} just above.

Though this picture does contain certain essential elements of the real solar atmosphere, the detailed structure is still poorly understood. For instance, the empirical construction of plane-parallel model chromospheres from observations (Vernazza *et al* 1973) and theoretical studies of chromospheric heating (Ulmschneider *et al* 1977) are very far from unification. Likewise it has proved impossible to construct a plane model of the transition region which satisfactorily represents all the observations. One reason is that the solar atmosphere is far from horizontally stratified, containing much fine structure due to convective processes and the influence of the solar magnetic field (see papers in Athay (1974), for example). Active regions and flares are necessarily more complex still (see papers in Newkirk (1974), for example) and any one-dimensional interpretation must be treated with great reservation. Nevertheless, even in complex geometries the fundamental division of a heated plasma into high- and low-temperature regions by this instability of its radiative properties must still persist (cf Somov and Syrovatskii 1976) and is a convenient one for our discussion of the observations of thermal emissions (§§2.2 and 2.3) and of flare heating models in §4.

A vital additional feature of the flare process is its transience, which results in effects occurring over distances less than a collisional mean free path and in times less than thermal relaxation times. This is the source of the various energetic non-thermal phenomena which especially distinguish the flare, and particularly of particle acceleration. The observations associated with these effects are described in §§2.3–2.5, after we have discussed (§2.2) observations of the magnetic environment of flares.

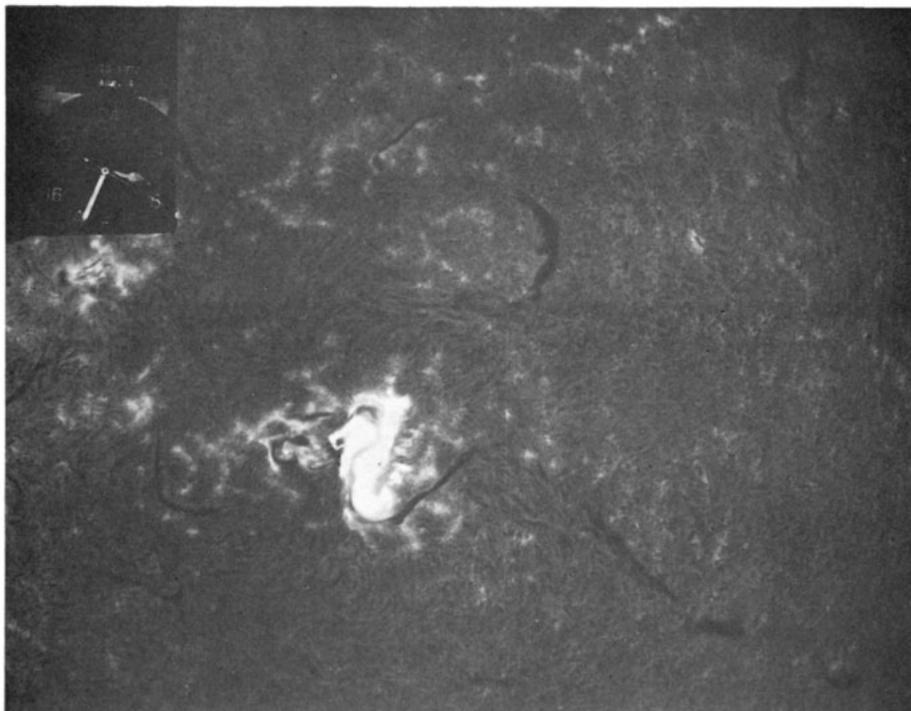*2.1.2. Low-temperature flare emission.* Morphologically, all of the chromospheric flare

Figure 3. A classic large double-ribbon optical flare which occurred around 06:34 UT on 4 August 1972. (Courtesy of H Zirin, Caltech Big Bear Solar Observatory.)

emissions appear to divide into two regions (cf Canfield *et al* 1978): a set of small ($\leqslant 10^{18}$ cm^2) bright kernels with lifetimes of a few minutes embedded in (or near) a more extensive ($\simeq 10^{19}$ cm^2) general flare brightening lasting tens of minutes or hours. General enhancement of the flare region often precedes the flash phase brightening of kernels and rapid increase in the general emission before the slow decay. Though the total energy emitted in the gradual component often substantially exceeds that in the impulsive brightenings (e.g. Švestka 1976), the large *power* of the latter and the rapidity of onset of the flash makes this stage the most demanding of theoretical mechanisms.

These two features of chromospheric emission are exhibited, for example, by radiation in Hα (Zirin and Tanaka 1973) and from the temperature minimum region (e.g. Machado *et al* 1978). Due to the ease of its observation, Hα flare emission has dominated the description of flare morphology. Detailed Hα flare morphology is extremely diverse but has provided no clear step forward in solving the central problem of flare energy release, and so is not discussed here (see Švestka 1976). The most salient features of Hα morphology seem to be the occurrence of brightenings on either side of the magnetic neutral line frequently present (cf §2.2), sometimes in the form of a double ribbon, following activation of a pre-flare filament (figure 3 (plate)). An important constraint on energy storage may lie in the occasional occurrence of two flares of very similar form in the same site (homologous flares) within a few hours.

Quantitative analyses of the Hα observations are relatively sparse. Numerous cases show the flash phase increase in Hα areas, linewidth and intensity over 10–100 s (e.g. Falciani *et al* 1968) and subsequent evolution over 10^3 s or more, but rarely a complete history of total Hα power against time (e.g. Dizer 1969). According to Zirin and Tanaka (1973) the energy in Hα is $\lesssim 10\%$ of the flare total. Spectral observations of Hα together with other hydrogen Balmer lines indicate (e.g. Švestka 1973) that they are formed around $T \simeq 8000$ K in a thin ($\simeq 10$ km) layer (or other structure) at an electron density of 10^{12}–10^{13} cm^{-3} inferred from application of the Inglis–Teller formula to Stark broadening of high-order Balmer lines. Recently, however, it has become clear that the electric-field fluctuations associated with quite modest levels of plasma turbulence might produce the same effect in much lower density plasma (e.g. Spicer and Davis 1975, Davis 1977). Similar ambiguity may arise in interpreting the red asymmetry of optical lines in terms of Doppler shifts. Hα observations also reveal mass motion phenomena in flares including wave motions emanating from disc flare sites (Moreton 1964) and mass ejection seen at the limb. These motions can involve masses of material up to about 10^{16} g and velocities up to 10^3 km s^{-1} which are compatible with observations of interplanetary blast material (e.g. Hundhausen 1972) (cf §2.5). The associated kinetic energy of up to 5×10^{31} erg thus comprises around 50% of the total thermal energy of a large flare, a fraction which is entirely reasonable on equipartition grounds for any heating mechanism, though the ejected mass component will depend on the openness (or disruption) of the magnetic field.

Machado *et al* (1978) have discussed the enhancement of the temperature minimum region emissions deep in the flaring atmosphere. On the basis of Ca K observations they conclude by plane-parallel atmosphere modelling that the minimum temperature rises some 250 K over pre-flare level and that it occurs at a column mass of the order of 0·03 g cm^{-2}, compared to 0·005 g cm^{-2} in the quiet Sun, results consistent with those inferred from Skylab observations of the Si continuum (Cook and Brueckner 1979). The dominant energy loss in these deep layers is H$^-$ ion emission and Machado *et al* (1978) have estimated that it may amount to several per cent of the total flare
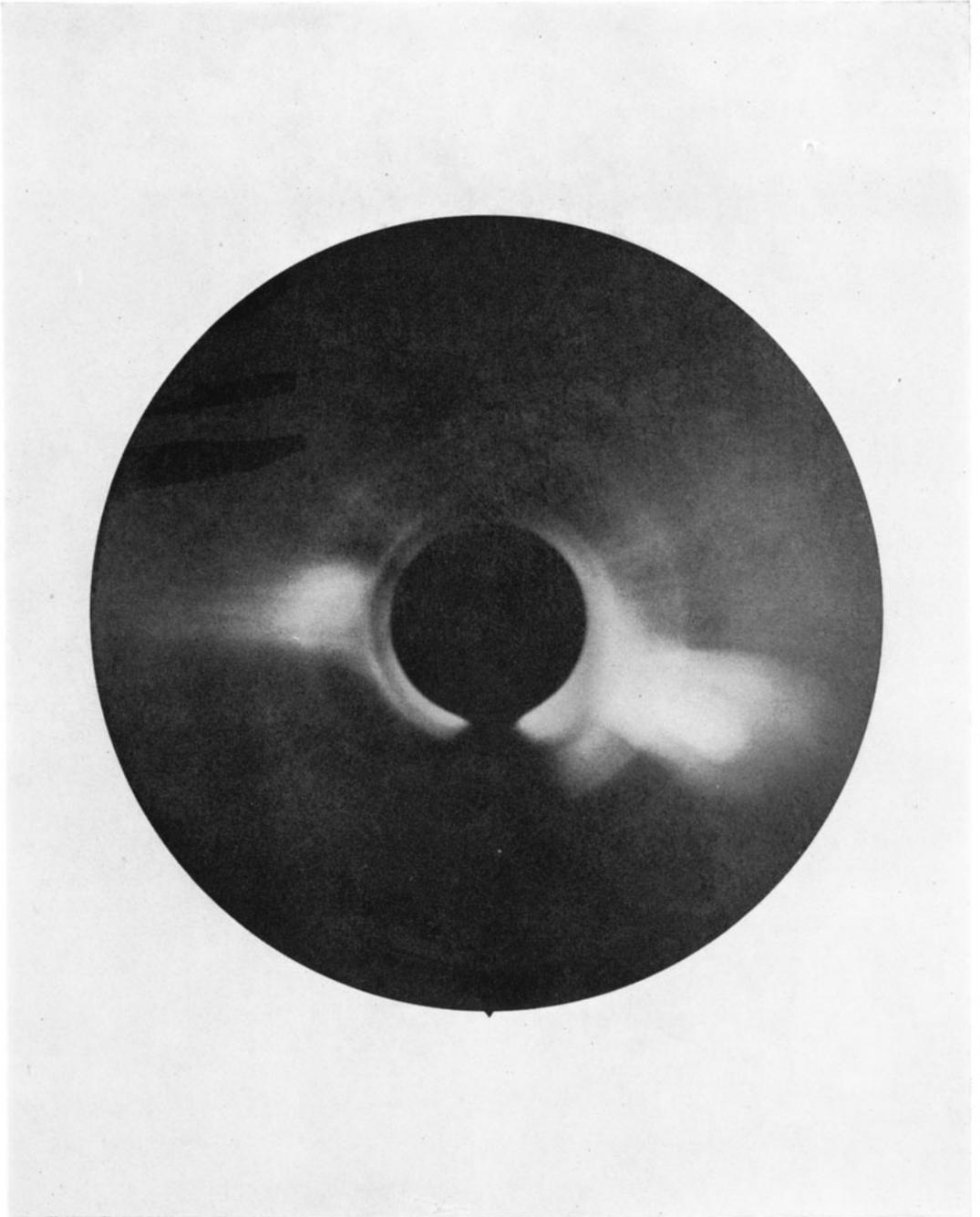
Figure 5. The coronal transient of 21 January 1974, 10:27 UT, as seen in white light by the Skylab coronagraph (from MacQueen *et al* 1976).

power. Zirin and Tanaka (1973) and Zirin (1978) have discovered the occurrence of small very short-lived ($\lesssim 10$ s) optical flashes (15 Å about $\lambda 3835$ Å) in precise synchronisation with hard x-ray burst peaks. Penetration of the flare to even deeper layers may be seen in the occurrence of 'white light flares' (e.g. Carrington 1859, Rust and Hegwer 1975) if these are interpreted as heating of near-photospheric layers. At this stage, however, some alternative non-equilibrium interpretation at higher levels (e.g. Hudson 1972) cannot be entirely excluded. Best-fit empirical models of the flaring chromosphere as a whole have been constructed by Machado and Linsky (1975) (cf Machado *et al* 1978) (see figure 4).

Analyses of the upper layers of the chromosphere in flares are still in a preliminary stage since typical emissions there, such as Lyman α, have only recently been observed in detail. A discussion of modelling of this regime of the flare has been given by Lites and Cook (1979) for example (cf Canfield *et al* (1978) and §5). Orrall and Zirker (1977)
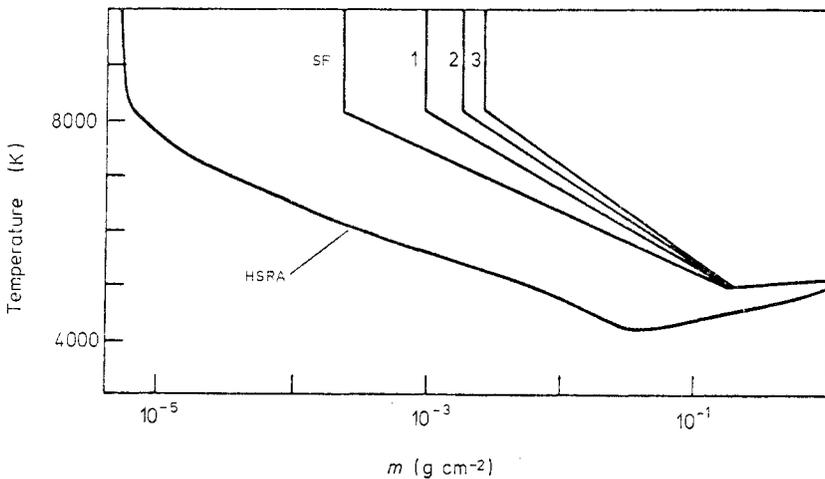

Figure 4. The temperature/height structure in chromospheric flares of various sizes according to the plane-parallel models derived empirically by Machado and Linsky (1975). An earlier quiet-Sun chromospheric model (cf figure 2) is also shown. The abscissa is the mass per cm² column overlying the depth concerned.

and Canfield and Cook (1978) have also set limits on low-energy photon fluxes using the absence of a significant non-thermal Lyman α wing in the observations.

An important additional means of studying flares at optical wavelengths is by means of the photospheric white light emission scattered by electrons in the corona, using a coronograph (e.g. Skylab—MacQueen *et al* 1976). Inter-relationships between coronal white light transients and flare radio events have been discussed in a variety of papers, as reviewed in Rust *et al* (1978). Quantitative aspects of coronograph data analysis have been discussed by Mouschovias and Poland (1978) and a very thorough investigation of the difficulties of deconvoluting three-dimensional coronal density structure from such two-dimensional data given by Wilson (1977). A typical coronograph image of a flare-associated transient is shown in figure 5 (plate).

2.1.3. High-temperature flare emissions. The short wavelengths characterising high-temperature radiation mean that this flare component has to be observed almost entirely by space-borne instrumentation (for radio observations see §2.4).

Reviews of the relevant observations up to and including those from Skylab ATM may be found in the various chapters of the Skylab Flare Workshop Report (Sturrock *et al* 1978), which we merely outline here. The EUV band (10–1000 Å) has been observed both from the ground and from space. Ground-based results (cf Donnelly 1970) are obtained indirectly by inference from sudden frequency deviations (SFD) of radio waves reflected from (or transmitted through) the ionosphere as it responds to the varying flare EUV flux. The technique has zero spatial resolution, little spectral resolution and rather poor absolute intensity calibration (due to uncertainties in modelling ionospheric response) but is an excellent means of observing EUV flux time variations in one event and for statistical studies of many events (Kane and Donnelly 1971, Donnelly and Kane 1978). Observations over a comparable wavelength range have also been made from space (e.g. Wood *et al* 1972, Emslie and Noyes 1978), but with much higher spectral resolution. The typical time evolution again shows a rapidly varying impulsive flash component closely synchronous with hard x-rays and microwaves (figure 6) and a gradual phase, the latter being shorter lived than the Hα flare. Spatially resolved data show strong localisation of the emission in sites not necessarily coincident with Hα kernels (see, for example, Neupert *et al* 1975, Widing and Cheng 1974). The total power of the EUV emissions responsible for SFD is very small (Emslie *et al* 1978, Donnelly and Kane 1978), suggesting an origin in very thin (transition) layers in the flare. In addition Donnelly and Kane (1978) have established a very substantial limb darkening of EUV intensities and constructed a simple chromospheric absorption model with non-plane-stratified geometry to describe this.
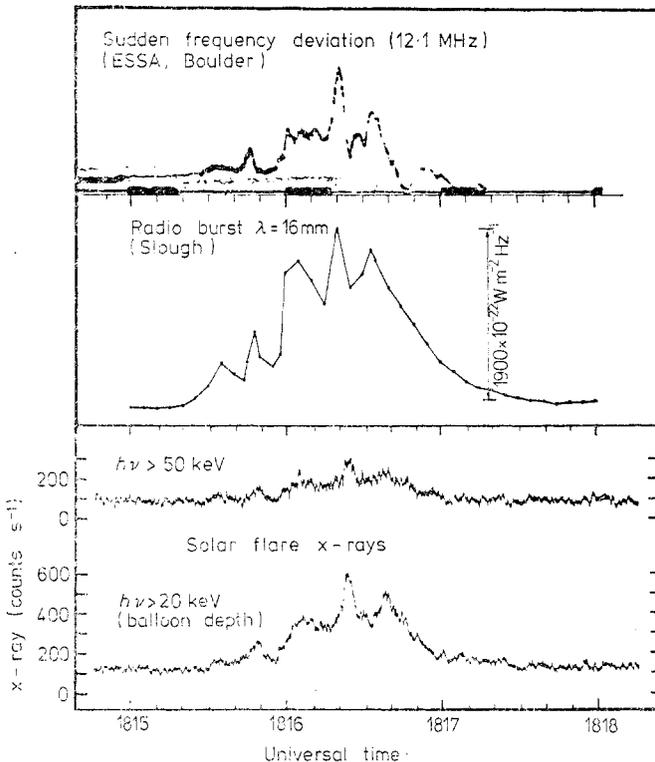

Figure 6. The close synchronisation of impulsive flare emissions in EUV, hard x-rays and microwaves illustrated here by the flare of 8 August 1968 (from Parks and Winckler 1971).

Observations of EUV lines with higher spectral resolution allow localisation of narrower flare temperature domains and inference of Doppler motions. The details of such line profile studies and their interpretation have been presented by Doschek and colleagues (as reviewed in Brueckner (1976) and Canfield *et al* (1978)) who have concluded that the line broadenings indicate either strong turbulent motions ($\simeq 500$ km s^{-1}) in flare loops or strong Stark broadening by plasma turbulence (Brueckner 1976). The latter interpretation seems excluded (for x-ray lines) by the calculations of Brown and Nakagawa (1978).

An important clue to the flare mechanism may lie in the gross morphology of hot EUV sources, which probably delineate the overall geometry of the sites of primary energy release, viz single or multiple arch-like structures (cf Cheng and Widing 1975, Brueckner 1976). Similar morphology is also seen at shorter x-ray wavelengths observed (over broader spectral bands) by means of the grazing incidence telescopes on ATM (e.g. Kahler *et al* 1975, Vorpahl *et al* 1975) (figure 7 (plate)). These arches are often low lying and of small dimensions ($\lesssim 10^4$ km). Considerable material has been published on the morphology of individual XUV events and correlations attempted between observed parameters (e.g. Vorpahl *et al* 1975), but without as yet providing any essential clue to primary flare problems. Finally, since most of these Skylab x-ray morphologies were observed after the flash phase, caution must be exercised in assuming that they characterise the overall magnetic-field geometry involved in the primary release, rather than just an aftermath. Specifically these data do not *in themselves* exclude primary energy release in a classical current sheet (e.g. Carmichael 1964, Sturrock 1966) with subsequent filling of an underlying arch with hot material (cf §§2.2 and 3).

Quantitative information obtainable in principle from the XUV data include particle density (n), temperature (T) and volume (V) of the source, and the time development of these parameters at each spatial point. Temperatures are inferred from spectral line intensity ratios (or continuum slope) on the basis of detailed steady-state ionisation equilibrium calculations for optically thin plasmas. Volumes can be inferred from spatially resolved data by assumption of comparable extent along and across the line of sight. In practice, the most intense XUV source regions are often too small ($\lesssim$ a few thousand km) to be resolved so only an upper limit to the volume can be set unless the density is independently available for use with the total emission measure n^2V (which determines absolute line intensities in a collision-dominated optically thin source). More commonly an observed value (or upper limit) for V is used to infer n (or a lower limit to it). Alternatively n has been estimated crudely by modelling the observed cooling curve for the x-ray flare in terms of some particular cooling process, usually classical conduction. Knowing T and having an estimate of the source size defines the plasma thermal energy density, hence n (e.g. Culhane *et al* 1970). Much better density diagnostics are density-sensitive spectrum line intensity ratios (forbidden/allowed) (e.g. Gabriel and Jordan 1969, Munro *et al* 1971). The method hinges on the existence of (forbidden) transitions whose spontaneous radiative lifetimes are comparable to the atomic collision time for densities in the range of interest and should thus be independent of the source volume (provided it is homogeneous). However, application of this method in its basic form (e.g. Parkinson *et al* 1977) has led to x-ray source densities ($n \geqslant 10^{13}$–10^{14} cm^{-3}) which are impossibly high at the temperatures involved ($T \geqslant 10^7$ K) on dynamic and energetic grounds (Brown and Nakagawa 1978). Since the explanation of this discrepancy cannot be found in plasma turbulent effects and probably not in radiative processes either, it most likely implies that the source
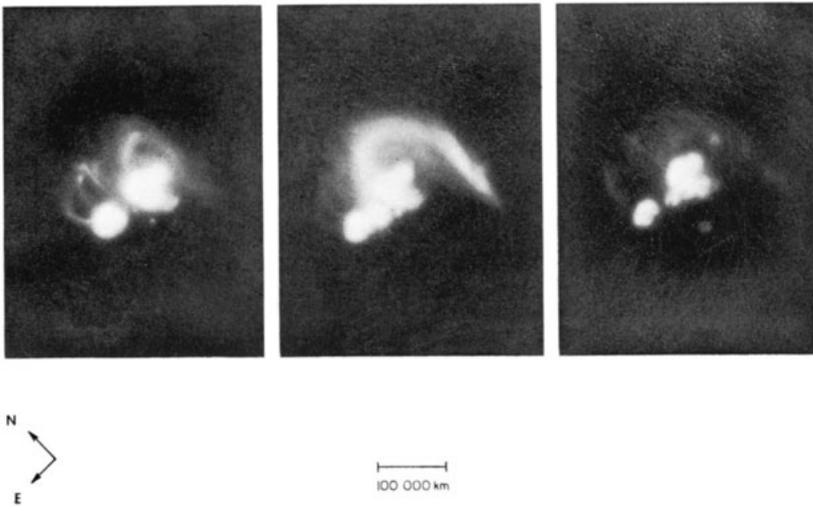
Figure 7. Several stages in the development of soft x-ray flare active region, 14–17 June 1973, as seen by the American Science and Engineering Grazing Incidence Telescope aboard Skylab. The images are separated by approximately one day each. (Courtesy of American Science and Engineering, Cambridge, Mass.)

comprises a large number of very small regions which successively emit while in a state of *transient* ionisation equilibrium arising from the dynamic nature of the primary energy release process (Brown and Nakagawa 1978, cf Shapiro and Moore 1977). On the basis of this interpretation the densities $n \simeq 10^{11}$ cm^{-3} inferred by the crude methods mentioned above seem the most likely (cf Moore *et al* 1978), with corresponding kernel size about 10^3–10^4 km (cf §4).

It is of course clear from the very large wavelength range over which a flare emits that the plasma is by no means homogeneous and, in particular, has neither a single density nor temperature. So far, neither instrumentation nor theoretical diagnostics have been capable of giving temperature or density measurements in volumes small enough to be homogeneous. (Indeed, they may possibly never be—cf §6.) Therefore much effort has been made to infer the inhomogeneous temperature structure within any observed volume from the deviations of line or continuum spectra from their isothermal forms. (As yet no such inhomogeneous analysis has even been formulated for the density structure.) The only quantity in this category which can, in principle, be derived from spectral information is the 'differential emission measure' or very crudely $\xi(T) = (\mathrm{d}/\mathrm{d}T)(\int_V n^2 \mathrm{d}V)$ which is the amount of the total emission measure $n^2 V$ per unit temperature range. (Optically thin collisionally excited radiations have power $n^2 f(T)$ per unit volume at temperature T and density n. The total emission from an isothermal volume V thus varies as $f(T) \int_V n^2 \, \mathrm{d}V$ where the integral is called the emission measure.) The completely general definition of $\xi(T)$ given by Craig and Brown (1976a) shows how far removed knowledge of $\xi(T)$ is from what we would really like to have, namely the distributions of T and n in spatial coordinates which govern conductive and convective energy transport. (Specifically $\xi(T)$ involves an inverse weighting with respect to the unknown temperature gradient, and summation over all disjoint isothermal surfaces.) Thus $\xi(T)$ is only useful when combined with rather strong modelling assumptions such as spatially monotonic T variations and simple relations between n and T (e.g. constant pressure). To make matters worse, it is in practice impossible to get more than a very crude assessment of $\xi(T)$ from the spectral data, since these are related by an ill-posed integral (or matrix) equation (Craig and Brown 1976a). The broad kernel of the equation (essentially the Maxwellian electron distribution) acts as a broad band filter on the temperature structure function $\xi(T)$, removing all information on its finer structure from the spectrum. This state of affairs means that any simple procedure such as model fitting of $\xi(T)$ to the data (e.g. Dere *et al* 1974, Underwood *et al* 1978) may yield an entirely spurious 'solution' for $\xi(T)$ through the misleading fit to the data which characterises all such problems (Craig and Brown 1976b) unless $\xi(T)$ happens to be smooth (cf Dere 1978). Furthermore, even the correct mathematical approach of optimised stable inversion using *a priori* constraints (e.g. Craig 1978) cannot yield detailed information on $\xi(T)$, which is absent from the data, except insofar as it is introduced in the constraints. Nevertheless, even very crude information on the form of $\xi(T)$ can place useful constraints on the mechanisms of energy redistribution in the flaring plasma (e.g. Craig *et al* 1978, Underwood *et al* 1978) as we discuss further in §5.

Finally, an important means to reduce these interpretive difficulties is to utilise the dynamic evolution of the spectrum rather than just its instantaneous form. In its simplest form this consists of returning to the homogeneous assumption and modelling the spectrum in terms of the time evolution of total emission measure $n^2 V$ and mean temperature T. Except at hard x-ray energies (cf Mätzler *et al* 1978) such analysis is found to show a steady increase of emission measure as the temperature falls away
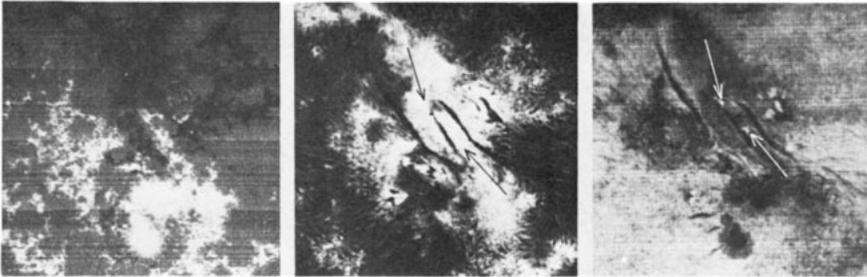
Figure 8. Three simultaneous views showing the photospheric magnetic field (bright—positive polarity; dark—negative polarity) and two line emissions for the small two-ribbon flare of 7 May 1974, 15:20 UT. The area shown is about 240 arcsec square. Arrows point to the flare kernels (after Rust 1976).

from its peak value (e.g. Datlowe *et al* 1974b, Zaumen and Acton 1974, Underwood *et al* 1978).

2.2. *Magnetic fields*

There are no direct measurements of magnetic fields in flares since line radiation with detectable Zeeman splittings which lead to a measurement only comes from the photosphere. Since the flare is a chromospheric and coronal phenomenon, our knowledge about magnetic fields in flares presently comes from an extrapolation of the photospheric field. It is to be hoped that this situation will change in the near future with ultraviolet (UV) polarimeters extending measurements into the transition region. Since until very recently only the line-of-sight component of the photospheric magnetic field could be reliably measured and since measurements were made only once per day, the problem of extrapolation was essentially hopeless. It is unlikely to improve in the near future because instruments which can determine *all* components of the magnetic field are not being designed to make measurements sufficiently rapidly. This state of affairs is due mainly to the precedence taken by fast scanning line-of-sight magnetographs which allowed only morphological relations between the photospheric magnetic field and flares to be established (Rust 1976). We briefly review these relations.

Flares occur in so-called active regions which are characterised by sunspots and bundles of emerging magnetic field called pores. The first question we can ask is whether the photospheric magnetic field changes at the time of a flare? The only known source of flare energy is stored magnetic energy (cf §3.1) and the 10^{32} erg required for a large flare could be obtained by lowering a 3000 G sunspot field by 20% in a volume $(10^4 \text{ km})^3$. Observationally the problem is that active regions are continuously evolving (irrespective of flare occurrence) and the best observations have only been made about once per hour. As a result, Rust (1976) concludes that the rates of change in sunspot fields required for flares do not differ significantly from normal active region growth or decay rates except possibly in the largest proton flares. This type of proton producing flare is so rare that convincing statistics on the rates of change of nearby sunspot magnetic fields have not yet been compiled.

A typical example of a magnetogram from a small flare and simultaneous observations in the lines of Hα and He 10830 Å is shown in figure 8 (plate). The magnetogram shows many lines of distinct polarity reversal called neutral lines which are a characteristic feature of every active region which produces flares. Flare activity is often but not invariably coincident with the location of neutral lines.

In He 10830 Å one can clearly see two patches of bright emission called flare kernels. These are also, but less distinctly, visible in Hα at the same locations. Each kernel lies just off the centre of a sunspot umbra. Further study showed that the flare kernels brightened close to emerging pores, implying that some flares are associated with emerging magnetic field. Flares are also associated with complex patterns of photospheric magnetic fields rather than absolute field strength. These two facts are all that can be said based upon direct measurement of magnetic fields. Considering the sizeable effort that has gone into obtaining these conclusions, the present developmental approach for vector magnetographs, which are not being designed to make measurements sufficiently rapidly to obtain new results for flares, is understandable because there is no guarantee that useful new results would be obtained. Nevertheless, it can be shown from equation (3.9) that to extrapolate the chromospheric and coronal

magnetic fields on a sound physical basis, both the photospheric vector magnetic and velocity fields are required. Thus the development of fast vector magnetographs operating simultaneously with vector velocity measurements appears to be well worth the effort.

An indirect method of inferring the magnetic fields in flares is to assume that $H\alpha$ fibrils and filaments outline magnetic-field structures (e.g. Tanaka and Nakagawa 1973). When these are combined with sunspot pictures and magnetograms, some general characteristics of the three-dimensional structure of the magnetic field may be inferred.

Filaments are large masses of cool gas which require magnetic support and are associated with flares. It seems reasonable to assume that the ionised material in filaments is frozen-in to the magnetic-field lines and constrained to move with them except where the magnetic field vanishes so that it becomes a 'tracer' of changes in the gross field topology. The beginning of a flare is sometimes signalled by the ejection of a filament to greater heights and its subsequent disappearance. The mass of a filament is sufficiently large that only magnetic forces could eject it. For example, the typical mass of a flare filament (which is also known as an active prominence) is $M \simeq 10^{16}$ g with an average density $n = 10^{10}$ cm^{-3}, volume $V = 6 \times 10^{29}$ cm^3 and temperature $T = 10^5$ K (Tandberg-Hanssen 1967). The energy required to eject this prominence to a height $h = 100\,000$ km is

$$\mathscr{E} = M g_\odot h = 2 \cdot 8 \times 10^{30} \text{ erg} \qquad (2.1)$$

where $g_\odot = 2 \cdot 8 \times 10^4$ cm^{-2} is the solar surface gravity. On the other hand, the total thermal energy of the prominence which could be turned into a pressure force to raise it is

$$\mathscr{E}_{\mathrm{T}} = nKTV = 8 \cdot 3 \times 10^{28} \text{ erg} \qquad (2.2)$$

which is a factor of 33 less than $\mathscr{E}$ (K is the Boltzmann constant).

The development of the $(\nabla \times \boldsymbol{B}) \times \boldsymbol{B}$ ejection forces could only occur by a change in the magnetohydrodynamic equilibrium which provides unambiguous evidence for a restructuring of the field (here $\boldsymbol{B}$ is the magnetic-field vector). However, several different restructurings could lead to the ejection of the filament and there is no observational way to decide on the relative likelihood of these possibilities.

In summary we know that some flares are associated with the emergence of new photospheric magnetic flux and that field restructuring occurs in some flares. However, we do not know whether the flare would have occurred without the emergence of new flux or whether the new flux acted as a trigger. For example, the pores of the new flux could simply have acted as a concentrating funnel for the energy released higher up which would produce the He 10830 Å and $H\alpha$ kernels, but only as a secondary effect. The emerging pores most likely only act as a trigger for the release of energy already stored higher in the atmosphere since it is unlikely that all of the emerging flux could concentrate in a small region in the low corona. Our knowledge of field restructuring in flares is equally ambiguous since we have no unique way of deciding which restructuring is necessary to fit the $H\alpha$ observations.

2.3. Hard x- and γ-rays

Much interest in flares has recently centred on the physics of hard x-ray burst production chiefly because the source electrons appear to carry more energy than any other energetic particle component in the flare, and possibly a substantial fraction of

the total flare energy (cf §5.1.1). Thus the properties of these electrons are intimately related to the efficiency of particle acceleration in the primary flare process (cf §5) and to the redistribution of primary energy through the thermal flare in the flash phase (cf §4). In some cases second-stage acceleration (§5) may also be indicated in the hard x-ray data. We therefore survey the observational material in some detail here. Quantitative theoretical aspects appear in §5. The most recent reviews of hard x-ray burst data and their interpretation are by Kane (1974), Brown (1975, 1976) and Kane *et al* (1978).

Since hard x-ray sources are currently observed without spatial resolution, the following description is of total flare emission. As the observed photon energy is increased, x-ray bursts exhibit an increasing dominance of the impulsive temporal component over the gradual evolution which dominates at lower energies (figure 9). This transition between 'soft' and 'hard' emission occurs at an energy typically in the range 5–20 keV and the change from gradual to impulsive behaviour is often taken to indicate a distinction of thermal from non-thermal emission (e.g. Kahler 1975) but this need not be the case (cf §5.1.1). At the higher energies ($\gtrsim 20$ keV) the impulsive time profile shows a complexity broadly increasing with burst size from simple single spikes lasting less than a minute (Crannell *et al* 1978, cf Kane and Anderson 1970, Datlowe *et al* 1974b, Hoyng *et al* 1976) through to large complex structures lasting
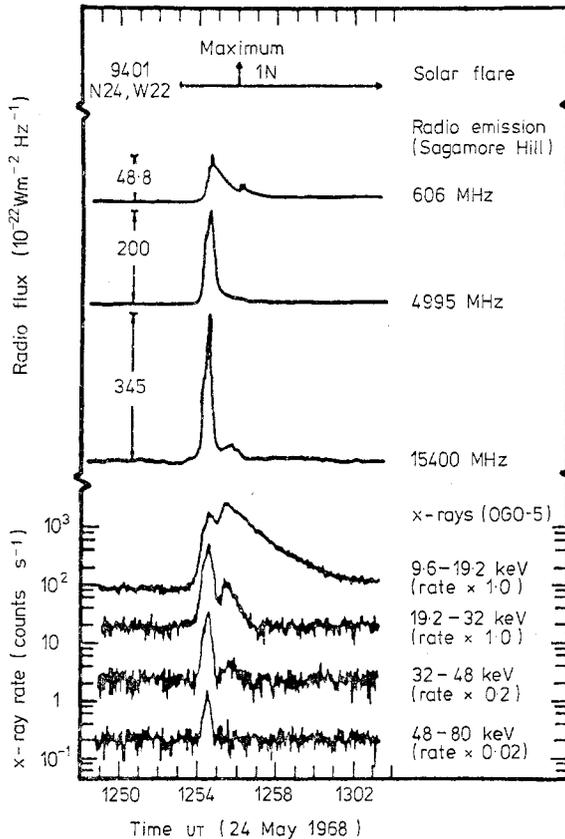

Figure 9. Time evolution of hard x-ray intensity at various photon energies in a typical small burst (from Kane and Anderson 1970).

tens of minutes, e.g. figure 10 (Hoyng *et al* 1976, cf Frost 1969). The prolonged bursts appear to show a second phase of emission (Frost and Dennis 1971) more gradual than the impulsive spikes, extending to energies of many hundred keV. Hoyng *et al* (1976) have shown by Fourier analysis and statistical arguments that a typical large burst involves no flux e-folding time scales less than about 10 s, down to their time resolution (1·2 s) and that there are periodic components present at larger time scales (cf also Frost 1969, Parks and Winckler 1971). The apparent simplicity of small bursts may merely indicate that an intrinsically complex structure is scaled down in time so that the details are unresolved in most cases (cf microwave bursts at high resolution, e.g. Kaufmann *et al* (1977)). Crannell *et al* (1978) have argued that simple spikes show a symmetry of rise and fall but their conclusion appears to be a statistical behaviour rather than a property of individual spikes. (The rise and fall times are estimated from about three points each.)

In judging the significance of some of these temporal data and also of the spectral distribution discussed below, it is important to recognise some practical limitations on them. The scintillation detectors used at higher energies are limited both in spectral resolution (typical effective channel widths in keV at photon energy ϵ keV being a few times $\sqrt{\epsilon}$) and in pulse pile-up at high photon flux whereby several low-energy photons accumulated in a single sampling time appear as a single one of high energy. The problem can be removed by use of a suitable window to exclude the low-energy flux, e.g. the results analysed by Hoyng *et al* (1976), though this permits observation
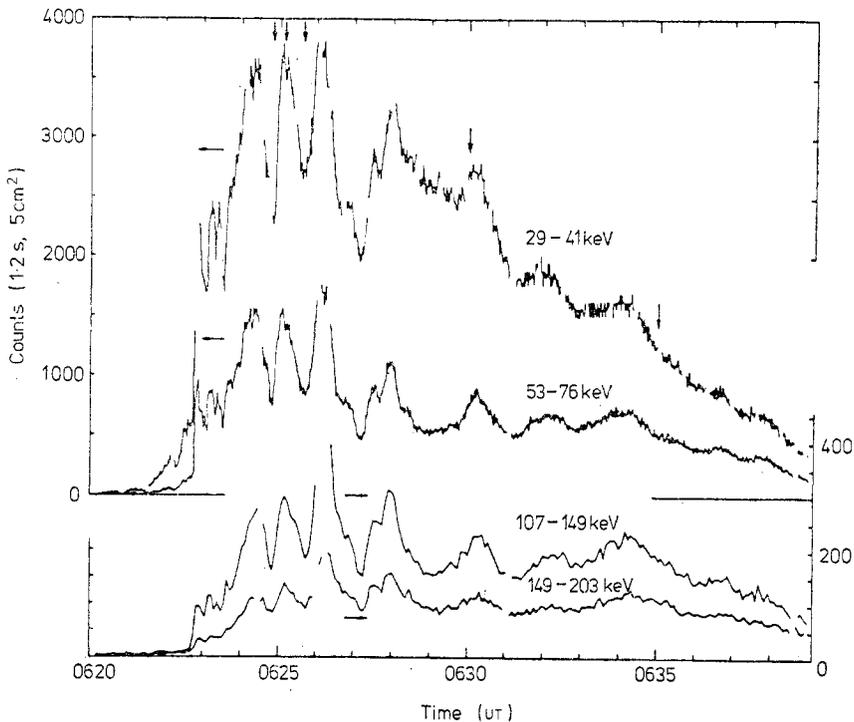

Figure 10. Time evolution of hard x-ray intensity at various energies in the large extended complex burst, associated with the 4 August 1972 flare of figure 3 (from Hoyng *et al* 1976).

10

only of larger events and severely limits interpretation of the spectrum (Brown 1978). Unless this and other precautions are taken (e.g. Frost *et al* 1971) pile-up can lead to entirely misleading results such as pile-up of a gradual soft x-ray component appearing in higher channels like a prolonged hard x-ray phase. In the case of the discovery of a second hard x-ray stage by OSO-5 (Frost and Dennis 1971) the precautions taken with the instrument (Frost *et al* 1971) and subsequent simulations (K J Frost personal communication), however, leave little doubt as to its reality.

For small hard x-ray bursts spectral information is generally flux-limited to a few channels, or well under one decade in energy. It is thus to be expected that any two-parameter function can be fitted quite well to the data. This is usually taken to be the power law form $I(\epsilon) = a\,\epsilon^{-\gamma}$ (photons cm^{-2} s^{-1} keV^{-1} at the Earth) where the index is a measure of how soft (large γ) or hard (small γ) the spectrum is. γ values occur almost uniformly in the range 2–6 (Kane 1974) except at small values, which occur only in very large events. Some data (e.g. Crannell *et al* 1978) have instead been better fitted by the exponential (thermal) form $(b/\epsilon)\exp(-\epsilon/\epsilon_0)$ with $\epsilon_0 \simeq 10$–100 keV. Where the photon flux is adequate to be detectable over greater than about one decade in energy, there is some evidence of a break (increased γ) in the power-law spectral fit above about 60–100 keV (Frost 1969, Kane and Anderson 1970, van Beek *et al* 1973). Evolution of the spectral hardness with time is a matter of some controversy. Kane and Anderson (1970) originally claimed that both rise and fall times of single spikes decreased with increasing photon energy, i.e. that γ decreased to burst peak and then increased. Datlowe *et al* (1974a) support this conclusion to some extent but Hoyng *et al* (1976) found no evidence for it. Perhaps the most convincing result on dynamic x-ray spectra is that for the very large event of 4 August 1972 (Hoyng *et al* 1976) where the flux was adequate for the accurate determination of γ every 1·2 s. Hoyng *et al* (1976) and Benz (1977) found a systematic, though complex, locus traced by the event in the flux/spectral index plane.

Attempts have been made to measure the linear polarisation of hard x-ray bursts (Tindo *et al* 1970, 1972a, b, 1973, Nakada *et al* 1974) with a view to assessing anisotropy effects in the source. The apparently positive results must be considered tentative since the degree of polarisation measured has declined from 30% or more to $\lesssim 10\%$ with improvement in the experiments. It is proposed (Tindo *et al* 1972b) that the plane of maximum intensity is approximately radial on the Sun's disc. Directivity of the emission has been investigated statistically in terms of the distribution of burst flares and spectral indices with flare location on the disc (cf Kane 1974, Datlowe *et al* 1975). Results are too marginal to be sure that they are not some effect of Hα visibility in locating the burst on the Sun.

Finally there is some indirect evidence on the source geometry. Takakura *et al* (1971) obtained a one-dimensional collimator observation showing the hard x-ray source to lie near a line through an Hα kernel in one flare. The close synchronisation of hard x-ray burst peaks, optical flashes (Zirin 1978) and EUV flashes (figure 6) indicates an intimate connection between chromospheric flare activity and at least some part of the hard x-ray emission. On the other hand, the observation of hard x-rays from flares behind the solar limb (e.g. Roy and Datlowe 1975, McKenzie 1975, Hudson 1973, 1978), their synchronisation with microwave bursts (cf §2.4) (figure 6) and association with interplanetary electrons (Lin 1974b) all suggest a source component rather high in the atmosphere ($\gtrsim 10^4$ km). Interpretation of these data is further complicated by the contribution to bursts from photons Compton back-scattered from the dense chromosphere (Tomblin 1972, Santangelo *et al* 1973), which

significantly modifies the intensity, spectrum, directivity and polarisation of the primary source, regardless of the x-ray emission mechanism (see, for example, Langer and Petrosian 1977, Bai and Ramaty 1978).

It is generally presumed that hard x-ray bursts are generated by collisional bremsstrahlung of energetic electrons (of energies slightly greater than the burst photon energies) in the flare plasma. The alternative mechanisms of synchrotron or inverse Compton radiation (Korchak 1971) require much less total electron energy (Brown 1976) and are not excluded by the hard x-ray data themselves. However, the individual electron energies required ($\gtrsim 1$ GeV for synchrotron and $\gtrsim 10$ MeV for inverse Compton radiation) do not seem consistent with microwave bursts nor with interplanetary electron fluxes. Secondly, they are very difficult to accelerate for various reasons: difficulty of containment; high energy losses; narrow range of wave phase velocities ($\simeq c$) capable of causing resonant acceleration (cf §5). Henceforth we will therefore adopt the bremsstrahlung interpretation. The actual flare luminosity in hard x-rays is very small, typical burst peak intensities corresponding to about 10^{22}–10^{25} erg s^{-1}. However, the power carried by the source electrons may be very much higher than the radiated power by a factor of 10^3–10^6, depending mainly on the energy losses present. It is for this reason that the energetics of hard x-ray electrons is an absolutely vital issue in considering the flare mechanism, as discussed in §5.1.1.

Observations have recently been successfully made at energies even higher than the 0·5–1 MeV hard x-ray continuum seen in large flares. Chupp *et al* (1973) have detected several γ-ray lines at these energies in the flare of 4 August 1972 including the 0·511 MeV positron annihilation line and the 2·23 MeV line of neutron capture. These were observed at the same time ($\pm$ 3 min) as the hard x-ray rise and in particular the 2·23 MeV line indicates the flash phase acceleration of $\gtrsim 30$ MeV protons. Interpretation of the γ-ray fluxes in terms of proton fluxes has been discussed in detail by Wang and Ramaty (1974) and Ramaty *et al* (1975), in terms of neutron production and capture processes, indicating a total number of > 30 MeV protons in this large flare around 10^{33} ($\simeq 5 \times 10^{28}$ erg). Ramaty *et al* (1975) have also put some rough constraints on the proton spectrum by analysis of the various γ-ray line intensities. Though these results may provide a more complete picture of the flare process and useful clues on acceleration mechanisms, the protons involved are not so important in terms of total energy as the hard x-ray source electrons nor in terms of individual energy as the flare cosmic-ray nuclei (§2.5). Finally Canfield and Cook (1978) have applied the Orrall and Zirker (1978) method of seeking a (Doppler-shifted) non-thermal Lyman α wing as evidence for beamed non-thermal protons in the flare but found none. Since the protons must have already been degraded to $\simeq 30$ keV before formation of Lyman α by electron pick-up, the relationship of this result to the (high) proton energies at injection is highly uncertain.

2.4. Radio bursts

Solar radio bursts which have been recently reviewed by Rosenberg (1976) are one of the secondary effects of fast particle production in flares which, along with interplanetary particles, give us some of the most stringent requirements for particle acceleration. Because all of the bursts can be produced by energetic electrons, the requirements can be limited to electrons. Since among several different kinds of bursts, only the electrons associated with type III bursts have been measured directly, the remaining bursts provide evidence on electrons which are either trapped near the

Sun or so rarely reach close to the Earth that no electron measurements have been made to date.

Until recently, radio bursts have been measured mostly in the microwave and metre wavelength bands from the ground. The advent of satellites allowed extension to hectometric and kilometric wavelengths. We start with microwave bursts and move to progressively longer wavelengths or lower frequencies. Microwave bursts are a type of continuum radio burst, so-called because they extend over a broad range of frequencies from a few tens of gigahertz to several hundred megahertz without any spectral structure. Microwave bursts can be classified into impulsive bursts, gradual bursts and microwave type IV bursts (Wild *et al* 1963). Impulsive bursts have a time scale of 1–5 min and brightness temperatures up to 10^9 K. (The brightness temperature of solar radiation is the equivalent temperature which a black body would have which emitted radiation of the same intensity at the same frequency.) Gradual bursts have a time scale of tens of minutes and brightness temperatures up to 10^6 K. Microwave type IV bursts have a time scale of 5 min to half an hour and brightness temperatures up to 10^9 K.

The impulsive microwave bursts are closely correlated with hard x-ray bursts (see figure 6). The accepted radiation mechanism for these bursts is gyrosynchrotron radiation due to electrons with energies greater than about 100 keV spiralling in a magnetic field (Ramaty 1969). The major problem in interpreting these bursts is that several factors affect their intensities and spectra, and we have no independent handle on many of them. The most important of these are non-uniformity of the magnetic field and various low-frequency absorption mechanisms such as gyrosynchrotron self-absorption (Takakura 1972). As a result interpretations of impulsive microwave spectra become extremely model-dependent and offer no unique way of increasing our knowledge of the electrons producing them. For example, it was once thought that hard x-ray bursts required at least 100 times more non-thermal electrons than implied by the impulsive microwave bursts (Holt and Cline 1968). Takakura (1973) has shown that the combined effects of a steeper electron distribution above 100 keV, gyrosynchrotron self-absorption, and a non-uniform magnetic field can remove this discrepancy. However, with so many free parameters the weight that can be placed on this result is small until the parameters are determined independently. Since the magnetic field is the key parameter and we have seen that it is extremely poorly known at low coronal heights (§2.2), this ambiguity is likely to remain for some time.

Gradual and type IV microwave burst interpretations are subject to the same uncertainties as for impulsive microwave bursts with the additional uncertainty of no information from hard x-ray bursts.

From the combined analysis of hard x-ray and impulsive microwave bursts (Holt and Ramaty 1969, Takakura 1972, Crannell *et al* 1978), we can place the following limitations on their source if it is assumed to be common. The electron density n_e $\lesssim 10^{10}$ cm^{-3} and the magnetic field $B \lesssim 350$ G. The total number of electrons with energies greater than 100 keV, $N(> 100 \text{ keV}) \lesssim 10^{34}$, and a power-law spectrum with exponent 3–5 is a possible, but by no means unique, fit to the data for the reasons given above and in §§2.3 and 5.1.1. The total amount of energy involved will be discussed in §5.1.1. The total electron energy required to produce a large impulsive microwave burst alone is 1–2×10^{29} erg.

Metre wave radio bursts have been classified into five types (Wild *et al* 1963) on the basis of their dynamic spectra as shown in figure 11. A dynamic spectrum is a plot of intensity at a range of frequencies against time made by a receiver whose frequency
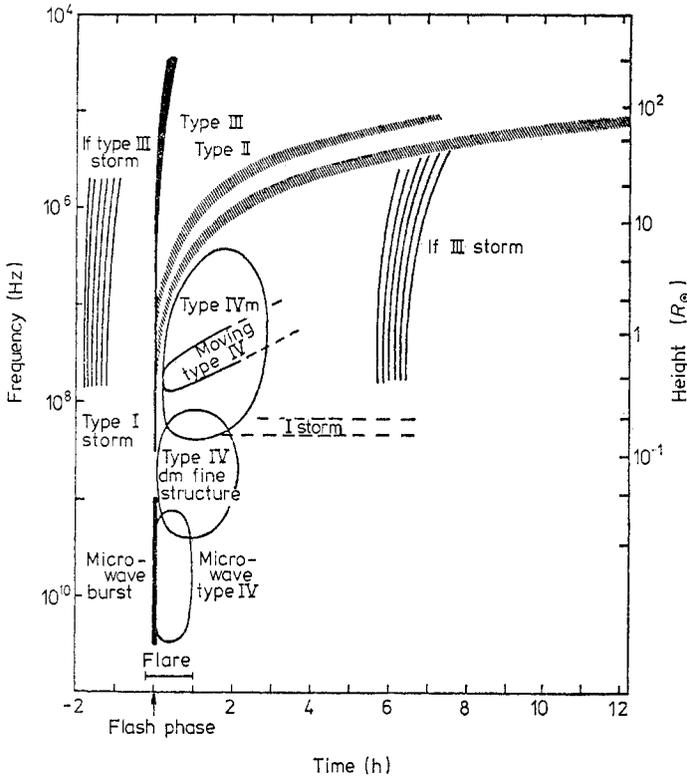
Figure 11. A schematic representation of the radio spectrum during and after a large flare. The low-frequency type III and type I storms preceding and following the flare are not necessarily ingredients. Only one type III burst has been drawn although a group of approximately ten occurs at the flash phase. Only the envelopes of the respective type IV bursts have been drawn, usually only parts of them are filled. The height scale on the right-hand side corresponds to the plasma level of the frequency scale on the left-hand side (after Rosenberg 1976).

is swept rapidly in time. Type I noise storms are the most persistent form of solar activity at metre wavelengths and are not flare-associated. They consist of a long series of bursts, sometimes accompanied by continuum radiation. Since they are not flare-associated, we shall not discuss them further but refer the interested reader to Elgaroy (1977) for more information. Type III bursts consist of a short duration intensity spike which drifts in a few seconds from high to low frequencies as shown in figure 11. It is known by *in situ* measurements in the solar wind that these bursts are caused by streams of electrons with energies 10–100 keV. This is consistent with the inter-pretation of the frequency drift as the decrease of plasma frequency with coronal density as such electrons traverse a density scale height $\sim 10^{10}$ cm. These electrons travel from the Sun to the Earth and an unknown distance beyond. When they travel essentially directly from the Sun to the Earth along the interplanetary magnetic field, they are called scatter-free events, and when they are scattered significantly by inter-planetary irregularities, they are called diffuse events (Lin 1974a). The scatter-free events provide us with one of the most direct links to electron acceleration processes at the Sun and simultaneously measured electron and type III events provide a direct test of the type III theory which will be discussed in §5.3.

Besides dynamic spectra, additional information about radio bursts can be gained from interferometry which provides source positions and sizes. Interferometry has provided us with the following additional information about metre wave type III bursts (Stewart 1974). Although the effect is rarely distinguishable on dynamic spectra due to the rapid drift, type III bursts produce radiation near both the fundamental and the second harmonic of the plasma frequency ω_{pe}, which is the characteristic oscillation frequency of the electrons, viz $\omega_{pe} = (4\pi\, n_e e^2/m_e)^{1/2}$, where n_e is the electron density. Thus, to the extent that fundamental and harmonic radiation can be distinguished in type III bursts and to the extent that the propagation of the radiation between the source and the observer is well understood, these bursts provide us with a density probe of the corona from the Sun to the Earth. Fortunately, there is evidence (Smith 1970) that radiation at metre wavelengths is primarily close to the fundamental and radiation at hectometric and kilometric wavelengths is primarily near the second harmonic (Fainberg *et al* 1972). However, progagation near the fundamental is complicated by scattering (Riddle 1974), so that this diagnostic is only reliable far from flare regions.

The total number of electrons above 20 keV required to produce a type III burst estimated from direct measurements at 1 AU (Lin 1974a) corresponds to a total energy of 5×10^{24} erg. These electrons must be produced in about 0·1 s as evidenced by the duration of a burst at 200 MHz. This time scale is the shortest known time scale for particle acceleration in solar flares (see, for example, Rosenberg (1976) for a discussion of type III fine structure).

Because other types of metre wave radiation have provided less quantitative information on the energy release process in flares, we treat them in less detail, referring the interested reader to Kundu (1965), Zheleznyakov (1970), Wild and Smerd (1972) and Rosenberg (1976) for more information. Type II bursts consist of slow drifting bands near the fundamental and second harmonic of the plasma frequency (figure 11). When the drift rate is converted into an effective radial velocity as for type III bursts, a velocity in the range 800–2000 km s^{-1} is obtained which has been identified with a collisionless magnetohydrodynamic (MHD) shock wave ascending through the corona (Wild *et al* 1963). The electrons producing the radiation are accelerated in the shock wave as a secondary effect and the main information we gain from type II bursts is that large flares must produce shock waves.

Type IV bursts are a very complex type of continuum radiation which typically occurs after type II bursts in large flares and is well correlated with large proton flares (Švestka 1976). Some of the radiation must be produced near the plasma frequency due to its high brightness temperature and some of it must be synchrotron radiation due to the high degree of polarisation (Wild and Smerd 1972). The part of the type IV burst with a high degree of polarisation moves progressively outward through the corona, in some cases with a velocity in the range 100–1400 km s^{-1}. It is called a moving type IV burst and in some cases, from the development of the polarisation, it is inferred that a plasmoid with its own self-contained magnetic field is ejected from the flare. It is not known whether the electrons trapped in the plasmoid and radiating are accelerated in the flare proper or as a secondary effect due to the plasmoid motion but energies of the order of 200 keV and magnetic fields of the order of 10 G are required (Robinson 1974). Other types of moving type IV bursts are expanding magnetic arches and advancing shock fronts (Wild and Smerd 1972). Acceleration must occur continuously in these types because they could not contain electrons for the half-hour involved. The main information to be gained from the moving

type IV burst is that flares give rise to a variety of MHD ejecta which at least in some cases must be the site of secondary acceleration processes. Another phenomenon relevant to particle acceleration in type IV bursts is regular intensity pulsations of a stationary part of the burst (McLean *et al* 1971). They occur as a peculiar phase of the event often simultaneous with the late part of a type II burst when the shock wave has reached a height of 1–2 $R_\odot$ above the photosphere and can be explained as the interaction of a moving shock wave with a stationary magnetic flux tube. The shock wave leads to standing MHD waves within the flux tube which modulate the radiation of the electrons trapped in the tube giving rise to the pulsations (McLean *et al* 1971).

Type V bursts are a continuum which sometimes occurs after a metre wave type III burst. The radiation must be near the plasma frequency or its second harmonic because of the high brightness temperature. Type V bursts are due to the trapping or delaying of part of the electrons ejected in the combined type III–V phenomenon, but the details of the delaying process remain obscure.

In summary, radio bursts are quite varied and provide us with important information about both primary and secondary particle acceleration processes in solar flares. They also indicate which kind of ejecta and MHD disturbances the flare must produce. Unfortunately, detailed theories of most of the bursts have too many free parameters to provide quantitative information except within rather broad limits. Thus the radio burst data must be combined with other observations, which can independently determine these parameters, to be of much quantitative use.

2.5. High-energy particles and mass ejecta

2.5.1. Introduction. High-energy electrons and ions observed by means of satellites provide us with a direct link to acceleration processes on the Sun in those cases where the propagation properties of the near-Sun plasma and interplanetary medium are well understood. As such they are potentially of great importance. The main problems in their interpretation are that the propagation properties of the near-Sun plasma and interplanetary medium are often poorly known and the big flares which are copious producers of high-energy particles also tend to eject large amounts of mass which disturb the quiet-time propagation properties. When the ejected mass which travels as a blast wave or a piston-driven shock reaches the vicinity of the Earth, it compresses the magnetosphere of the Earth which in turn leads to a perturbation of the ionosphere giving rise to communication blackouts and other effects (Matsushita 1962). The ejected mass also makes it harder for galactic cosmic rays to reach the Earth, an effect known as the Forbush decrease (Forbush 1946). Such terrestrial disturbances and also direct interplanetary measurements indicate a total mass of 10^{15}–10^{16} g and velocities $\simeq 1000$ km s^{-1} in the blast (Hundhausen 1972). Solar high-energy particles are generally broken into four groups: non-relativistic electrons ($E \lesssim 200$ keV), relativistic electrons, protons and other nuclei. We consider electrons, and protons and other nuclei in turn, concentrating again on those observations which put well-defined requirements on the flare mechanism.

2.5.2. Non-relativistic electrons. Non-relativistic electrons are the high-energy particles most often accelerated in solar flares and originate in the range from phenomena too weak to be classified as a flare to the most energetic proton flares. Their properties have recently been reviewed by Lin (1974a,b). Unlike the other high-

energy particles which generally travel through the interplanetary medium diffusively, non-relativistic electrons often travel directly along the field lines in the so-called 'scatter-free' mode, i.e. they appear to have mean free paths $\gtrsim 1$ AU. Flare-related non-relativistic electron events observed near the Earth are impulsive in nature, as shown in figure 12, with a rapid rise and gradual decay. The time profile depends on the amount of scattering in propagating to 1 AU. If it is small as in figure 12(*a*), the event will show a sharp, rapid rise and rapid initial decay (both $\sim$minutes at $\gtrsim 40$ keV) followed by a long decay ($\sim$ hours). If the amount of scattering is large then a diffusive profile as in figure 12(*b*) will be observed with rise times from a few minutes to a few hours and decay times about an order of magnitude longer. If the event is seen at several different energies, the faster electrons are observed to arrive prior to the slower electrons.

The energy spectra of non-relativistic electron events at the Sun (injection) can be determined only with a model for electron propagation in the interplanetary medium. Unfortunately there is a lack of measurements of the interplanetary magnetic-field structure on the scale of non-relativistic electron gyroradii (~ 50–500 km in a 5×10^{-5} G field). Thus theories relating the particle propagation to the interplanetary field fluctuations (reviewed by Jokipii (1971)) cannot be confirmed for these electron energies. The observed characteristics of non-relativistic electron events are consistent with a highly variable amount of scattering in the interplanetary medium with little, if any, propagation across interplanetary field lines. Thus the transverse diffusion coefficient $\kappa_\perp \ll \kappa_\parallel$, so that $\kappa_\perp$ can be neglected relative to $\kappa_\parallel$ (the parallel diffusion coefficient) hereafter called κ. Non-relativistic electrons measured near the Earth are emitted over a range $\sim 60^0$ (full width at half-maximum) of solar longitude which is concentrated to the west, i.e. to those field lines which map directly from the Sun to the Earth. According to the simple scattering model of Fisk and Axford (1969), reviewed by Lin (1974a), for diffusive propagation the maximum intensity is independent of the value of the diffusion coefficient κ. Thus, even if electrons of different energies have different values of κ, the maximum flux is directly related to the number of electrons emitted by a constant independent of κ and the energy spectrum constructed from this flux in each energy channel is representative of the injection spectrum. This is true provided the spatial dependence of the propagation does not vary substantially with energy which should be a good approximation for diffusive events. A typical spectrum constructed in this way is shown in figure 13. For scatter-free events there are too many uncertain factors to relate the measured spectrum to the injection spectrum.

The properties of injection electron spectra constructed in the above manner are the following.

(i) The differential spectra are power laws in the range ~ 5–100 keV of the form $dn/dE \propto E^{-\delta}$, where E is the energy and n is the number of electrons per cm³. Values of δ range from 2–5 with most of the events falling between 3 and 4·5.

(ii) In the case of electrons unaccompanied by protons, called pure electron events, there is a rapid steepening of the spectrum at energies above ~ 100 keV as shown in figure 18. In the case of electrons accompaned by protons, called mixed events, the electron spectra extend in a power law to relativistic energies.

(iii) The spectrum below ~ 5 keV departs from a power law, tending to roll off at low energies.

(iv) The number of electrons injected into the interplanetary medium may be estimated from the maximum intensity observed at each energy which yields $\sim 10^{33}$
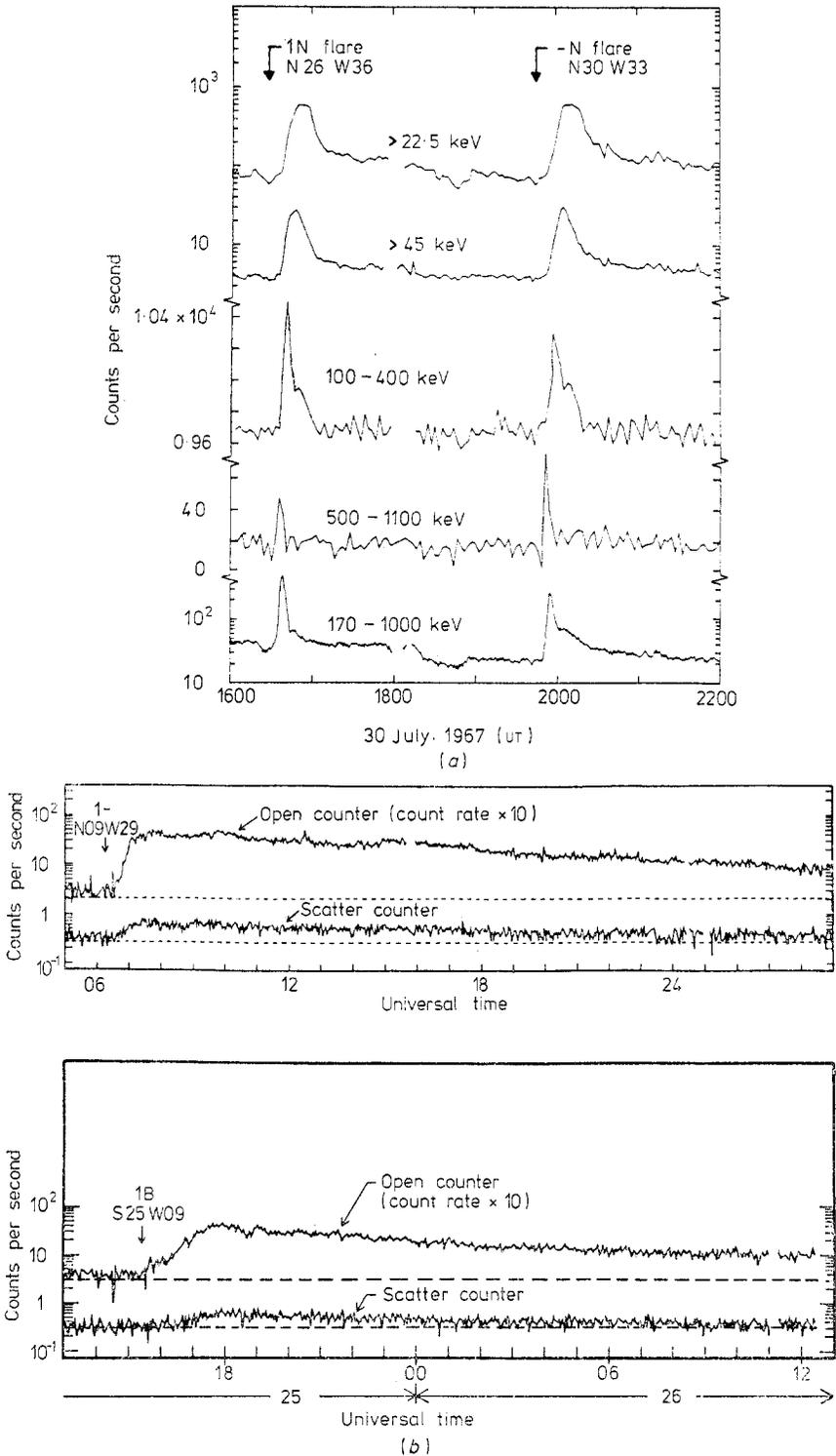
Figure 12. (*a*) Two scatter-free impulsive solar electron events. The sharp initial peak in the high-energy channels place an upper limit on the duration of the electron injection into the interplanetary medium of $\leqslant 3$ min. (*b*) Two diffusive electron events. The scatter counter is only sensitive to > 45 keV electrons while the open counter counts both > 40 keV electrons and > 0.5 MeV protons (after Lin 1974a). Upper part: 27 December 1965. Lower part: 25–26 June 1966.

Figure 13. The electron energy spectra of the 30 July 1967 events of figure 12(*a*) illustrating the steepening in the spectrum above ~ 100–200 keV typical of pure electron events (after Lin 1974a).

electrons above ~ 20 keV for a pure electron event and as high as $\sim 10^{36}$ electrons above ~ 20 keV in a large mixed event. According to the simple propagation theory reviewed by Lin (1974a) only 0.1–1% of the electrons accelerated escap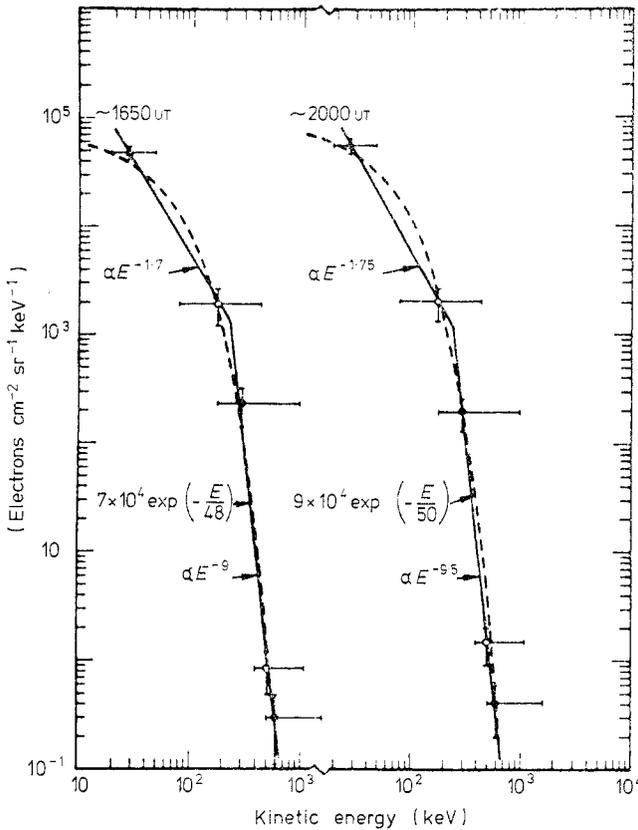e to the interplanetary medium. Thus 10^{35}–10^{36} electrons above ~ 20 keV must be accelerated at the Sun in a small flare with a power-law spectrum.

The density of the region in which the escaping electrons are accelerated has been estimated by noting that none of the low-energy electron spectra observed shows a turnover above ~ 6 keV and assuming that energy losses are collision-dominated (Lin 1974b). Using the expression of Trubnikov (1965) for the energy loss in ionised hydrogen leads to a spatial loss rate

$$\mathrm{d}E/\mathrm{d}x = -2.6 \times 10^{-18}\, n_\mathrm{i}/E \qquad \text{keV cm}^{-1} \qquad (2.3)$$

where n_i is the ion density. Integrating from height h to 1 AU

$$\int_{E_1}^{E_2} E\,\mathrm{d}E = -2.6 \times 10^{-18} \int_h^{1\,\mathrm{AU}} n_\mathrm{i}(x)\,\mathrm{d}x = K(h) \qquad (2.4)$$

and

$$E_2 = (E_1{}^2 - 2K)^{1/2} \qquad (2.5)$$

where E_1 is the initial accelerated energy of the electron and E_2 is its energy at 1 AU. If the spectrum of the accelerated electrons is

$$\mathrm{d}n_1/\mathrm{d}E_1 = AE_1^{-\delta} \qquad (2.6)$$

then the spectrum observed at 1 AU will be

$$\mathrm{d}n_2/\mathrm{d}E_2 = AE_2/(E_2^2 + 2K)^{(\delta+1)/2} \qquad (2.7)$$

which has a maximum at

$$E_{2M} = (2K/\delta)^{1/2}. \qquad (2.8)$$

The observed lack of a turnover above ~ 6 keV implies

$$\int_n^{1\,\mathrm{AU}} n_1(x)\mathrm{d}x \lesssim 3 \cdot 5 \times 10^{19}\ \mathrm{cm}^{-2} \qquad (2.9)$$

which corresponds to $\lesssim 60\ \mu\mathrm{g\ cm}^{-2}$ of hydrogen, equivalent to an ambient density of less than $\sim 10^{10}\ \mathrm{cm}^{-3}$ at the *effective* acceleration region. It is possible, however, that the electrons are accelerated at higher densities and have their energy redistributed at lower densities. The *effective* acceleration region is the redistribution region in this case.

2.5.3. Relativistic electrons. In the case of mixed events, the electron spectrum does not steepen above ~ 100 keV, but extends smoothly up to relativistic energies as shown in figure 14. The properties of the relativistic electrons have recently been reviewed by
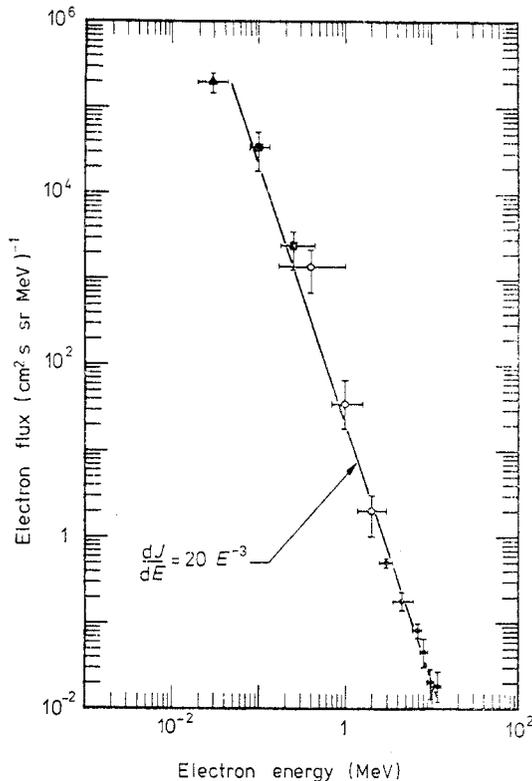

Figure 14. The differential electron energy spectrum for the mixed electron–proton event of 28 May 1967 compiled from four different detector systems aboard IMP-4 showing its uniform nature (after Lin 1974a). ▲, University of California; ○, University of Chicago; ●, Goddard Space Flight Center.

Simnett (1974) including so-called co-rotating events which are infrequent electron events of duration greater than one week. Since the co-rotating events are not directly flare-related, we shall limit our discussion to impulsive events as with the non-relativistic electrons. As compared to the non-relativistic electron events which occur for about one flare out of 40, mixed events only occur for about one flare in 230. Energy spectra for relativistic events are constructed in the same manner as for non-relativistic diffusive events. Since essentially all relativistic electron events are diffusive with a rise time typically of ~ 1–4 h, and a decay time to background of ~ 2–3 d
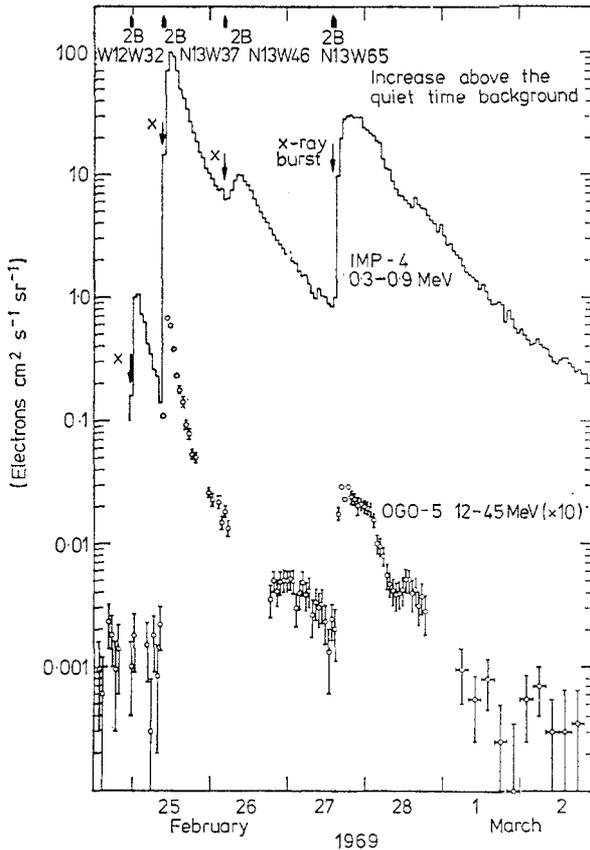
Figure 15. A series of impulsive electron events from flares in MP 9946 which was at west solar longitude during the period of interest (after Simnett 1974).

as shown in figure 15, this should be a good approximation. The spectra are power laws of the form $dn/dE \propto E^{-\delta}$ with $\delta \sim 2$–$4\cdot 5$ with steepening toward higher energies in some events.

Simnett (1974) gives no information on the number of relativistic electrons required for an impulsive event, but does quote Holt and Ramaty (1969) in estimating the escape efficiency as 1%. From spectra such as figure 14 we can deduce that the number of relativistic electrons with $E > 200$ keV is about 10^2 smaller than the number above ~ 20 keV so that as many as $\sim 10^{34}$ electrons above ~ 200 keV are produced in a large mixed event which arrive later and last longer than non-relativistic electrons.

2.5.4. Protons. Like relativistic electrons, protons are a much rarer phenomenon than non-relativistic electrons. These events are always diffusive in nature which justifies the use of the simple scattering theory to deduce spectra. However, their spectra are more consistent with exponentials in rigidity R than power laws in energy as can be seen in figure 16, where the rigidity

$$R = pc/Ze \qquad (2.10)$$

(here p and Z are the momentum and atomic number of an ion, respectively). Over limited ranges power laws in energy also fit the observations. Like relativistic electrons, non-relativistic protons arrive later and last longer than non-relativistic electrons.
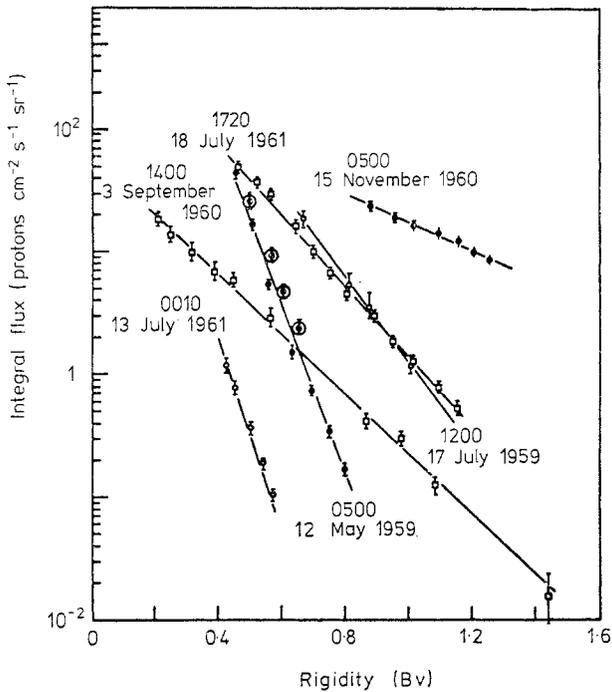
Figure 16. Exponential rigidity spectra for solar cosmic ray events at the universal times indicated (after Freier and Webber 1963).

The number and energy of escaping protons above 0.1 MeV for the 28 September 1961 event has been calculated by Biswas and Radhakrishnan (1973). They found $\sim 2 \times 10^{33}$ protons escaped which had an energy of $\sim 4 \times 10^{27}$ erg. Only 0.01% of the protons injected escape so that $\sim 1.5 \times 10^{37}$ protons must be injected with an energy of $\sim 3 \times 10^{30}$ erg. These numbers are quite sensitive to the spectral form at injections, but should be good to within an order of magnitude for this class 3 flare. The largest events release about 10^{34} protons into interplanetary space and involve an energy release of more than 10^{31} erg. At 10 MeV there are typically between 10 and 100 times more protons than relativistic electrons.

The relativistic BeV proton fluxes from large solar flares ('solar cosmic rays') present a severe constraint on the acceleration mechanism since the energy per particle is so large. These protons are measurable on the ground by neutron monitors and thus their existence was known long before the advent of satellites.

2.5.5. Other nuclei. Solar cosmic rays are also rich in other nuclei, often with strange and highly variable relative abundances. While they may provide valuable clues to the acceleration process, they are not really important energetically because even He is at most about 10% as abundant as H and often less so (Lanzerotti 1973). Because most of the theories for abundances and preferential acceleration of heavy ions seem rather contrived (e.g. Cartwright and Mogro-Campero 1973), we do not feel that the data give any unambiguous clues which justify inclusion here, though interesting in themselves (see, for example, Crawford *et al* 1975, Colgate *et al* 1978). The theory of Colgate *et al* (1978) involves nuclear reactions which destroy light elements but leave heavy nuclei intact. This behaviour seems common to most nuclear reactions which can occur under even extreme flare conditions so that the need for preferential acceleration of heavy nuclei is unchanged.

In short, non-relativistic electrons and protons of all energies provide the most stringent requirements on the flare acceleration mechanisms in terms of energy. Other nuclei possibly provide more subtle and difficult constraints on the acceleration mechanisms, but there is no certainty at the present time that some of these difficulties could not be due to propagation effects.

3. Theory of primary energy release mechanisms

3.1. Magnetic-field configurations

As noted in §2.2 there are no direct measurements of magnetic fields in the low corona where the bulk of the flare energy is believed to be released in solar flares. The flare energy, however, *must* be derived from stored magnetic energy as may be seen by comparing the other forms of energy available with the average energy released per unit volume in a flare W_F. In a large flare the total energy released is about 10^{32} erg, while the maximum possible volume in which this energy release occurs is 10^{29} cm³. Thus $W_F \geqslant 10^3$ erg cm⁻³. The other forms of energy in the flare plasma are thermal and gravitational. With reasonable upper limits to the mean density n and temperature T in the pre-flare plasma, $n = 3 \times 10^{11}$ cm⁻³ and $T = 10^6$ K (Dere *et al* 1977) at a height $h \leqslant 4000$ km, the thermal energy density is

$$W_T = 3nKT = 126 \text{ erg cm}^{-3} \qquad (3.1)$$

and the gravitational energy density

$$W_G = n \, m_i \, g_\odot h = 5 \cdot 5 \text{ erg cm}^{-3} \qquad (3.2)$$

where m_i is the proton mass. (This is in an arch. In an open structure hydrostatic equilibrium demands $W_G \simeq W_T$.) Thus, thermal and gravitational energy combined fail by more than an order of magnitude to power a flare. The flare energy density $W_F \geqslant 10^3$ erg cm⁻³, on the other hand, corresponds to a magnetic field $B > 160$ G which is quite compatible with any reasonable extrapolation of the measured photospheric magnetic field to the low corona.

With the lack of direct observations and the demonstrated importance of magnetic fields for solar flares, theoreticians have resorted to the so-called 'cartoon approximation' in which a given overall magnetic-field configuration is empirically adopted with more or less loose observational justification (cf de Jager 1968). The first type of configuration shown in figure 17(*a*) is the large-scale current sheet put forward by Sweet (1958) and Parker (1963) and developed by Sturrock (1966, 1968, 1973) largely on the observational grounds that the high-energy particles and ejected plasma from a
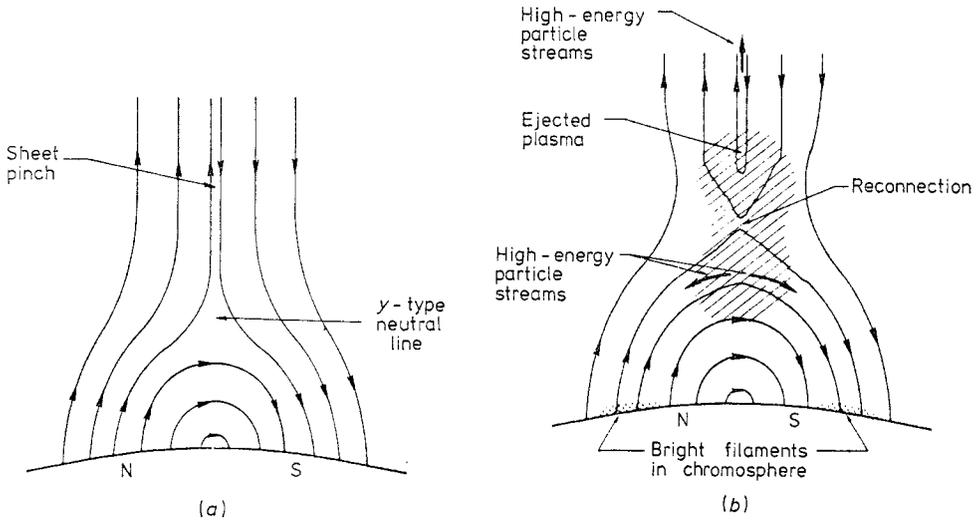
Figure 17. (*a*) Schematic representation of the magnetic-field pattern above a bipolar magnetic region, showing the transition from closed field lines to open field lines with a sheet pinch or current sheet (after Sturrock 1968). (*b*) Schematic representation of reconnection of magnetic-field lines during a flare showing ejection of plasma and heating of chromosphere by particle streams to form two bright filaments (after Sturrock 1968).

flare escape into the interplanetary medium. They can only escape along open magnetic-field lines, such as occur naturally in a reconnected current sheet as shown in figure 17(*b*) (but compare with below). This model has been developed by Heyvaerts *et al* (1977). The second type of configuration is an arched flux tube carrying a current as shown in figure 18. This was first put forward by Gold and Hoyle (1960) and Alfven and Carlquist (1967), but only developed on realistic grounds by Spicer (1977). The firmest observational basis for this geometry lies in the soft x-ray pictures from Skylab as reviewed by Spicer and shown in figure 19 (cf figure 7). The Skylab observations strongly suggest (cf §2.2) that flares appear to occur in either a single arch or a series of arches, often called an arcade, and that the energy release is near the top of the arch or arches. In the case of a single arch the observations are quite incompatible with the current sheet model because the disc view (early phase) drawing of figure 19 would require a sheet of very small extent in the direction out of the paper in figure 17(*a*) and there is nothing to contain the energy flowing down in this direction so that one would
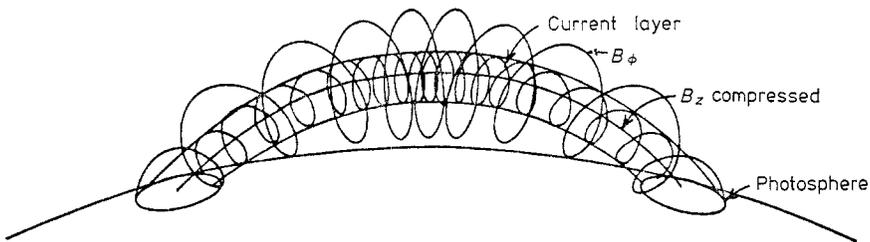
Figure 18. An initially twisted flux tube or current carrying filament with the induced B_ϕ component compressing the B_z field (after Colgate 1978).

not expect to see the two kernels. In the case of a classic two-ribbon flare in Hα either a current sheet or an arcade of loops is compatible with the observations in this respect.

However, there are theoretical considerations which definitely favour the arch or arcade configuration. We shall see in §3.2 that the process involved in both configurations is the same, namely magnetic-field reconnection in steady-state (Vasyliunas 1975) and dynamic forms (Furth *et al* 1963, Drake and Lee 1977) which converts magnetic energy into thermal and particle energy in extremely small volumes. In order to obtain a significant energy conversion rate the number of these volumes should be very large. Even with an anomalous resistivity due to microinstabilities (cf §3.2) a current sheet is quite thin for low corona parameters. The half-thickness of the diffusion region of the current sheet where energy dissipation occurs is (Vasyliunas 1975)

$$z^* \sim \eta c^2 / 4\pi v \tag{3.3}$$

in the case that it is determined by the resistivity η. Here v is the velocity of field lines flowing into the sheet which is of the order of 10^7 cm s^{-1} for reconnection rates under low corona conditions. The resistivity due to ion-acoustic turbulence is (Smith and Priest 1972, cf Spicer 1977)

$$\eta = 6 \cdot 71 \times 10^{-7} \, n^{-1/2}. \tag{3.4}$$

For $n = 3 \times 10^{11}$ cm^{-3}, $\eta = 1 \cdot 2 \times 10^{-12}$ s and $z^* = 8 \cdot 7$ cm with the above value of v. The maximum area of the sheet consistent with observations is about 10^{18} cm^2 or 10 000 km on a side. If the whole sheet were filled with diffusion regions (which is extremely unlikely), the total volume would be $1 \cdot 7 \times 10^{19}$ cm^3. One could argue that the assumed density is too high for the corona, but the density in a current sheet is higher than in the surrounding plasma because the plasma pressure at the centre of the sheet must balance the magnetic pressure of the reversing magnetic field for an initial equilibrium
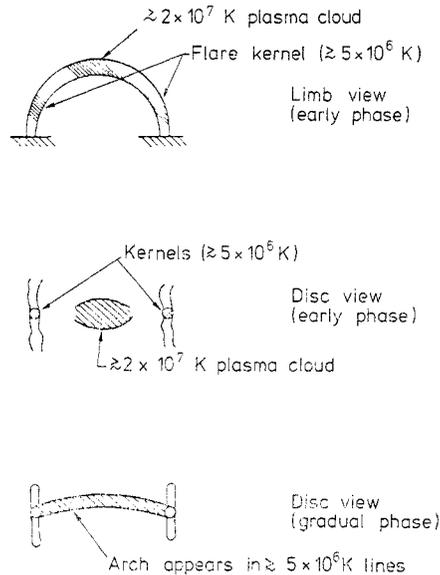

Figure 19. Schematic of flare geometry as observed in soft x-rays (after Spicer 1977).

to exist, i.e. for a fully ionised plasma

$$B^2/8\pi = 2\,n\,KT. \tag{3.5}$$

With the observed orders of magnitude for n (3×10^{11} cm^{-3}) and T ($\simeq 10^6$ K) $B \simeq 50$ G. Hence either a higher n or higher T or both are required to create a current sheet with a larger reversing field. The typical time scale for a small flare is 100 s and the typical energy release is 10^{29} erg which requires an energy release rate of 10^{27} erg s^{-1}. It is shown in §3.3 that the largest energy release rate which can be obtained in a sheet of this size is $\sim 10^{25}$ erg s^{-1} unless mass motions of ~ 1000 km s^{-1} exist, which are unobserved in most small flares. Hence current sheet models have difficulty explaining a small flare. It does not help that magnetic field can be brought in from a much larger region because it must still be dissipated in a very small region and there are limits to the rate at which this can realistically be done (§3.3).

A magnetic arch or arcade, on the other hand, can easily have a fairly large dissipation volume since the ratio of its volume to surface area is large. It can be argued that the volume/area ratio for a current sheet is misleading because a much larger volume of magnetic field can be processed by the sheet. The essential difference, however, is that in the current sheet, the magnetic field can only be transformed to other forms of energy in the sheet or, if we include the shock waves in a Petschek type of configuration as in figure 20, in a thin layer. It will be shown in §3.3 that the factor which throttles the rate of energy release in a current sheet is the rate at which material can be removed from the sheet. In §3.2 we show that in a magnetic arch magnetic energy can be transformed into other forms at every point in the arch. Thus, even though the rate of energy release is throttled in the same manner in an arch, it does not lead to the same limitation as in a sheet because different regions in the arch can become simultaneously unstable. (This is equivalent to saying that a complex sheared field geometry allows a large effective area of reconnecting surface to be convoluted inside a fixed total volume. The increased dissipation rate of the sheared field model can then in part be regarded as analogous to accelerating a chemical reaction by fine crushing of a solid reagent.) This is its advantage and the reason why the typical observed volume of an arch of 10^{25} cm^3 can result in a sufficiently fast energy release for a small flare. Small regions of the surface of the arch can be blown open by kink instabilities (§3.2) to allow some escape of particles as required by the observations.

Thus, while there are no direct observations of the magnetic field in solar flares, there are clear theoretical guidelines for selecting a hypothetical magnetic geometry. A magnetic arch or arcade is the most likely on these grounds and we shall adopt it as a working model in the following subsections.

3.2. Plasma instabilities and the trigger mechanism

With some idea of the relevant magnetic-field configuration, we can analyse relevant plasma instabilities for converting magnetic energy into the forms required for a flare. Since a flare can be triggered as evidenced by both impulsive beginning phases and by shock waves from one flare giving rise to another, some subsidiary instabilities which can trigger the primary instability will also be considered. The main energy conversion mechanism is the tearing mode instability (Furth *et al* 1963, Drake and Lee 1977) which is a dynamic form of magnetic-field reconnection (Vasyliunas 1975). Because it will be helpful in analysing how the tearing mode instability is throttled in its non-linear state in §3.3, we shall begin by considering steady-state or static

11

reconnection. The basic configuration is shown in figure 20 (plate). Oppositely directed magnetic fields are brought into a small region where the field lines and fluid are de-coupled, called the diffusion region with a velocity v. They are reconnected in this region, with the conversion of a large fraction of the magnetic energy into kinetic and thermal energy, and flow out with the velocity v_x which is taken as the Alfven velocity $v_A = B (4\pi n m_i)^{-1/2}$. The rate of reconnection is then determined by the ratio

$$v/v_x = v/v_A = M_A \tag{3.6}$$

called the Alfven Mach number. In the case that the resistivity is anomalous, the half-thickness of the diffusion region is given by equation (3.3) and the half-length of the diffusion region

$$x^* = z^*/M_A = \eta c^2/4\pi v_x. \tag{3.7}$$

In the model of Petschek (1964) for the non-linear stage of this steady-state recon-nection there are slow mode shock waves emanating from each corner of the diffusion region which have recently been discussed by Soward and Priest (1977). Thus far these shock waves have been considered as infinitely thin layers whose purpose is to turn and accelerate the flow, and heat the plasma, as will be discussed in §3.3.

Moving from steady to dynamic reconnection, we come to the model of Syrovatskii (1966, 1969, 1972), in which the dynamic evolution of a potential magnetic field con-taining a neutral point is followed. In this model the region of dissipation of magnetic energy can be large but, as discussed by Kaplan *et al* (1974), in this case even motion at the velocity of light cannot take away the total gas mass sufficiently rapidly to explain a flare (see also §3.3). This brings us to the tearing mode instability (Furth *et al* 1963, Rutherford 1973, van Hoven and Cross 1973, Schnack and Killeen 1978). Here, instead of pushing oppositely directed magnetic fields together we ask how a system with initially oppositely directed magnetic fields forming a current layer will evolve as shown in figure 21. The magnetohydrodynamic (MHD) equations, with which we can
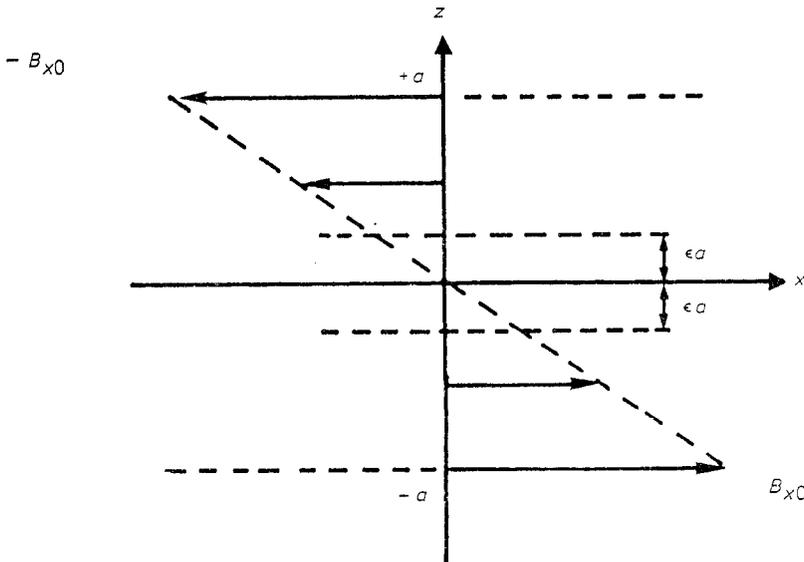

Figure 21. The sheared magnetic-field equilibrium at the beginning of the tearing mode instability (after Spicer 1976).
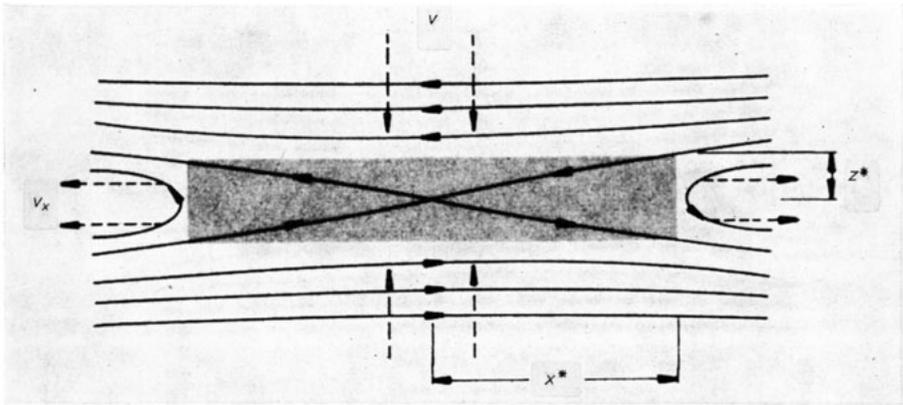
Figure 20. The Petschek model of reconnection with the diffusion region shaded. The magnetic-field lines are full and the flow lines are broken. Fluid enters with velocity v and flows out with velocity v_x (after Vasyliunas 1975).

trace the evolution of this system, are (Furth *et al* 1963)

$$E + \frac{1}{c}\,v \times B = \eta j \tag{3.8}$$

$$\frac{\partial B}{\partial t} = \nabla \times (v \times B) + \frac{c^2}{4\pi}\,\eta \nabla^2 B \tag{3.9}$$

$$\nabla \times \left(\rho\,\frac{\mathrm{d}v}{\mathrm{d}t}\right) = \nabla \times \left(\frac{1}{4\pi}\,(\nabla \times B) \times B\right) \tag{3.10}$$

and

$$\nabla \cdot v = 0 \tag{3.11}$$

where E is now the vector electric field, $\rho = nm_1$ is the mass density and η is a constant resistivity. The tearing mode instability exists only for finite η. It can be seen from equation (3.8) that the effect of η becomes important at the neutral layer at $z \sim 0$ where the magnetic field $B_x \sim 0$. On the other hand, at distances far from the neutral layer the $v \times B$ term can dominate, and the plasma can be regarded as lossless. Thus, to understand the physical process of the instability we divide the current layer into two regions, with resistivity at $|z| < \epsilon$, and that with no resistivity at $a > |z| > \epsilon$. As shown by Hasegawa (1975) the growth rate for the instability can be found by solving for the perturbed magnetic field B_{1z} in these two regions and matching the solutions at $z = \pm\epsilon$.

This procedure shows that the driving force of the instability is the non-uniform magnetic field with $B_x/B_x'' < 0$, where B_x'' is the second derivative of B_x with respect to z.

The instability occurs for a wavelength in the x direction longer than the thickness, i.e. $k < 1/\lambda \sim a^{-1}$. As a consequence of the instability, x-type neutral points tear the sheet current into a number of smaller segments as shown in figure 22. In other words, the tearing mode instability filaments the sheet current into a number of current strands. A more complete analysis including non-constant resistivity (Spicer 1976) shows that the places where the sheet tears, or the x-type neutral points form, are those where $k.B = 0$. The reason why these are the most unstable points can be understood by noting that we should have included a term $i\,(k.B)v$ on the right-hand side of equation (3.9) which is an additional damping term to the diffusion. The vanishing
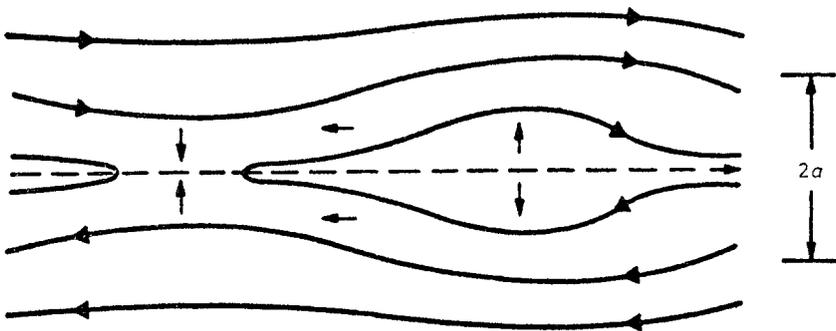
Figure 22. The tearing mode instability of a current sheet with the pattern of fluid flow indicated by the small arrows (after Hasegawa 1975).

points of $k.B$ correspond to those for which this damping vanishes and hence are most unstable.

Unfortunately the single tearing mode in sheet geometry saturates with uninteresting energy release on the time scales required for a flare (Rutherford 1973). However, a slight variation of this mode, called the double tearing mode, does result in energy release of interest on the required time scale by means of two closely adjacent tearing nodes driving each other (Schnack and Killeen 1978).

We turn now to the more realistic geometry of a current carrying arch with a longitudinal B_z magnetic field and a transverse B_ϕ field as shown in figure 18, which was considered in detail by Spicer (1976, 1977). As reviewed in Spicer (1976), to be initially stable, an arch, with uniform current density and a length of ten times its radius r_0, must have a longitudinal field which is a minimum of 1·6 times the surface transverse field produced by the current. This result is a particular case of the Kruskal–Shafranov stability condition (Kruskal and Tuck 1968, Shafranov 1956) which states that $2\pi r_0 B_z (B_{\phi_0} L)^{-1}$ must exceed unity for stability where B_{ϕ_0} is the azimuthal field B_ϕ at $r = r_0$. Thus the magnetic field is helical and most unstable at those points where $k.B = 0$ which we shall call singular points. Let m be the azimuthal mode number and ϕ the azimuthal angle. The singular points may be characterised as those values of r for which $m\phi + kz$, the phase of the pertubation, is constant along a line of force, i.e. the pitch of the lines of force exactly matches the pitch of the perturbation. Here the cylindrical tearing mode (Furth *et al* 1973) filaments a current surface into current ribbons as shown in figure 23, converting magnetic energy into thermal, kinetic and particle energy. For a given (k, m) a helical current layer will exist at the singular surface with a null $k.B$, permitting the layer to tear as in the sheet tearing mode. Many singular layers, corresponding to any pair (k, m), are possible, and these layers are closely spaced in the arch. The high m and large k modes have the largest growth rate, but the small m and k modes involve the largest displacements of the magnetic-field lines. Thus, with the great number of possibilites, the arch can
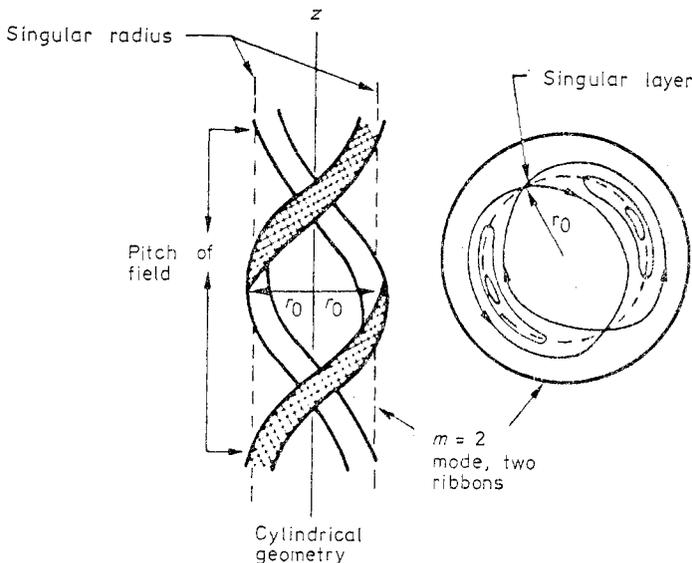

Figure 23. An $m = 2$ mode perturbation of a current filament showing how the filament can tear into strands (after Spicer 1976).

simultaneously and successively tear at many different r, ϕ and z values. As yet there has been no solution of the very complicated numerical problem required to study this phenomenon in detail for solar conditions. As noted by Spicer (1976) results for cylindrical tearing modes are strongly dependent on the boundary conditions, but growth rates tend always to be higher than for the current sheet.

One distinction between the tearing mode in sheet geometry and the cylindrical tearing mode which dissipates the B_ϕ field or the current in an arch should be kept in mind. Because the fields are reversed *ab initio* in the current sheet, whereas the field reversal must develop from a perturbation in the current carrying arch, the parameters of tearing modes are much more sensitive to the initial configuration in the arch than in the current sheet. However, instabilities in Tokamaks which have been directly traced to cylindrical tearing modes leave no doubt that such perturbations can easily occur (e.g. Hutchinson 1976).

What can facilitate the production of such perturbations and act as a trigger for the instability? The instability is resistive in nature and anything which tends to increase the amount of dissipation at one point will enhance the instability. Thus instabilities which rapidly increase the resistivity and/or steepen the current profile locally are the most likely candidates. The value of steepening the current profile locally can be understood from an analysis of the tearing mode in sheet geometry (Hasegawa 1975) which shows that the tearing mode is driven by the non-uniform magnetic field with $B_x/B_x'' < 0$. An instability which rapidly increases the resistivity locally is the ion-acoustic instability (Smith and Priest 1972). Computer simulations have shown that a magnetic field with $\omega_{ce}^2 \ll \omega_{pe}^2$, where $\omega_{ce} = eB/m_e c$ is the electron cyclotron frequency, has little influence on this instability (Biskamp and Chodura 1973). Thus we can use the field-free results which have recently been considered in the context of steady-state reconnection by Coroniti and Eviatar (1977) and in the context of dynamic reconnection by Smith (1977a). The current in a tearing current layer as in figure 21 can be expressed in terms of a drift velocity v_D which is related to the magnetic-field gradient as

$$j = ne\, v_D = cB/4\pi a. \tag{3.12}$$

The instability condition for the ion-acoustic instability is (Stringer 1964)

$$v_D > c_s + (c_s - v_i) \left(\frac{m_i}{m_e}\right)^{1/2} \left(\frac{T_e}{T_i}\right)^{3/2} \exp\left[-\left(\frac{T_e/2T_i + \tfrac{3}{2}}{1 + k_s^2\, \lambda_{De}^2}\right)\right] \tag{3.13}$$

where

$$c_s = [K(T_e + 3T_i)/m_i]^{1/2} \quad \text{and} \quad v_i = (KT_i/m_i)^{1/2} \tag{3.14}$$

are the sound speed and ion thermal velocity, respectively, k_s is the wavenumber of the unstable ion-acoustic wave and $\lambda_{De} = v_e/\omega_{pe}$ is the electron Debye length. Here $v_e = (KT_e/m_e)^{1/2}$ is the electron thermal velocity and $k_s = 0.5\, \lambda_{De}^{-1}$ for the fastest growing mode (Stringer 1964). In simple terms when $T_e/T_i = 1$, $v_D > v_e$ is required for instability and when $T_e/T_i \gg 1$, $v_D > c_s$ is required.

The instability criterion (3.13) places fairly severe requirements on the magnetic-field gradient B/a through equation (3.12). However, these requirements can easily be met for short wavelength or large k tearing modes close to the tearing points where $\mathbf{k}.\mathbf{B} = 0$. The result of the ion-acoustic instability is to produce waves which saturate due to the trapping of ions in the waves (Biskamp and Chodura 1973) which leads to an anomalously large resistivity (Smith and Priest 1972)

$$\eta_A = (1.6/\omega_{pe})\, (m_e/m_i)^{1/2} = 6.71 \times 10^{-7}\, n_e^{-1/2}. \tag{3.15}$$

The ratio of this resistivity to the classical resistivity η_c due to Coulomb collisions (Spitzer 1962) is

$$\eta_A/\eta_c = 3 \cdot 21 \; T_e^{3/2} \; n_e^{-1/2}. \tag{3.16}$$

Suppose there is some electron heating in establishing the gradients necessary for the ion-acoustic instability so that $T_e \sim 10^7$ K which is consistent with pre-flare brightenings (see §2.1). For a density of 3×10^{11} cm^{-3}, $\eta_A/\eta_c = 1 \cdot 8 \times 10^5$ which increases the growth rate of the double tearing mode by a factor ~ 200. The time for development of the ion-acoustic instability is of the order of 2000 ω_{pi}^{-1} or $1 \cdot 7 \times 10^{-6}$ s for the above density. Such a large change in so short a time could trigger a flare.

An instability which steepens the current profile is the superheating instability discussed by Spicer (1976). It is a form of thermal instability and works as follows. We have a current j_z flowing along the B_z field in figure 18. If the temperature of the current carrying plasma is perturbed on a scale for which thermal conductivity can be neglected, then the current density will grow when the Joule heating exceeds the radiative power loss. Since the classical resistivity varies as $T_e^{-3/2}$ (cf equation (3.16)) and j_z is inversely proportional to η_c, a temperature increase will result in a decreased resistivity, an increased current density and thus an increased Joule heating. This in turn leads to a further decrease in the resistivity, a further increase in the current density, etc, so that the current grows at the expense of the external perturbing thermal input. We shall not derive this instability which was first noted by Kadomtsev (1966) and as a thermal instability by Coppi and Friedland (1971), but refer the interested reader to Spicer (1976) for details. The criterion for instability is

$$\frac{j_0^2}{T_0\sigma_0} \frac{\mathrm{d} \ln \sigma_0}{\mathrm{d} \ln T_0} > \frac{\mathrm{d} Q}{\mathrm{d} T_0} \tag{3.17}$$

where subscript 0 refers to initial values, the electrical conductivity $\sigma = \eta^{-1}$, and Q is the heat flux.

There are several infinite conductivity MHD instabilities such as the kink instability which may also play a role in flares. However, these instabilities do not really represent a *dissipation* of magnetic energy, but a conversion of magnetic energy to kinetic energy. Thus their role could also be a triggering one, but they cannot serve as the primary energy release process.

3.3. Dissipation mechanisms and the rate of energy release

As alluded to repeatedly in §§3.1 and 3.2 many of our ideas about magnetic-field configurations and plasma instabilities are conditioned by the dissipation mechanisms to which they lead. With a current sheet the rate of energy release in a sheet of surface area S is

$$\frac{\mathrm{d} W}{\mathrm{d} t} = \frac{B^2}{8\pi} S(2 \, v_\mathrm{I}) \tag{3.18}$$

where v_I is the velocity of the incoming fluid. It is often assumed that v_I can be set equal to $\sim 0 \cdot 1 v_\mathrm{A}$ which is reasonably consistent with the results of Soward and Priest (1977). However, the results of Soward and Priest apply to a Petschek type reconnection configuration as shown in figure 20. The region of actual dissipation of magnetic energy is very small with a half-thickness z^* given by equation (3.3) and a half-length x^* given by equation (3.7). For $n = 3 \times 10^{11}$ cm^{-3} and a 500 G reconnecting

field, $v_A = 2 \times 10^8$ cm s^{-1} and for $v_I = v = 10^7$ cm s^{-1}, $x^* = 174$ cm with $\eta = 1 \cdot 2 \times 10^{-12}$ s found in §3.1 for the anomalous resistivity. There can only be one such dissipation region in the whole sheet so the maximum value which can be used for S in equation (3.18) is $2\,x^*L = 3 \cdot 48 \times 10^{11}$ cm^2, where $L = 10^9$ cm is the depth of the sheet. The rate of energy release for the above values of B and v_I is $6 \cdot 9 \times 10^{22}$ erg s^{-1} which is far too small to explain even a small flare. It may be argued that the slow shocks which extend from each corner of the diffusion region in figure 20 enhance the energy release (over $S = L^2$). They do result in the conversion of magnetic energy, about half of which goes to the kinetic energy of fluid motion and about half of which goes into heat (Kantrowitz and Petschek 1966). However, since most of the magnetic energy conversion now occurs in the slow shocks, half of the energy released goes into mass motion at a velocity of 2000 km s^{-1} over an area of 5×10^{16} cm^2 for the above values which would certainly be observable. In fact, there are many small flares with an energy release of 10^{29}–10^{30} erg where essentially no mass motion greater than ~ 20 km s^{-1} is observed.

What happens if we try to slow down the mass motion to, say, 20 km s^{-1} by impeding the outflow, for example, to keep velocities consistent with observation? The angle which the shocks make with the x direction in figure 20 would decrease to almost zero and the following constraint on the inflow velocity would result:

$$v_I = \frac{2z^*}{L}\, v_F \tag{3.19}$$

where v_F is the outflow velocity. With $v_F = 20$ km s^{-1} and $z^* = 8 \cdot 7$ cm from equation (3·3), $v_I = 3 \cdot 5 \times 10^{-2}$ cm s^{-1} which leads to a rate of energy release of $6 \cdot 9 \times 10^{20}$ erg s^{-1}, again far too small for a flare. For large flares with an energy release of $\sim 10^{32}$ erg, half of which is observed to be in mass motions, this argument cannot be used to eliminate the Petschek mechanism. However, as shown by Steinolfson *et al* (1977) these mass motions can also result simply by heating up the low corona and including material boiled off the chromosphere. Alternatively we could consider the dynamic dissipation mechanism of Syrovatskii (1966) (see §3.2) in which the whole sheet is filled with a dissipation region and there are no slow shocks. However, equation (3.19) can be used in this case to determine the maximum possible inflow velocity with $v_F = c$ since it cannot exceed the speed of light. The area of outflow, $2\,z^*L = 1 \cdot 7 \times 10^{10}$ cm^2, is now so small that it would be unobservable so that the argument used for the Petschek configuration does not apply. With the above volume of z^* and L and $v_F = c$, $v_I = 5 \cdot 3 \times 10^2$ cm s^{-1} which leads to a rate of energy release from equation (3.18) of $1 \cdot 1 \times 10^{25}$ erg s^{-1} which is still two orders of magnitude too small to explain a small flare. Pustil'nik (1975) tried to rescue the current sheet by noting that currents will be large on the edges of the sheet which will tend to expand the region of anomalous resistivity and thus thicken the sheet. However, as pointed out by Pikel'ner and Tsytovich (1976), the current in the middle of the sheet will become so low that the ion-acoustic instability will be turned off and the sheet will collapse again to a half-thickness z^* of the order of that given by equation (3.3) so that this mechanism does not help.

We next consider what throttles the rate of energy release in tearing mode instabilities. The non-linear analysis by van Hoven and Cross (1973) showed that this throttling takes about $5\gamma^{-1}$ where γ is the growth rate (Schnack and Killeen 1978) and that the back pressure of flow in nearby tearing modes slows the energy conversion. Essentially all of the reversing magnetic field is converted into thermal and kinetic energy during this time. Setting $a = 2z^*$ with $z^* = 8 \cdot 7$ cm as determined from equations

(3.3)–(3.4) appropriate for an anomalous resistivity and with $v_A = 2000$ km s^{-1}, $\gamma = 6.9 \times 10^6$ and $5\gamma^{-1} = 7.2 \times 10^{-7}$ s. Thus if a current with $B_\phi = 500$ G is flowing in an arch of volume 10^{25} cm^3, only a volume of 7.2×10^{16} cm^3 needs to be occupied by the tearing regions at any one time to release 10^{29} erg in 100 s. The actual rate of energy release in one of the tearing regions is 1.4×10^{10} erg cm^{-3} s^{-1}, but in 7.2×10^{-6}s the energy in this region is burnt up with insignificant net mass motion and another region becomes unstable. For flares with a larger energy release a larger volume and/or magnetic field is easily sufficient. For example, with the same $B_\phi = 500$ G, a volume of 10^{27} cm^3 is sufficient to supply 10^{32} erg in 10^3 s. This most likely occurs due to multiple arches. Thus there is essentially no throttle to the rate of energy release in an arch for flare time scales, but the energy release takes place in many small volumes which sequentially become unstable.

Spicer (1976, 1977) suggested that a significant amount of the energy stored in the subphotospheric component of the current dissipating in the atmosphere can also be tapped during a flare. This statement requires qualification. Magnetic-field lines (and hence magnetic energy) cannot be transported faster than the Alfven speed v_A since they are tied to the fluid (e.g. Krall and Trivelpiece 1973). The total transport time is $\int dz/v_A \simeq H/v_A$, where H and v_A are the density scale height and Alfven speed in the photosphere, since v_A increases with altitude much faster than H does. We thus estimate (with $B \simeq 2000$ G and $n \simeq 10^{17}$ cm^{-3} in the photosphere) a total time of 40 s for subphotospheric magnetic energy to reach a height of 2000 km. The process is thus likely to be relevant only in larger flares of longer durations where the photospheric fields achieve the high value we have used.

The actual dissipation products for the magnetic or current energy converted in the tearing mode instability are heating, particle acceleration and fluid acceleration. Thus, while the tearing mode does seem to yield the necessary power, the theory of this instability is not presently sufficiently developed so that it can be shown that the amount of energy dissipated in each of these three dissipation products is correct for explaining the phenomena observed in flares.

4. Theory of secondary energy redistribution in the flaring atmosphere

4.1. Introduction

As we have emphasised throughout, the foremost theoretical flare problem is that of sufficiently rapid primary energy release. However, we have also shown in §3 that, by its very nature, the necessary field annihilation involves length scales well below the spatial resolution of instruments, available both currently and in the foreseeable future (cf §6). Consequently, attempts to support or discount any particular model in terms of overall field morphology may be highly misleading since the essential physics of the mechanism lies in the substructures. At present we therefore seem faced with only one alternative diagnostic approach and that is to investigate the properties of primary release models in terms of their secondary effects on the solar atmosphere on resolvable size scales. It is with this topic that we concern ourselves here.

Unfortunately such modelling of secondary energy redistribution processes has not yet been linked with specific primary release models. Rather, the procedure has been to adopt semi-empirical forms of energy input, based as much on other flare observations as on primary release mechanisms, model the atmospheric response to this and compare the result to flare atmospheres inferred empirically from observations

(cf §2 and Canfield *et al* 1978). The work can be subdivided temporally into flash and decay phase modelling and spatially into different atmospheric regions (similar to the discussion of observations in §2). Here we will again concentrate on the different spatial regimes of atmospheric heating in the *flash* phase, but will give some attention to the decay phase also.

4.2. Energy transport mechanisms

Energy is transported in flares by means of the usual three thermal mechanisms of conduction, convection and radiation, and in addition by the non-thermal particles present under flare conditions. (These are, of course, in addition to the magnetic energy carried by currents, which we regard as part of the primary process.) Energy is also present in the diverse plasma wave modes available but not in sufficient amounts to play an energetically significant role outside the primary release site, though important in diagnostics such as modelling of type III radio bursts.

Each of these four transport processes may play the role of either an energy input or an energy loss process in different regimes and at different times in the flare. At present the conventional wisdom is that in the flash phase primary energy release occurs in a highly localised region of the low corona and that in the first instance all four modes transport energy away from this site into the surrounding atmosphere. An informative introduction to the way in which the different modes may dominate in different atmospheric regimes is given by considering some of the first quantitative work on this problem by Korchak (1971), Brown (1973a) and Hudson (1972, 1973) (cf Dubov 1963). These authors took the observational interpretation (cf §5.1) that the hard x-ray-emitting electrons comprised a non-thermal stream and carried a substantial proportion of primary energy release away from their acceleration site, and by collisional absorption produced much of the flash phase heating. (Some support for this scheme stems from the observed synchronisation of the hard x-ray burst with impulsive thermal emissions at widely varying locations in the atmosphere, requiring a rapid energy transport mode.) Since the collisional degradation of one electron increases with the column mass $Y(\text{g cm}^{-2})$ traversed and decreases with the initial electron energy E_0, the atmospheric heating rate produced by an injected power-law electron spectrum $\sim E_0^{-\delta}$ suggested by hard x-ray data (cf §§2 and 5.1.1) the energy deposition by electrons should decline as $Y^{-\delta/2}$ (e.g. Brown 1973a) in a one-dimensional geometry (vertically descending electron stream).

Considering the remarks of §2.1 on the role of radiative instability in determining the structure of the solar atmosphere, it is clear that the decline of electron input with depth Y, and the rise of radiative losses (as the density increases with Y), results in the division of the flaring atmosphere into hot and cool regimes. (The effect of the particle input is thus to push the transition region to deeper layers during the flare (Brown 1973a).) Simplification of the energy transport problem was then attempted on the grounds that the low values and gradients of temperatures and pressure in the cool region permit convection and thermal conduction to be neglected there. With the further simplifying argument that the electron flux varied only on time scales longer than the chromospheric radiative response time, Brown (1973a) constructed steady-state models of the flaring chromosphere in which electron deposition everywhere balanced radiactive losses.

While this description gave a first crude description of the flare chromosphere, it failed to predict satisfactory Hα profiles (Canfield 1974) and disagrees with the

empirical models of figure 4. Furthermore, though the analysis gave a correct order of magnitude estimate of the extra mass of plasma heated to coronal temperatures (Brown 1973a, cf Lin and Hudson 1976), it provided no means of analysing the detailed properties of the hot flare plasma where conduction and convection are important. Since these first efforts, much development has gone into modelling the hot flare plasma, improving models of the chromospheric heating, and to understanding the interface of these hot and cool domains. In the following subsections we briefly review recent progress in each of these areas.

4.3. Heating of the low chromosphere

Canfield (1974) showed that the discrepancy between the predicted and observed Hα profiles might stem from the fact that Brown's net radiative losses were too high by a factor of 10–100, at some temperatures, compared to a detailed radiative transfer estimate. The source of this error lay in an inconsistency between Brown's energy equation and his non-LTE ionisation equation, arising from his neglect of the energy *input* by Balmer continuum absorption which controls the ionisation (cf Brown *et al* 1978). The result was that the depth of flare penetration was underestimated, resulting in too small a Stark broadening of the Hα line. When this error is corrected the dominant radiative losses are found to be Hα and Lα lines rather than the continua (though a numerical error in Canfield (1974) in fact underestimated continuum losses). As both of these are optically thick, solution of the energy equation by local balance considerations has to be abandoned and replaced by an iterative radiative transfer approach to the whole atmosphere. Brown *et al* (1978) have derived new electron-heated chromosphere models on this iterative basis and find that they *can* now adequately reproduce observed Hα profiles (figure 24) but not the complete chromospheric structure (Machado and Noyes 1978) provided they incorporate macroturbulence of velocities $\simeq 70$ km s^{-1} along the line of sight and horizontal inhomogeneity of the Hα emission.

However, in the one clean case of simultaneous observation of hard x-rays and Hα flash kernel spectra in the large flare of 7 August 1973 (Zirin and Tanaka 1973, Hoyng *et al* 1976), it is found that satisfactory heating of this region of the chromosphere requires $\lesssim 10\%$ of the electron energy flux implied by a non-thermal thick target explanation of the hard x-rays. This fact may have important implications for the primary flare mechanism (cf §5.1 and Brown *et al* 1979). Zirin and Tanaka (1973) and Zirin (1978) have presented convincing direct evidence for the heating of very small optical kernels by electron streams, in the form of their synchronisation of these optical flashes with hard x-ray spikes (cf §2). On the other hand (cf Švestka 1976) not all flares exhibit detectable hard x-rays, but given the small amount of energy required for *kernel* heating (as against the gradual Hα flare) this may be a threshold effect.

It must be borne in mind also that consistency of the electron-heated models with Hα spectra in no way establishes the mechanism unambiguously since the atmospheric structure may be insensitive to the input mechanism. Specifically, heating of the chromosphere by absorption of XUV radiation from the overlying hot flare plasma may also be effective (Machado *et al* 1978) particularly for the prolonged diffuse Hα emission (cf §2). This hypothesis would overcome the possible difficulty involved in Švestka's (1970) proposal for conductive heating of the chromosphere in the decay phase, viz the small conductive flux possible at the low temperatures of the Hα flare. In any event, any additional contribution to low chromospheric heating must further
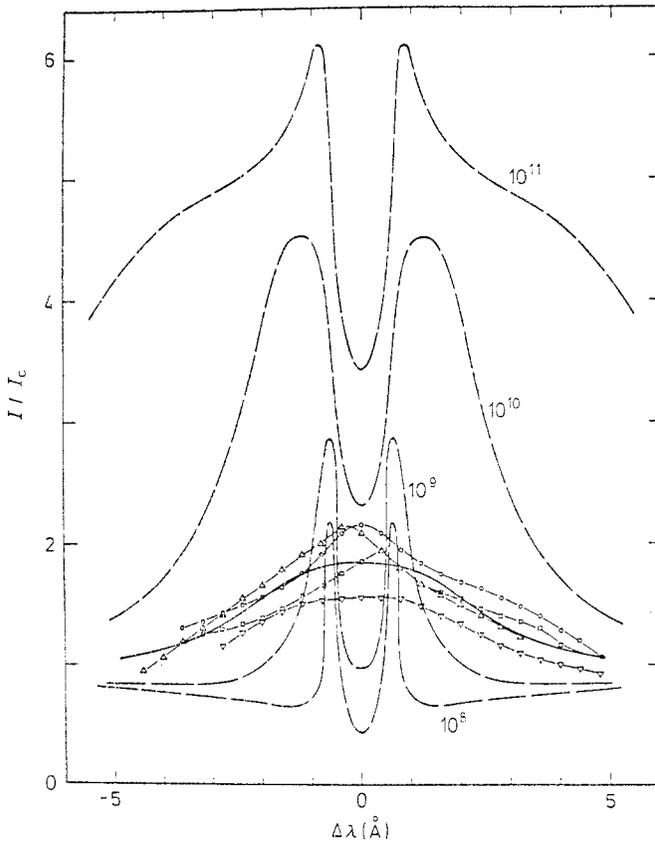
Figure 24. Comparison of observed Hα line profile data points in the flash kernels of the 7 August 1972 flare with the theoretical (broken curves) profiles predicted for chromospheres collisionally heated by a variety of electron bombardment fluxes (from Brown *et al* 1978). The full curve is the best fit model, allowing for inhomogeneity effects. Computed: ——, – – –; observed; △, ○, □, ▽.

reduce the non-thermal electron flux required below the small value inferred by Brown *et al* (1978).

4.4. *Heating of the photosphere and temperature minimum*

Penetration of flare heating right down to the photosphere in some cases is suggested by the occurrence of flares in white light (cf §2.2.1) and confirmed by the enhancement of the temperature minimum region (see § 2.1.2). Najita and Orrall (1970) and Švestka (1970) first calculated the particle flux and spectrum required for heating the photosphere and concluded that the region could not be heated by electrons descending from above since there is insufficient flux available among electrons of high enough energy ($\gtrsim 1$ MeV) to penetrate the large column mass involved. Instead they suggested that a reasonable flux of protons of $E \geqslant 10$ MeV could produce the effect, but recent γ-ray measurements (§2.3) suggest that this flux does not exist. The possibility that white light flares are not photospheric but are produced by non-thermal ionisation higher in the atmosphere is not yet excluded (cf Hudson 1972, Brown 1973b, Korchak 1974, Lin and Hudson 1976).

Machado *et al* (1978) have given a simple description of the temperature minimum region by balancing model flare inputs against the dominant H-negative losses. They conclude that electrons are not capable of heating even this level consistently with observed hard x-ray fluxes, with stability of the beam against wave–particle losses, and with the degree of heating observed in the overlying regions (cf §4.3). In addition they note (cf Cook and Brueckner 1979) that some of the temperature minimum emissions last much longer than the hard x-rays.

An input model based on absorption of soft x-rays emitted in the corona (cf Somov 1975, Henoux and Nakagawa 1977, 1978) also does not have the necessary penetrating power to heat the temperature minimum. For proton bombardment Machado *et al* (1978) argue that extrapolation of known power-law proton flux spectra does not give enough protons in the energy range necessary to heat this region and discount a bump in this part of the proton spectrum as very unlikely, though the relevant proton flux is not actually observed by γ-ray techniques nor by non-thermal Lyman α production. They propose instead some direct *in situ* heating of the region by Joule dissipation of flare currents deep in the atmosphere. An alternative requiring investigation is enhanced dissipation of acoustic flux bottled up below the active region (cf Pneuman 1969).

4.5. Heating of the upper chromosphere

This domain of the flaring atmosphere presents greater difficulties of theoretical modelling since the heating there may be too rapid to justify a quasi-steady-state treatment while the density is still high enough for radiative transfer effects (particularly in Lyα) to be important. Furthermore this region lies at the bottom knee of the transition region (cf figure 4) where conductive deposition is important due to the large $\nabla^2 T$. Nor is it clear what role may be played here by the impact of shocks driven down from the hot flare plasma (cf below). Preliminary quasi-static modelling of this region has been described by Lites in Canfield *et al* (1978). Attempts have also been made at simultaneous modelling of the hot and cool regions in both static and evolutionary limiting cases.

Shmeleva and Syrovatskii (1973) and others assumed a localised primary energy deposition and derived a temperature structure of the remainder of the atmosphere by steady-state solution of the energy equation (conduction versus radiation). However, in the coronal domain transient convection effects are certainly important (Craig and McClymont 1976) and in the deep chromosphere optical thickness and the nature of the lower boundary condition complicate the problem (cf Brown 1977). It does, however, provide a possible explanation of the observed spectral similarity between flare and quiet Sun transition region structures since the solution in this domain is independent of the form of the energy input apart from a scaling factor (cf Craig *et al* 1978). Secondly, Antiochos and Sturrock (1978) have described the peneration of the transition region to deeper atmospheric layers during a flare as a process of 'evaporation' to high temperatures (by radiative instability—see §2) of the chromosphere by the enhanced conductive flux from the corona.

The aim of such modelling is to explain the appearance in the flaring corona of a large plasma mass not previously present. It should be noted, however, that the mass so produced by 'evaporation' is highly insensitive to the form of chromospheric heating driving it and depends essentially only on the location of the peak in the radiative loss curve (Sweet 1969). It is thus misleading to suppose that any heating

model (e.g. particles—Brown (1973a), or conduction—Antiochos and Sturrock (1978)) is tested by the mass which appears in the corona.

4.6. Coronal heating

The chief simplifications which occur in modelling coronal heating are that the region is optically thin and that radiative losses can mostly be ignored except as diagnostics. On the other hand, the problem is outstandingly difficult in most other respects. Firstly, this is the region of strongest and most uncertain heating both from primary release of magnetic energy and from collisional heating by electrons in the low-energy end of their steep spectrum. Consequently the response is essentially transient and is dominated by conduction and convection processes driven by the steep temperature and pressure gradients set up. Secondly, the comparatively low density results in long collisional mean free paths and relaxation times. Consequently there is a strong possibility of generation of plasma wave modes which may drastically modify transport coefficients (e.g. the conduction coefficient, cf Spicer (1977), Brown *et al* (1979) and Smith and Lilliequist (1979)). In addition the resulting deviation from steady-state ionisation equilibrium complicates x-ray spectral diagnostics of the plasma (cf Shapiro and Moore 1977, Shapiro and Knight 1978 and references therein).

Most theoretical modelling to date has ignored the latter factors and concentrated on the former, and has been carried out under one of two further simplifying assumptions. Firstly, that in the immediate spatial and temporal vicinity of the primary heating the magnetic field is strong and simple enough to confine the mass motions and conduction to a one-dimensional loop. Or, secondly, that the region considered is sufficiently remote from the primary energy release that the latter may be regarded as a localised point explosion (or piston). Broadly speaking, the former approximation should describe initial evolution of flare loops in the low corona such as observed in the XUV by Skylab while the latter relates to the interplanetary motions, subsequent to the flare, as observed at the Earth, by white light coronograph and radio telescopes.

The first calculation of particle heating of coronal loops is probably that by Jefferies and Orrall (1965) in connection with optical emission from loop prominences containing energetic protons. Cheng (1971) calculated the heating of an x-ray loop during degradation of electrons trapped within it, neglecting all loss processes which naturally yields excessive temperatures. Strauss and Papagiannis (1971) (cf Davis *et al* 1977) analysed a somewhat less unrealistic situation in which the density was held constant (no convection) but conduction was permitted, giving plausible peak temperatures.

Kostjuk and Pikelner (1975) and Craig and McClymont (1976) first undertook a complete calculation involving convection and conduction. The latter paper adopted a loop model of initially uniform density (justified by the large coronal scale height) and infinite length, heating being maximal at and symmetric about the loop centre, while the former started with a detailed quiet model atmosphere. Energy, momentum and continuity equations are then solved numerically to yield space–time profiles of density, temperature and velocity. These two papers (and subsequent work by Kostjuk (1976a,b) and Bloomberg *et al* (1977)) are essentially in agreement as regards the main hydrodynamic behaviour of the hot loop material, as summarised below, but Kostjuk also gives (unconfirmed) results on the evolution of the descending driven thermal and acoustic fronts on reaching the density discontinuity of the transition region (cf Nakagawa *et al* 1973).

As far as the upper reaches of a heated loop are concerned, the evolution is essentially independent of the details of a localised energy input. On a time scale of the order of the transit time for a sound wave through a heating length scale substantial pressure and temperature gradients drive hydrodynamic and thermal fronts down both limbs of the loop. In this stage convective energy redistribution is at least as important as conduction. Craig and McClymont (1976) find that a regime is then approached in which the pressure becomes close to spatially uniform behind the shock front formed. They have also computed the resulting differential emission measure functions $\xi(T)$ (cf §2) for the entire loop, for comparison with spectral observations but note that some of the detailed (shock) structure in $\xi(T)$ which characterise the model would be very difficult to discern in spectral data for the reasons we discussed in §2. Typical results for the temperature profile are shown in figures 25(a) and (b) which emphasise the importance of convection, while figure 26 shows the evolution of $\xi(T)$. Detailed spectral diagnostics for particular x-ray lines and for the hard x-ray continuum have also been computed by Davis *et al* (1977) incorporating a full Fokker–Planck treatment of the collisional interaction of a beam of electrons and plasma particles without
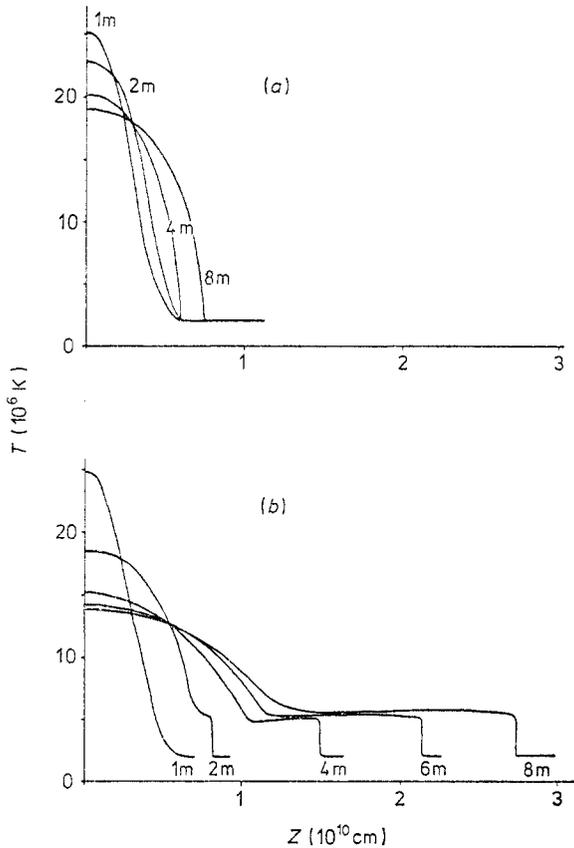
Figure 25. Time evolution of the temperature profile along a flux tube of initially constant density ($n=10^{11}$ cm^{-3}) heated by an input energy flux of 10^{12} erg cm^{-2} s^{-1} as described by Craig and McClymont (1976). In (a) convection has been artifically suppressed, leaving a conduction-dominated situation, while in (b) convection is included.
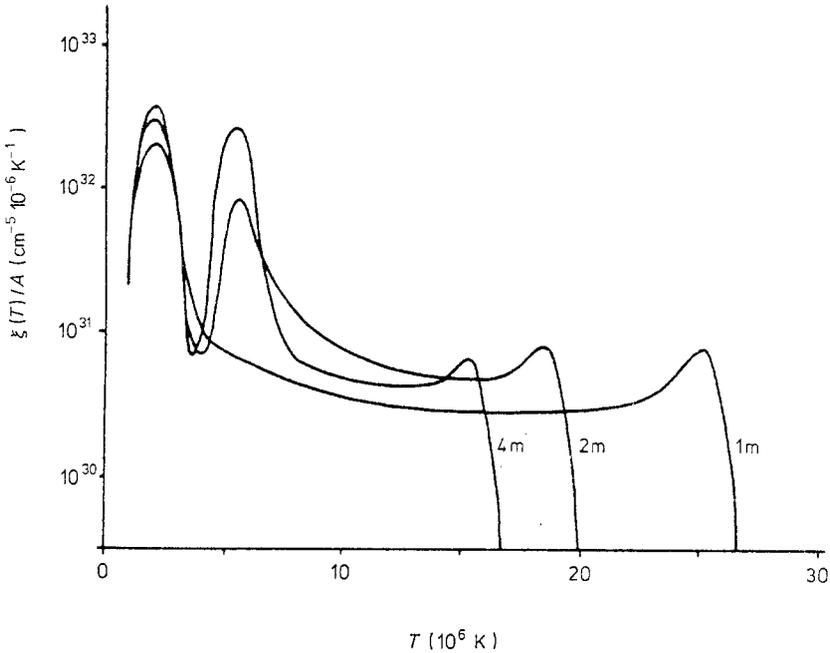
Figure 26. Time evolution of the differential emission measure function $\xi(T)$ for the heated flux tube model described in figure 25(*b*).

convection and by Bloomberg *et al* (1977) including convection, but unfortunately the calculations neglect heavy-element radiative losses (cf §2.1), which may invalidate their results at lower temperatures. In addition, their use of a fixed lower boundary temperature at 3×10^5 K means their work does not synthesise the hot and cool flare regions as they claim. An essential difficulty with all such computational models is that comparison of theory and observations requires detailed computation for each set of input parameters, so that it is hard to establish trends and to identify the key model factors. This has led to attempts at approximate analytic solutions in which the parametric dependence is explicit. For instance, Zaumen and Acton (1974) calculated the relationship which should exist between emission measure and temperatures during decay of an isothermal x-ray-emitting volume by conduction into a cool surrounding plasma, namely that emission measure should rise as temperature falls (as is usually observed). Following their proof that convective energy transport is more important (initially) than conduction, Craig and McClymont (in Canfield *et al* 1978) obtained a simple analytic (separable) solution of the hydrodynamic equations, based empirically on their (1976) computational result on the existence of a constant pressure regime. This led them to the same conclusion as Zaumen and Acton (1974) concerning the increase of emission measure during cooling though the time scale is shorter and the function relationship different in the convective case, but also yields an analytic form for the differential emission measure $\xi(T,t)$. These distributions must ultimately relax in the flare cooling phase toward a quasi-static conduction-dominated solution, the form of which is almost independent of the form of any energy input present and which is also applicable to active region loops (Landini and Fossi 1975, Craig *et al* 1978 Rossner *et al* 1978). Similar calculations of $\xi(T)$ have been performed for static and evaporative conductive cooling situations by Antiochos and Sturrock (1978).

The relationship between these models and observations remains rather obscure in our view. In particular, Dere *et al* (1974) published $\xi(T)$ profiles, inferred from observations during the entire progress of a flare, showing $\xi(T)$ to be monotonically decreasing (above $T \simeq 10^6$ K) at all times. On the other hand the results of Underwood *et al* (1978) show an increasing $\xi(T)$ in $T \simeq 10^6$–10^7 K. While the methods used in both of these cases to derive $\xi(T)$ are highly ambiguous for the reasons discussed by Craig and Brown (1976a) (cf §2) it seems doubtful that this large discrepancy can be so explained since it involves the gross form of $\xi(T)$ and not its details. Rather it must be tentatively attributed to variations between individual flares (the Underwood *et al* analysis is based on a subflare). Underwood *et al* (1978) claim further that the shape of their $\xi(T)$ results exclude both the Antiochos and Sturrock (1978) evaporative models of flare cooling. Perhaps more convincingly Underwood *et al* (1978) note that the emission measure declines with time as the flare cools, contrary to conductive and convective cooling models (see above) and also to evaporative models. With the added fact of a high inferred density, they therefore conclude that the late phase of flare cooling must be radiatively controlled, a conclusion reached observationally earlier by Zaumen and Acton (1974) and highly plausible on theoretical grounds since convection and conduction must decline in importance with the relaxation of pressure and temperature gradients. Estimation of the dynamical conductive and radiative cooling times from the parameters proposed by Underwood *et al* (1978), however, shows all three to be comparable ($\simeq 100$ s) even at this stage of the cooling process. Somov and Syrovatskii (1976) have recently discussed conditions under which radiation is a controlling factor.

A discussion of the heating and cooling of x-ray kernels at the very earliest stages of their formation has been given by Brown and Nakagawa (1978). They show that for all densities $\gtrsim 10^{10}$ cm^{-3} the size of these kernels is so small that they will be cooled both convectively and conductively in times much less than the observed burst duration. They conclude that this gives direct evidence that flare primary energy release must be a dynamic process occurring successively at many small sites within the observed volume (cf §3) and propose that this explains the persistence of a transient ionisation state throughout the burst necessary to interpret the high apparent densities measured (see §2). The implied highly inhomogeneous structure of a single loop in active regions also would provide a satisfactory alternative to the proposal by Craig *et al* (1978) and Rossner *et al* (1978) that many separate loops (each of monotonic temperature structure) are required to explain the observed $\xi(T)$ for active regions.

In the other limiting regime of outer coronal mass motions, extensive numerical computations have been carried out by Nakagawa *et al* (1975) and Wu *et al* (1975), for example (for a detailed survey see the Rust *et al* (1978) chapter in Sturrock *et al* (1978)). These calculations now encompass MHD effects, as well as pure hydrodynamics, and simulate some of the broad features of outer coronal and interplanetary disturbances following flares. However, they have no direct bearing on the primary energy release process which interests us here and provide no clear interpretation of many of the observed features of white light coronal transients (cf discussion in Rust *et al* (1978)). Essentially these models devolve about the response of matter spatially remote from a driving force which is either instantaneous (point explosion) or prolonged (piston). An alternative approach is again to seek analytic solutions here in terms of the classical self-similar hydrodynamic solutions for a point explosion (Sedov 1959, Obashev 1966, Korobeinikov 1969, Vesecky and Meadows 1969, Guseinov 1973).

5. Particle acceleration

5.1. *First phase acceleration*

5.1.1. The implications of hard x-ray bursts. We have seen in §2.5 what direct observations tell us concerning high-energy particles in flares. However, we note that in principle the chief problem of particle acceleration does not lie in their high individual energies since a 100 G field reduction over dimensions of the order of 10^4 km in 10^2 s induces a voltage drop of the order of 10^{10} V (cf Sweet 1969). Rather the chief problem lies in accelerating sufficient particles (cf §5.2). The most stringent demand posed in this respect has been the electron flux suggested by hard x-ray burst analysis in the past decade. It is therefore our intention in this section to review critically hard x-ray burst interpretation (cf recent reviews by Brown (1975, 1976)).

The power-law spectral character $\sim \epsilon^{-\gamma}$ of most hard x-ray bursts has led to the almost universal belief that they are non-thermal in character. That is, they arise by collisional electron–ion bremsstrahlung (cf Haug (1976) for contribution of electron–electron bremsstrahlung) of an energetic electron population of spectrum $E^{-\delta}$ in a cool background plasma, rather than by thermal bremsstrahlung which is characterised by a spectrum $\sim \exp\left(-\epsilon/KT\right)\epsilon^{-1}$ for an isothermal plasma at temperature T. This belief persists despite the fact that the flare plasma is patently non-isothermal and despite the proof on qualitative grounds (Chubb 1971), empirical grounds (Davis and Rogerson 1977) and analytically (Brown 1974) that a suitably inhomogeneous thermal plasma can mimic a wide range of spectral forms, and particularly a power law. This danger of analytic oversimplification in, for example, spectral interpretation is one which pervades astrophysical modelling and is additional to the *numerical* problems of uniqueness in spectral data fitting raised by Craig and Brown (1976a,b) and discussed in §2 for soft x-rays but which are equally relevant here (e.g. Brown, 1975, 1976, 1978, Craig 1978). On the other hand, direct interplanetary evidence on the existence of some accelerated electrons does exist. Therefore, for the moment we will adhere to the non-thermal interpretation and pursue its consequences.

The earliest observations of hard x-ray bursts led Takakura and Kai (1966) to propose that electrons were accelerated up till burst maximum and thereafter simply decayed in a coronal trap of low density ($\sim 10^9$–10^{10} cm^{-3}) such that the collision time equalled the burst decay time (10–100 s). The microwave burst properties of this model were further developed by Holt and Ramaty (1969). Such a collisional model should produce a hardening of the burst spectrum during the decay contrary to observations (Kane and Anderson 1970). This difficulty might be overcome if electrons of different energy move in regions of different density in the source (cf Benz and Gold 1971, Brown 1972a) such as might occur if the acceleration occurs within small magnetically confined kernels in a highly inhomogeneous volume (cf Spicer's (1976) arch model). Brown (1973b) pointed out that complex burst time profiles were incompatible with the low density trap model unless the trapping field itself was time-varying and so would drive the electrons present, the situation being no longer collision-dominated. Analysis of burst data have now in fact been carried out for model situations in which collisions are negligible in comparison to betatron (Brown and Hoyng 1975) and Fermi (Brown and McClymont 1976) action of a time-varying trap field on the electrons. In addition, electrons may of course be further continuously accelerated by the primary processes described in §5.2 due to field dissipation in the trap. The situation of a high density static trap ('thick target trap'—see below) into which electrons are accelerated continuously has been modelled by Bai and Ramaty

(1979) who proposed trap density $\simeq 3 \times 10^{10}$ cm^{-3} for the events considered. Furthermore, precipitation of electrons into the dense trap limbs, due to pitch angle diffusion, inevitably plays a role since the collisional deflection and energy loss times are comparable (Hudson 1972, Melrose and Brown 1976). This led Melrose and Brown (1976) to establish a relationship between the hard x-ray burst dynamic spectrum from a trap with precipitation and the rate of electron acceleration in the trap, based partly on an analogy with magnetospheric particle precipitation (Kennel and Petscheck 1966) and partly on Brown's (1971) solution of the hard x-ray spectrum deconvolution problem.

An alternative regime is one in which electrons are accelerated in the corona with small pitch angles so that they are directly injected into the dense chromosphere, this situation being commonly known as the continuous injection or thick target model (e.g. Brown 1971, Hudson 1972, Petrosian 1973). The latter name derives from the fact that the bremsstrahlung target is collisionally thick, absorbing the entire electron energy. One immediate difficulty of the thick target with no trapping is that the microwave emission is too weak due to the small pitch angles and high densities involved. This might be overcome by an asymmetric magnetic arch situation, envisaged by Mätzler (1976). Because the transit time of an electron from corona to the chromosphere is less than burst time resolution (i.e. $\lesssim 1$ s), the thick target electron injection rate can be identified with the simultaneous rate of bremsstrahlung emission without the time deconvolution necessary in a trap plus precipitation regime. Furthermore, because bremsstrahlung radiation and collisional energy loss have the same dependence on ambient density the bremsstrahlung emission is essentially independent of the density distribution in the thick target (Brown 1971, Hudson 1972). Consequently it is possible if energy losses are collision-dominated in this case to express the electron injection (acceleration) spectrum at any instant uniquely in terms of the bremsstrahlung spectrum via an integral equation which, with the non-relativistic Bethe–Heitler cross section, reduces to Abel's equation and so can be inverted analytically (Brown 1971). In the particular case of a power-law photon flux spectrum $\sim\epsilon^{-\gamma}$ the corresponding electron injection flux spectrum is $\sim E^{-\gamma-1}$ whereas in the trap situation the flux spectrum in the trap is $E^{-\gamma+1}$, the difference arising due to the energy-dependent loss rate in the thick target. Whether the thick target model represents the actual flare situation or not, it is useful in that the electron flux required *is* independent of the target so it can conveniently be used for x-ray burst data analysis as illustrated in figure 27 (from Hoyng *et al* 1976) Further, because the collisional energy losses in a thick target represent the minimum total loss (i.e. no collective losses to plasma waves, for example) for a *non-thermal* situation, the model defines the maximum bremsstrahlung yield, i.e. the minimum electron flux required, which we reconsider shortly.

The observational status of these two models is still debated (cf Kane *et al* (1978) in Sturrock *et al* (1978)). Briefly the observation of hard x-ray emission from flares behind the limb suggest a trap component as do the synchronous microwave burst data. On the other hand, the close association of bursts with chromosphere optical and UV flashes points to a thick target component. The bremsstrahlung polarisation and anisotropy properties have also been calculated (Elwert and Haug 1970, 1971, Haug 1972, Brown 1972b, Petrosian 1973, McClymont 1976). Marginal agreement is found between the polarisation data (see §2.3) and the radial polarisation plane predicted for the thick target but is not confirmed by directivity studies. In principle such data could tell us the electron velocity distribution at acceleration. Datlowe and Lin (1973) have argued that neither a trap nor a thick target model agrees with the observed relationship between hard x-ray spectral indices and spectral indices for
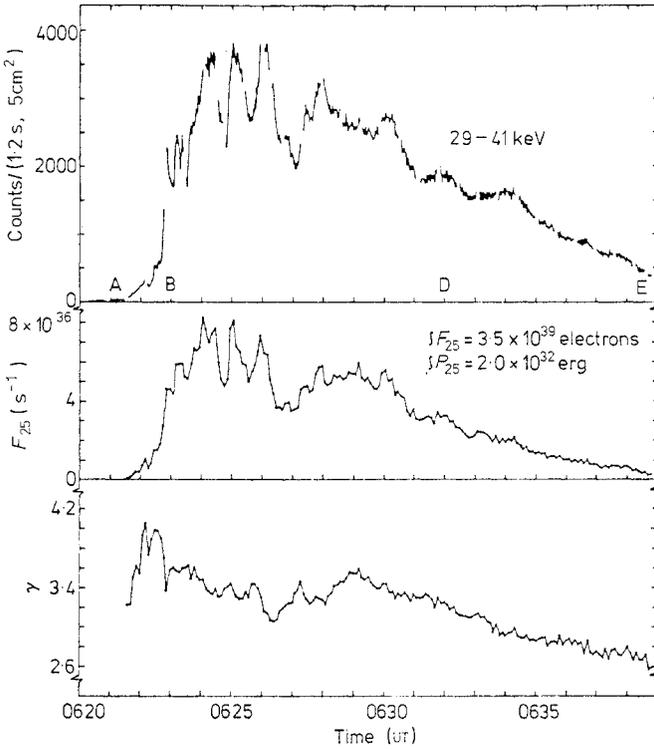
Figure 27. Collision-dominated thick target electron injection rate F_{25} (t) above 25 keV, and x-ray spectral index $\gamma(t)$, throughout the large flare of 4 August 1972 (cf figure 10) shown together with the lowest channel hard x-ray flux (from Hoyng *et al* 1976). The electron spectral index $\delta = \gamma + 1$. Also indicated are the total number and total energy of electrons ($\geqslant 25$ keV) injected.

interplanetary electrons, the escape probability of the latter being assumed energy-independent. Since only $\lesssim 1\%$ of electrons are required to escape in this situation, their assumption seems weak. In any case Melrose and Brown's (1976) analysis of the trap plus precipitation situation shows that it would be compatible with the Datlowe and Lin (1975) result. Further support for this mixed model comes from the modelling of impulsive EUV bursts (Kane 1974, Donnelly and Kane 1978).

What the thick target and trap situations have in common is the theoretical difficulty of providing sufficient electrons for their requirements. For, apart from a spectral factor of order unity, the two models require the same total electron numbers N_0. This is so because although the high ambient density n_0 of the thick target permits proportionately fewer electrons N_1 instantaneously present in the source, for a prescribed instantaneous burst intensity ($\sim n_0 N_1$) the collisional lifetime of these electrons is correspondingly lower and they have to be replenished that much more frequently throughout the burst duration (cf Brown 1971, 1975). It follows that the total number of electrons required is the total number of photons emitted times the ratio $Q_{\text{coll}}/Q_{\text{Bremss}}$ of the energy-loss cross section to the bremsstrahlung cross section, or the inverse of the bremsstrahlung efficiency. In the energy range 10–100 keV this efficiency is of the order of 10^{-4}–10^{-5}. The resulting electron numbers range fron 10^{36}–10^{39} above 25 keV with an associated acceleration rate $F \simeq 10^{34}$–10^{36} s^{-1} and a total energy 10^{29}–10^{32} erg (cf Hoyng *et al* (1976) and figure 27).

All these figures rise rapidly if the steeply decreasing non-thermal spectrum extends much below 25 keV. The difficulty is that the total energy involved is comparable to the entire energy dissipated in the flare, demanding an improbably high acceleration efficiency from the primary dissipation process. Furthermore, we have seen in §4 that perhaps only a fraction ($\lesssim 10\%$) of this energy is actually needed for electrons to heat the chromospheric flashes associated with the hard x-rays. The number of electrons is also formidable, comprising virtually all the electrons in the flare volume. Colgate (1978) has recently argued that the vast self-magnetic energy (greatly exceeding that of the flare field) of these electrons in the form of a directed flux (current $\simeq 10^{15}$–10^{17} A over an area $S \simeq 10^{19}$ cm^2) invalidates non-thermal models, as had been noted earlier by Hoyng *et al* (1976), Brown and Melrose (1977) and Knight and Sturrock (1977) (cf Melrose 1974). The result is that a neutralising reverse current is set up among the ambient electrons, this current being capable of supplying the required large numbers of electrons to the coronal acceleration site. One limiting factor is the stability of this reverse current against losses by plasma wave generation (Hoyng *et al* 1978, Hoyng and Melrose 1977, Hoyng 1977a). Specifically, if the required reverse current electron flux $n_0 v_0 = F/S$ has a velocity v_0 exceeding the ion-sound speed it is unstable to generation of ion-sound waves. It is of interest that the stability criterion is just about satisfied by typical thick target fluxes F/S for acceleration sites located at reasonable coronal densities $n_0 \gtrsim 10^9$–10^{10} cm^{-3} (Brown and Melrose 1977). Thus the primary problem does not lie in the total electron *numbers* but rather in the efficiency demanded by their total energy.

A further important consideration to emerge recently (Knight and Sturrock 1977) in relation to reverse currents is the role of ohmic dissipation of the reverse current as an energy loss to the electron beam. The non-zero resistivity of the solar atmosphere demands that a small electric field be maintained, to drive the reverse current against the resistance, which decelerates the beam electrons, their energy appearing as Joule heating of the reverse current. From available analyses (Knight and Sturrock 1977, Hoyng and Melrose 1977, Hoyng *et al* 1978, Emslie 1979) it appears that the reverse current loss may completely dominate the beam at the start of the flare and in many cases remain important even after flare heating, particularly high in the atmosphere. The most important implication is that, due to the reduced lifetime of electrons, conventional descending electron-beam models are even less efficient as hard x-ray sources than was previously thought from collisional thick target analysis, and so may possibly be precluded on the grounds of impossibly high electron fluxes. This leaves open in such cases the possibilities that: electron beams exist but are accelerated in very dense regions rather than in the corona; a thick target situation exists in the corona but with electrons trapped (zero current) due to their large pitch angles (Bai and Ramaty 1978); the hard x-rays are quasi-thermal.

We consider here the third option, namely that the collisional energy losses themselves are reduced if the source electrons are part of a relaxed distribution of all the electrons (with mean energies $\gtrsim 10$ keV or $T > 10^8$ K) present rather than comprising a fast component in a cool background.

As discussed at length by Brown *et al* (1979) and Smith and Lilliequist (1979) (also by Crannell *et al* (1978) and Mätzler *et al* (1978) under different assumptions—see below), the essential point of these 'thermal' models is that in a confined thermal plasma with equal electron and ion temperatures (T_e, T_i) the only electron energy loss is by bremsstrahlung radiation. Thus 100% of the electron energy goes into radiation as against $\leqslant 0.01\%$ for a non-thermal component as described above. Even if $T_i \ll T_e$

the non-radiative (electron–ion collision) energy loss is still about 100 times lower than in the non-thermal case. There are two snags encountered in applying this argument to the flare x-ray burst problem (Kahler 1971a,b, 1975). Firstly, for acceptable plasma densities, the electron–electron collisional mean free path and relaxation time exceed, respectively, feasible source sizes and burst durations so that the distribution cannot be collisionally relaxed. Secondly, the hot plasma cannot be confined completely and it should rapidly cool not collisionally but by conduction and convection of thermal electrons along the magnetic field at a speed $\sim (kT_e/m_e)^{1/2}$. For plausible source lengths this cooling occurs on times much less than some burst durations (even when the conduction problem is solved self-consistently with constraints posed by the observed burst spectrum on the temperature structure (Brown 1974)). This demands constant supply of energy throughout the burst, so defeating the efficiency advantage of the thermal source, unless the rapid cooling can be inhibited by anomalous conductivity effects (Brown 1974). Brown *et al* (1979) and Smith and Lilliequist (1979) have now shown that such inhibition of conduction should indeed occur. The initial rapid efflux of hot electrons, across a steep temperature front, drives a neutralising reverse current (among the electrons of the cool surroundings) which, having a velocity much in excess of the ion-sound speed $v_s \simeq (kT_e/m_i)^{1/2}$, generates ion-sound waves. In turn this ion-sound turbulence decelerates the hot electrons and inhibits their outflow much more effectively than collisions can. Additionally, interaction of the electrons through the wave turbulence allows a relaxed electron distribution (not necessarily Maxwellian) to be achieved without high densities. It is shown by Brown *et al* (1979) on approximate analytic grounds, and by Smith and Lilliequist (1979) using detailed computation of the transport coefficients, that the net result is the bottling up of the hot electrons in a near-isothermal state behind a very thin collisionless conduction front which spreads the energy into the surrounding plasma at speeds around v_s (43 times slower than the free expansion rate). The resultant improvement in bremsstrahlung efficiency over collision-loss-dominated non-thermal situations varies directly as the source density n_0. A value of $n_0 \gtrsim 10^{11}$ cm^{-3} is consistent with that of soft x-ray kernels and results in a total thermal electron energy requirement of a few per cent of the non-thermal model result. This is also compatible with the numbers of interplanetary electrons seen directly in space and (possibly) required to heat chromospheric flashes (Brown *et al* 1978). However, the high density and small volume of the thermal source would produce much too small a microwave flux from the region. But in the Brown *et al* (1979) model electrons of velocities $\gtrsim 2.6$ times the mean electron velocity escape through the turbulent front to fill a larger volume, which may be the microwave source. Another problem is that any conductive cooling process should result in an increase of emission measure as the temperature falls, directly contrary to the observation that hard x-ray bursts soften during their decay (cf §2.3). Brown *et al* (1980) have shown how the basic concepts of the model can be generalised to satisfy this observation if the hard x-ray burst arises from rapid sequential production of many hot cores (such as would occur in Spicer's (1976) arch tearing mode model), each cooling as just described but in times much less than the burst duration. The evolution of emission measure versus temperature is then governed by the rate of production of cores of different emission measures and temperatures.

An alternative form of thermal model which obviates the two problems discussed above has been proposed by Crannell *et al* (1978) and Mätzler *et al* (1978). Instead of heating the source by a dissipative process and cooling it by conduction, these authors propose that these stages result, respectively, by successive adiabatic compression and

expansion of a plasma kernel by magnetic-field variations. This model has the property that the emission measure n^2V rises and falls (due to density squared variation overwhelming the volume variation) with temperature as observed. The authors propose a rather low density ($\approx 10^9$ cm^{-3}) in order that the model directly produce the correct microwave burst flux but still have to invoke escape of high-energy electrons to explain the microwave burst decay profile. What seems to be lacking theoretically is a satisfactory explanation of why the dissipative conductive cooling process, which controls the Brown *et al* (1979) thermal model, should not occur along the field lines. In addition the maximum field compression factor possible (cf Brown 1973b) is insufficient to produce the necessary temperatures without some dissipative preheating being involved.

The (dubious) rise/fall symmetry of hard x-ray spikes (cf §2.3) is neither a necessary nor a sufficient condition for an adiabatic process. For instance, dissipative heating of a region of size L may be characterised by a time scale $\tau_1 \simeq L/v_A$ where v_A is the Alfven speed while in Brown *et al* (1979) the cooling time is $\tau_2 \simeq L/v_s$. Approximate equality of rise and fall times τ_1 and τ_2 would follow in situations where magnetic and gas pressures are comparable ($v_A \simeq v_s$). Nevertheless, aside from the time scale involved, the actual locus followed in an emission measure against temperature plot of burst evolution is remarkably similar between rise and fall (figure 28) and demands an explanation involving some element of reversibility.

At the present stage of development of thermal hard x-ray burst models we can thus conclude that they seem capable of reducing the electron energy requirements by at least an order of magnitude from non-thermal model values and permit this energy to appear as a bulk electron energisation with random motions rather than as ordered electron motions.
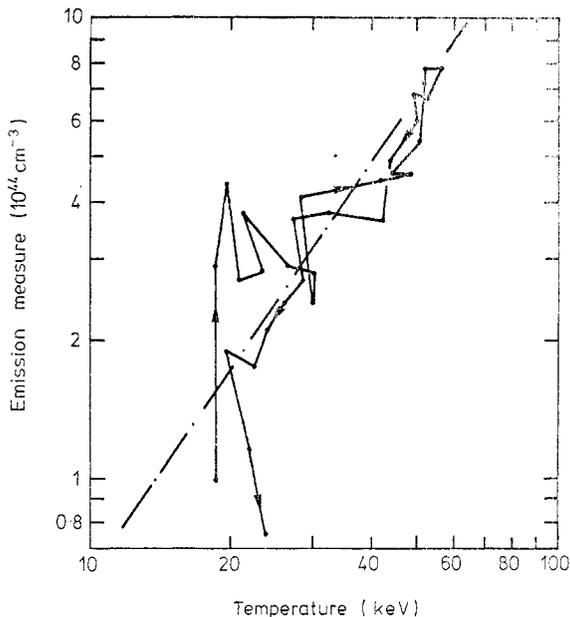

Figure 28. Locus described in the plane of total emission measure versus temperature inferred by Crannell *et al* (1978) and Mätzler *et al* (1978) for the simple spike hard x-ray burst of 1 March 1970.

5.1.2. First phase mechanisms. From §§5.1.1 and 2.5 we may conclude that the *minimum* requirements of the flare flash process in terms of high-energy particles are: bulk energisation of 10^{36}–10^{38} electrons to a near-isotropic distribution with an effective temperature of up to 5×10^8 K to explain the hard x-rays; *acceleration* of 10^{35}–10^{37} electrons above 25 keV with ordered velocities and a quasi-power-law spectrum to explain interplanetary observations; injection of sufficient electrons and ions into a regime capable of producing second-stage acceleration to produce particle bursts in the cosmic-ray regime. The last topic is considered in §5.1.3.

The first requirement does not appear to pose any serious theoretical difficulties. Specifically, the van Hoven and Cross (1973) tearing mode analysis showed that about 90% of the total energy release went into heating (i.e. bulk energisation) though this may be somewhat reduced by a full kinetic treatment of particle runaway and acceleration. Similar results are obtained by numerical simulation (Boris *et al* 1970) and in the laboratory (Hirosa and Skarsgard 1976). Thus if a sufficient total energy release rate for the flare is achieved (§3) the necessary result follows. (It remains of course to be shown that the heating will produce hard x-ray emission kernels of the correct temperature but, in view of the very small volumes in which energy is initially released in the tearing mode, the necessary high temperatures in the impulsive emissions seem qualitatively plausible.) This dominance of bulk energisation as an energy sink lies in the sequence of events following application of an electric field to the plasma. The initial free acceleration of electrons is limited by wave generation as they approach a critical drift velocity v_D. However, the resultant heating of the plasma to a mean thermal electron velocity $v_e \simeq v_D$ removes the instability and the electrons again freely accelerate until they approach a new critical velocity for a further instability to set in (cf, for example, Kaplan *et al* 1974, Smith 1977c). The net result is that *acceleration can only occur to the extent that the plasma is also heated*, which means that 50–75% of the energy intended for acceleration goes to heating, depending on the number of thermal degrees of freedom.

Acceleration itself can occur either by the action of a direct electric field along the magnetic field, arising from changing magnetic fields or by charge separation, or stochastically by wave–particle interactions. As regards the former, polarisation fields may be discounted under normal conditions due to the rapidity with which currents driven in the highly conducting plasma neutralise the field. Substantial reduction of the conductivity due to plasma instabilities does slow this neutralisation and has been observed to result in the formation of an electrostatic double layer in the non-linear stage of the Buneman instability (Quon and Wong 1976). However, coronal conditions permit the Buneman instability only for very large current densities in small regions such that even the resulting 10^5 reduction in conductivity is inadequate to produce flare energy release rates (Smith and Priest 1972). Thus the necessary fields must arise from the magnetic-field motions which occur in the tearing mode (e.g. Drake and Lee 1977) or in current sheets (Vasyluinas 1975) resulting in acceleration as described by Smith (1974a) for purely two-dimensional magnetic fields. Motion of lines of a field B_f into a region of (Z direction) field reversal at velocity v_f induces an electric field

$$|E| = v_f B_f / c \qquad (5.1)$$

felt by a particle of mass m for a time

$$\tau_A = \pi \, mc/e \, B_y \qquad (5.2)$$

where B_y is the average value of the transverse magnetic field which the particle experiences before it leaves the reversal region. The resulting particle energy gain is

$$\Delta E = 2m \, v_f{}^2 \, B_f{}^2/B_y{}^2 \tag{5.3}$$

where thermal velocities and microinstabilities are neglected. The effect of the former is a spread about energy gain (5.3) due to a distribution of the effective B_y sampled. Microinstabilities also lead to a spread in ΔE through scattering of some particles prematurely out of the field reversal region and prolongation of the presence of others. Additionally microinstabilities may dramatically increase ΔE by causing sudden jumps in v_f through dynamic changes in conductivity, an effect which may explain some pulsating phenomena in flares (Smith 1977a). Smith (1974b) tabulated the energy gains arising from (5.3) for a variety of v_f and B_y with $B_f = 500$ G. While gains $\Delta E \gtrsim 100$ keV do appear they occur only for $B_y \lesssim 0.5$ G (and $v_f \gtrsim 100$ km s^{-1}), i.e. only for electrons very close to the neutral sheet. Most electrons will experience higher B_y and consequently smaller ΔE (e.g. $\Delta E \simeq 1$ keV for $B_y = 5$ G with the other parameters unchanged). Thus it seems that direct electric fields (induced by moving magnetic fields) can accelerate electrons but probably not with the correct distribution and with a substantial fraction of the energy expended going into heating. We expect this to be true even when more realistic three-dimensional configurations are considered such as tearing modes in an arch (Spicer 1977) where the $|E|$ of (5.1) acts along B. To achieve the necessary electron distribution, acceleration and redistribution by wave–particle interactions probably must occur (see below). However, some degree of such direct acceleration and heating must occur first since, as we shall see, wave interactions can only act on an existing non-thermal or multi-thermal particle component. One important further restriction can be placed on the direct acceleration process, related to the problems of reverse current mentioned in §5.1.1. The condition arises from the fact that for the accelerated current to arise, its magnetic energy density must not exceed its kinetic energy density. For example, in a cylindrical beam (current) of electrons of density n_b the radius R must satisfy

$$R \lesssim \frac{6cm}{(n_b/10^9)^{1/2}} \tag{5.4}$$

under typical solar conditions. The smallness of this dimension shows that many such beams must be involved (cf Hoyng 1977a).

Eastwood (1972, 1975) pointed out that the distribution of particles accelerated along a given field line is at any instant quite narrow and directed opposite to the incoming particles. Since the velocity distribution peaks will be separated by more than v_e, the situation is unstable to generation of electron plasma waves and ion-acoustic waves (Stringer 1964) with very small growth times. As in the application of a direct electric field to a plasma, these waves probably ensure that a substantial fraction of the energy expended goes into heating. If the resulting turbulence were isotropic the particle scattering produced would prevent effective particle acceleration except for a few ions with high injection velocities (Friedman 1969). In fact, in a sufficiently strong field, the turbulence is much nearer to one-dimensional along the field (cf Kaplan and Tsytovich 1973, Smith 1974b). It is just such anisotropic wave turbulence which can further accelerate the particles.

Since, even to produce interplanetary electron fluxes, a fairly efficient acceleration or redistribution process is probably required we may restrict ourselves to those wave modes which carry a large fraction of their energy in electric fields, viz electron plasma

waves and whistlers. However, whistlers are probably of secondary importance since they can only be effectively excited by quite anisotropic distributions of electrons already accelerated (Melrose 1974) which are most likely absent after the initial acceleration. Therefore electron plasma waves which can resonantly interact with particles from $\sim 2v_e$ up to c are the prime candidate for wave acceleration (Benz 1977, Hoyng 1977a,b, Smith 1977b,c). Thus we consider mechanisms for electron plasma wave generation in more detail.

First, electron–electron or electron–ion two-stream instabilities require at least one streaming component. In the former case the stream density must be comparable to the ambient plasma density and of velocity $> v_e$. This is tantamount to already having the desired accelerated electrons. In the latter case, the low-frequency waves ($\ll \omega_{pe}$) generated do not propagate outside their region of generation and so can only redistribute the electrons of $v > v_e$ producing them. Second, quasi-linear relaxation of a *low-density* stream produces plasma waves of frequency $\simeq \omega_{pe}$ which can propagate into the plasma and interact with other electrons. Again streaming particles of $v > v_e$ are required and the quasi-linear relaxation process can at most redistribute already non-thermal energy when combined with the fourth candidate (below). Third, in cases of isotropic electron distributions, the presence of an extended non-thermal tail or of a gap in their energy distribution can also produce waves close to ω_{pe} but of much lower intensity than in the case of quasi-linear relaxation. The fourth, and only possibly viable mechanism in the presence of only electrons streaming at $v \lesssim v_e$, is non-linear amplification of anisotropic ion-acoustic wave energy (cf Tsytovich *et al* 1975). It remains unclear, however, whether other processes inhibit such amplification. For instance, in many laboratory experiments (e.g. Hamburger and Jancarik 1972) on shock and turbulent heating involving high levels of anisotropic ion-acoustic waves, high intensities of electron plasma waves were also found only in conjunction with the presence of suspected runaway ($v > v_e$) electron streams. Thus quasi-linear relaxation of the initially accelerated electrons appears as the best candidate for generation of electron plasma waves.

A possible manner in which this may occur was pointed out by Smith (1977c). As noted above plasma which is bulk-energised to high temperatures and accelerated must be confined to relatively small volumes and is continuously leaking out from these volumes either *en masse* or as a select group of high-velocity electrons. In a three-dimensional approach to the former case it is likely that hot electrons will interact with cold electrons, leading to a quasi-linear situation. In the latter case it is certain that a weak beam or quasi-linear situation will develop. Either or both of these quasi-linear situations will relax, leading to the production of electron plasma waves. It was estimated by Smith (1977c) that 10–20% of the initially accelerated electron energy could be redistributed to electrons by the process described below.

Irrespective of the wave production mechanism, we can give a brief account of how the necessary efficiency of acceleration or redistribution is achievable *if* a sufficient energy density W_p of electron plasma waves is postulated *ad hoc* (cf Kaplan *et al* 1974, Hoyng 1977a,b, Smith 1977b,c). In the first instance the waves are likely to be produced with large wavenumbers k (low phase velocities). Non-linear scattering processes (Tsytovich 1970) transform the waves to small k until

$$W_p/n_e K\, T_e > k^2\, \lambda_{De}{}^2 \text{ or } (\Delta k)^2\, \lambda_{De}{}^2 \tag{5.5}$$

where $\lambda_{De} = v_e/\omega_{pe}$ is the electron Debye length and Δk is the width of the plasma wave spectrum in k space. At this point the ponderomotive force of the plasma waves

overcomes the dispersive properties of the plasma, forcing the plasma waves to form spatially isolated regions of depleted density and high wave intensity called solitons (Zakharov 1972). Subsequently further density depletion and intensity increase drives the plasma waves trapped in the solitons to large k where they are heavily Landau-damped by the tail of the electron distribution which absorbs the wave energy and extends up to several tens of keV for typical solar conditions. The process is especially effective in the presence of a moderate magnetic field with $\omega_{ce} \lesssim \omega_{pe}$ (ω_{ce} is the electron cyclotron frequency), non-linear scattering and modulational instability occurring in two narrow cones along the magnetic field (Smith 1976b, 1977c). As noted by Kaplan *et al* (1974) and Smith (1977c) the process is capable of explaining not only the high efficiency of acceleration but also plausible electron spectra including power laws and the preferential acceleration of electrons in most flares. Under most flare conditions ion-sound turbulence cannot survive long enough to raise ion energies up to $v_i > v_e$ ($E \gtrsim 10$ MeV) necessary for electron plasma waves to act on them.

5.2. Second phase acceleration

An additional requirement on the first phase acceleration mechanism is that it must provide particles satisfying the injection conditions for second phase acceleration which is defined as the phase in which relativistic ($E > 200$ keV) electrons and protons ($E > 200$ MeV) are accelerated. The evidence for two phases of acceleration is summarised in §2.5. To understand the injection conditions we must first consider second phase acceleration processes. From the 4 August 1972 flare it is known that these processes must follow the first phase processes essentially immediately ($\lesssim 2$ min) (e.g. Bai and Ramaty 1978) and provide on this time scale ~ 1 MeV electrons required by type IV radio bursts (see §5.3) and several MeV ions to produce γ-ray bursts (see §2.3). Fermi (1949) proposed that these particles could be accelerated by making collisions with moving magnetised blobs. Thompson (1955), Davis (1956) and Parker and Tidman (1958) realised that effective acceleration by this mechanism, now known as Fermi acceleration, requires effective scattering of the particles to keep them isotropic. The theory of this so-called 'resonance scattering' has since been developed in detail and forms an important, often unrecognised, ingredient in present theories of acceleration by hydromagnetic turbulence which is just generalised Fermi acceleration. We shall first consider the Fermi mechanism assuming the presence of resonant scattering and then consider the source of this scattering.

The basic configuration for Fermi acceleration is that a particle collides with a magnetic field moving at velocity V and is reflected by it. To find the change ΔE in the particle's energy we use the fact that in the rest frame of the field lines the particle's energy does not change (Morrison 1961, following Melrose 1980) which leads to

$$\Delta E = \frac{2}{1 - V^2/c^2} \left(\frac{V^2}{c^2} E + \boldsymbol{p} \cdot \boldsymbol{V} \right) \tag{5.6}$$

where E and p are the initial energy and momentum of the particle. A specific model must be used to evaluate equation (5.6). A general result is that Fermi acceleration can be described by the Fokker–Planck equation:

$$\frac{\partial N(E)}{\partial t} = \frac{1}{\tau_A} \left\{ -\frac{\partial}{\partial E} \left(\frac{v}{c} E N(E) \right) + \frac{1}{4} \frac{\partial^2}{\partial E^2} \left[\left(\frac{v}{c} \right)^3 E^2 N(E) \right] \right\} \tag{5.7}$$

where $N(E)$ is the number of particles with energy E, τ_A is the acceleration time and v

is the particle velocity. A typical model-dependent value of τ_A for bounces off a magnetic field with a mean spacing of L is (Fermi 1949)

$$\tau_A = 3Lc/4V^2. \tag{5.8}$$

This time is far too long for large L, e.g. with $V = 100$ km s^{-1} and $L = 10^9$ cm, $\tau_A \simeq 2 \cdot 3 \times 10^5$ s $\simeq 2$ d.

The way to decrease τ_A is to decrease L and increase V as discussed in Sturrock *et al* (1978, chap 4). This has been done by Kulsrud and Ferrara (1971) using hydromagnetic waves which have frequency ω and wavevector k. They noted that equation (5.7) could be written in the form of an isotropic diffusion equation in momentum space, i.e.

$$\frac{\partial f(p)}{\partial t} = \frac{1}{p^2} \frac{\partial}{\partial p} \left(p^2 D(p) \frac{\partial f(p)}{\partial p} \right) \tag{5.9}$$

where $f(p)$ is the particle distribution function and $D(p)$ is the diffusion coefficient. Kulsrud and Ferrara (1971) calculated $D(p)$ using quasi-linear theory and found

$$D(p) = p^2 \int \frac{d^3 k \, d\omega}{(2\pi)^4} \, \gamma(k, \omega) \frac{\langle \delta B(k, \omega) \, \delta B^*(k, \omega) \rangle}{B_0^2} \tag{5.10}$$

where $\gamma(k, \omega)$ is the acceleration rate which depends explicitly on the assumed scattering rate ν, B_0 is the magnitude of the ambient magnetic field and $\delta B(k, \omega)$ is the Fourier transform of the magnetic fluctuations associated with the turbulence. Angular brackets here denote an ensemble average. Melrose (1974) proposed that the scattering rate ν can and does adjust itself to the value which corresponds to Fermi acceleration in Kulsrud and Ferrara's theory. Basically the proposition is that the turbulence causes the particles to become anisotropic, and anisotropic particles in turn generate the resonant waves which scatter them as discussed in the following paragraph. The resulting acceleration rate for $\omega/k = v_A$ is

$$\gamma(k, \omega) = \frac{\pi}{2} \omega \frac{v_A}{v} \tag{5.11}$$

which is independent of ν and essentially the same as the one occurring in the early theories of Fermi acceleration referenced above. The equivalence of equations (5.8) and (5.11) can be shown by equating $\omega/k_\parallel = v_A$ to the speed V and $\lambda = 2\pi/k_\parallel$ to the length L.

Although ν does not appear in equation (5.11), this equation is valid only if the scattering rate lies in the range

$$\omega \, (v_A/v) \, (\delta B/B)^2 \ll \nu \ll \omega \, (v/v_A). \tag{5.12}$$

This resonance scattering (Melrose 1974) is possible only when the gyroradius of the particle is comparable with the wavelength of the resonant wave and the speed of the particle is much greater than the phase speed of the wave (so that the wave looks like a stationary spatially periodic magnetic fluctuation to the particle). These conditions can be satisfied only for

$$v \gg v_A \qquad \text{for non-relativistic ions}$$

and $$\tag{5.13}$$

$$v \gg 43 \, v_A \quad \text{for non-relativistic electrons}$$

the latter being able to resonate only with whistler waves. The particles must become sufficiently anisotropic to generate the waves which scatter them which requires

$\delta B/B_0 \lesssim v/v_{th}$, where v_{th} is the threshold speed given by changing (5.13) into equalities. Thus for $\delta B/B_0 \ll 1$, we require $v \gg v_A$ for ions and $v \gg 43 \, v_A$ for electrons before the scattering, and hence the acceleration, can be effective. Another requirement is that the resonant waves grow sufficiently fast which places a lower limit on the density of accelerated particles (Melrose 1974).

When these requirements are satisfied, the rate of Fermi acceleration is roughly

$$\gamma = \omega(v_A/v) \, (\delta B/B_0)^2 \qquad (5.14)$$

where $\omega = k \, v_A$ is the typical frequency of the hydromagnetic turbulence. For $v_A \sim 10^3$ km s^{-1} corresponding to $B_0 = 10$ G and $n = 4 \cdot 8 \times 10^8$ cm^{-3}, and $v \sim c$, an acceleration time of less than 10^2s requires $(\delta B/B_0)^2 T \gtrsim 1$, where T is the typical period of the turbulence in seconds. The threshold criterion (5.13) in this case corresponds to electrons and protons of energy greater than 5 keV. There is radio evidence for large amplitude turbulence with periods between a few tenths of a second and a few seconds (McLean *et al* 1971, Gotwols 1972) which could account for second phase acceleration provided it contains sufficient power.

5.3. Radio bursts as a diagnostic of plasma processes

As noted in §2.4, there are only a few types of radio bursts with sufficiently unique and complete observations to allow construction of self-consistent theories which provide plasma parameters from the bursts. We limit the discussion here to type III bursts and moving type IV bursts of the plasmoid type. Type III bursts occur near the fundamental and second harmonic of the plasma frequency $\omega_{pe} \propto n_e^{1/2}$ and thus, to the extent that one can distinguish between fundamental and harmonic emission, provide an immediate determination of the electron density. The theory of type III bursts has been reviewed by Smith (1970, 1974a), Melrose (1980) and Smith and Nicholson (1979). They occur due to a two-stage process. In the first stage an electron stream excites electron plasma waves which satisfy the dispersion relation

$$\omega_p{}^2 = \omega_{pe}{}^2 + 3v_e{}^2 \, k^2 \qquad (5.15)$$

where ω_p and k are the frequency and wavenumber of the wave, respectively. In the second stage these plasma waves are converted into radiation by two non-linear processes. The first process which produces radiation near the fundamental is

$$p + i \rightarrow t + i' \qquad (5.16)$$

where p is the plasma wave, t is the transverse electromagnetic wave and i and i' are the polarisation cloud of an ion before and after conversion. This process is known as scattering on the polarisation cloud of an ion. The second process which produces radiation near the second harmonic is

$$p + p' \rightarrow t \, (2\omega_p). \qquad (5.17)$$

The radiation from both processes can be either spontaneous or induced. Spontaneous radiation near the fundamental has a brightness temperature $T_B \lesssim 10^9$ K and thus to explain the observation of T_B as high as 10^{11} K the radiation must be amplified. The absorption coefficient for process (5.16) can be negative, but a complicated transfer problem must be solved to determine the resulting intensity including the effects of density inhomogeneities in the background plasma (Smith and Riddle 1975). Such studies show that the only way the amplification can be sufficiently effective is for

intense plasma waves to occur in clumps of 100 km characteristic size at the 80 MHz plasma level. This conclusion was also reached by Melrose (1980).

The mechanism for clumping the plasma waves is presently unknown, but such a spatially inhomogeneous distribution of plasma waves has been directly observed at 0·45 AU (Gurnett and Anderson 1976). At this distance from the Sun, however, the radiation is near the second harmonic and produced by process (5.17). For this process the absorption coefficient is always negative so that amplification is impossible. Spontaneous radiation by process (5.17) can only lead to sufficiently high T_B if the plasma waves p and p' meet almost head on and both have high effective temperatures. The only plasma waves which have high effective temperature are those excited by the electron stream and they are initially confined to a fairly narrow cone about the stream direction (Magelssen and Smith 1977). Thus some of these waves must be scattered into a cone opposite to the direction of the stream (the so-called backward cone) to obtain the required brightness temperature. The process for this scattering is

$$p + i \rightarrow p' + i' \tag{5.18}$$

which is the same as process (5.16) except that a plasma wave is scattered into another plasma wave rather than a transverse wave. As shown by Smith (1977d), this process is sufficiently effective for the plasma wave intensities measured by Gurnett and Anderson (1976) to produce a quasi-isotropic distribution of plasma waves and to explain the observed radiation via process (5.17). Consequently type III radio bursts provide a direct test of some non-linear plasma processes.

Type III bursts also provide a natural laboratory for study of the interaction of an electron stream with a plasma. However, realistic theories for this interaction have only been worked out in the one-dimensional approximation (Magelssen and Smith 1977). At 1 AU electron streams associated with type III bursts are observed to have only a 2:1 anisotropy in the anti-solar direction. There are no simultaneous observations of the angular distribution of plasma waves, apparently because of their clumping property discussed above, but it is expected on theoretical grounds (Smith 1977d) that this distribution should also be quite broad. Future work must therefore take the actual two-dimensional nature of these distributions into account. The one-dimensional results (Magelssen and Smith 1977) indicated that the stream should have no problem in travelling from the Sun to the Earth for stream densities consistent with observations. Although a large level of plasma waves is produced by the front of the stream, these are largely reabsorbed by the rear of the stream so that the stream suffers small energy loss at any given point, which is consistent with observations. However, we cannot tell at present to what extent this effect will persist when two-dimensional effects are taken into account.

Moving type IV bursts of the plasmoid type (see §2.4) provide another example of a burst where the theory is sufficiently developed to allow determination of the plasma properties. The burst polarisation increases with time from $\sim 10\%$ to 70% or higher and is in the extraordinary mode. Such high polarisation can only be consistent with gyrosynchrotron emission from mildly relativistic electrons ($E \gtrsim 100$ keV). Detailed numerical calculations of this emission have been made by Dulk (1973) which indicate that the internal magnetic field of the plasmoid must be of the order of 10 G at a height $\sim 2R_\odot$ to explain the high degree of polarisation. One of the interesting inferences from this theory comes from the observation that the whole source is often polarised in the same sense. The sense of polarisation depends on the sign of $\cos\theta$, where θ is the angle between the line of sight and the direction of the magnetic field.

Thus it should be opposite for regions of the source with opposite signs of the components of the magnetic field along the line of sight. In a plasmoid the self-contained magnetic field must have approximately equal regions with oppositely directed fields to satisfy the Maxwell equation $\nabla . \boldsymbol{B} = 0$. Thus one would except approximately equal regions of opposite polarisation unless the field lines directed towards the observer, say, were much more concentrated than the field lines directed away from the observer. Since the power radiated per electron varies as B^2, the emission from the stronger field regions would then dominate. The fact that the whole source is often polarised in the same sense suggests that this must often be the case.

Thus we have two examples of bursts where the theory is sufficiently well developed to provide information about the states of the plasma and the plasma processes occurring in these bursts. Of these the type III example is more important since the relatively weak electron stream only slightly perturbs the ambient coronal plasma through which it propagates. With the moving type IV plasmoid, on the other hand, we only gain information on a rather exotic and rare piece of coronal plasma, the plasmoid itself.

6. Conclusions and prognosis

We return finally to an attempt to answer the question posed in §1 as to how much more of the flare problem we have managed to understand in the past decade or so. As far as the central problem of obtaining a high enough dissipation rate in the flash phase is concerned, it now seems clear that this is only possible (even in principle) either by field *reconnection* in a Petschek (1964) type situation or by dynamic field *annihilation* in many magnetic islands generated by the tearing mode instability within a sheared magnetic geometry, such as in the arch model envisaged by Spicer (1976). The fact that most flares exhibit energy release in the form of hot kernels without necessarily involving strong mass motions seems strongly to favour the latter alternative, as does the predominant appearance of loops in the XUV. Furthermore the richness of theoretical possibilities intrinsic to the tearing mode/arch geometry model offers an encouraging interpretation of the great diversity of ways in which different flares manifest themselves. On the other hand, no complete description of the arch model is yet available in terms of the evolution of the current carrying arch toward the unstable onset of the flare.

The problems of particle acceleration are in an even less satisfactory state. The theory of first phase acceleration is a complicated sequence whose details have only been sketched in and are badly in need of quantitative calculation. While direct electric-field acceleration in quiet steady two-dimensional current sheets has been worked out, the problem of such acceleration in turbulent dynamic three-dimensional current sheets has hardly been touched. The efflux of heat and particles from these regions into the surrounding plasma and their subsequent interactions there are problems which must really await solution of the first problem. About all we can say is that we know of an efficient redistribution process by Langmuir plasma waves which can lead to observed electron distributions if these are not already produced in the initial acceleration. Numerical simulations of the processes in current sheets still suffer from the unsatisfactory restriction of two dimensions. On the diagnostic side, the recent emergence of reverse currents as a possible controlling feature in the energy losses of accelerated electron streams has thrown considerable uncertainty even into the number

of accelerated electrons required, i.e. into the extent to which the electron distribution is non-thermal. Second-stage acceleration presents less difficulties of principle, as has been known for a long time, and hinges mainly on receiving the necessary input from the first stage.

Probably the most productive area in flare theory recently has been that of modelling of energy redistribution in the flaring atmosphere. This comparative success is in part attributable to the novelty of this area of flare research and its stimulation by the greatly improved flare atmosphere diagnostics made available by recent observational techniques. It is also partly due, however, to the *relative* simplicity of the problems tackled which consist essentially of modelling the response of the atmosphere, to *prescribed* inputs, in terms of radiative and hydrodynamic/magnetohydrodynamic processes, the physics of which are well understood. On the other hand, while this approach has yielded useful new information on the spatial and temporal distribution of energy deposition in the flaring atmosphere, it is limited insofar as a variety of input mechanisms may be compatible with these distributions.

It seems reasonable to conclude that a good deal of progress has been made observationally, theoretically and in the laboratory toward our understanding of flares but that, in the key problem areas of dissipation rate and particle acceleration, either no adequate theory yet exists or no adequate tests to distinguish rival theories is available. We finally therefore give a little attention to discussing possible reasons for the continuing lack of success in these respects, despite the effort expended, and to our prognosis of possible progress in the subject. Aside from the interest of such questions to the solar physics community, they present a sobering object lesson to astrophysicists generally on the difficulty of understanding an astrophysical phenomenon when one is allowed to observe it in detail.

Among the many reasons for our failure to solve the flare problem, we would regard the following as particularly relevant and instructive.

(a) *Plasma physics is complicated.* This remark is not intended trivially but to emphasise that we should not be too surprised that the behaviour of a magnetised solar plasma is hard to understand, nor too simplistic in our attempts to interpret our observations. As a particular solution of a set of some fifteen or so coupled non-linear partial differential equations together with atomic and plasma coefficients, a solar flare may have many facets hardly amenable to being readily parametrically modelled. Thus, while simple modelling procedures may be desirable and instructive, the principle of simplicity is not necessarily recognisable in a macroscopic system but only in the physical laws governing it on a suitably localised scale.

(b) *The flare process itself destroys much of the necessary information.* In common with all instabilities, the solar flare is an entropy increasing process. Herein may lie an explanation of the point already mentioned, that modelling of observations of the thermalisation of flare energy permits considerable ambiguity in the primary dissipation mechanism. Once energy is thermalised, all detailed information on where it came from is already lost. This intrinsic difficulty can only be overcome either by looking for subtle diagnostic clues still present in the aftermath or by looking at the unstable device just before it goes unstable. Either way, in the case of a flare, will require the use of instruments of comparatively high sensitivity to detect the small fluxes involved.

(c) *Characteristic scales may lie below instrumental resolution.* Despite the steady progress of instrumental resolution in space, time, wavelength and polarisation, it is essential to recognise that certain key features of flare mechanisms may occur on such a

small size or time scale that we have virtually no prospect of observing them directly. Two striking examples are the thickness of a neutral sheet, which might be as small as 10^{-6} arcsec as seen from the Earth, and the plasma period which characterises some acceleration mechanisms and which scarcely exceeds a nanosecond anywhere in the solar atmosphere. Nor is the problem only instrumental since, for example, the sharpness of temperature and density discontinuities, which could indicate the thermal conductivity across a shock, is spatially smeared because of the finite range of temperatures over which any spectrum line forms. In Spicer's (1976) model the *total* volume dominated by the tearing mode with classical conductivity is about 10^{23} cm^3 ($\simeq 0.5$ arcsec on a side) while that with anomalous resistivity is only 10^{20} cm^3 ($\simeq 0.05$ arcsec) or, respectively, 1% and 10^{-3}% of the minimum volume resolved by Skylab. (The volume of each separate energy release site is much smaller still.) Evidently a very advanced instrument indeed would be needed to check the model by measurement of the different temperatures and densities characterising these volumes. On the theoretical side a serious effort is needed to predict model features which might be distinguished in practice, such as the energy distribution over different readily observable bulk flare products, e.g. thermal and mass motion energies.

If we really wish directly to confirm the occurrence of the tearing mode instability as against any proposed alternative then it may not be sufficient to make observations of the overall geometry of the dissipation region since, for example, Spicer's tearing mode results could apply within any sufficiently sheared field geometry and are in no way uniquely associated with an arch. Thus one might have to envisage resolution of high-temperature plasmas of better than 10^{-2} arcsec.

(*d*) *Lack of flash phase observations.* As we have maintained throughout this review, the problem of flash phase energy release is the essence of flare physics. It is thus of paramount importance that observations be made in this phase with adequate spatial and temporal resolution in the right place and at the right time. The brevity of the flash phase rise (typically 1–2 min) means that this is no small request. The instruments which have systematically achieved flash phase coverage are those which cover the whole Sun almost continuously and whose detection thresholds are only exceeded during flares, such as hard x-rays and γ-ray detectors which currently lack spatial resolution.

With spatial resolution there is generally a scan time restriction on time resolution, or a restriction of the field of view. The latter option raises the problem of choice of the region of observation which underlay the resounding failure of the Skylab package to observe flare flash phases except in a few lucky cases. Given the uncertainty of our knowledge of flare precursors, it is clear that the flash phase duration will be shorter than the time taken to slew an instrument onto the relevant region or to change it over to flare mode and the only answer is a committed flare instrument package directed continuously at a preselected flare active region.

(*e*) *Lack of observational coordination.* In the flare problem and astrophysics generally, we are primarily concerned with the testing of different models against data, and we have already discussed in §2.1.3 by means of a particular example how great ambiguity can arise in such problems. This may alternatively be expressed as a lack of linear independence of the information vectors over the limited observational range measured (Craig and Brown 1976a), allowing the immediate generalisation of considering observations made at widely differing wavelengths (or with widely differing mechanisms supplying the photons, etc) as being virtually orthogonal information vectors. The information yield of a set of n such *independent* observations clearly

increases with n. However, suppose that each of the observations alone can distinguish between m aspects of the flare model hypotheses. Then, if the observations are made in an uncoordinated way because they are not synchronised in space or time (or because data are not exchanged) the information yield rises like the product mn. If the observations are properly coordinated, however, the yield increases as m^n. Even with $m = 3$, $n = 6$ (one each radio, optical, uv, soft x-ray, hard x-ray and γ-ray instruments) the yields are in the ratio $1/40$ (cf the detailed joint study by Crannell *et al* (1978) of hard x-ray and microwave bursts). An example of the loss incurred by the absence of data coordination is the failure to provide adequate hard x-ray coverage during the Skylab period and consequent difficulty in testing electron heating models for the optical and xuv flares (cf Canfield *et al* 1978). This problem can only be overcome by ensuring: availability of adequate instrumental coverage; sufficient prior planning of coordinated observing modes to answer specific scientific questions; an efficient communications network; that flare theorists and modellers establish the relevant scientific questions in observationally answerable form. The lack of one or more of these essential ingredients is, in our view, the source of disappointing results yielded by many past efforts at international cooperative programmes.

To close we would like to indicate the extent to which present research efforts and future instrumentation are directed at overcoming the more tractable of the difficulties listed above. It is possible that the next few years will see considerable progress in these regards due primarily to NASA's dedication of their Solar Maximum Mission entirely to flare and active region studies. This satellite, due for launch in February 1980, carries a package of instruments spanning most of the range of energetic emissions from the uv through to γ-rays with the highest resolution in space, time and spectrum possible within the technical limitations of the satellite. The originally intended package is summarised in table 1, and includes the first ever imaging facility in the hard x-ray range (this will be augmented by high spatial resolution now available from large microwave arrays), the first uv polarimeter (capable of determining magnetic fields at higher altitudes in the solar atmosphere than any optical polarimeter) and improved resolution in all wavelengths. Unfortunately a large slice of the xuv spectral range has already been lost by the cancellation of the spectroheliometer but may be partly compensated by Soviet instrumentation in orbit at the same time. Likewise, further space coverage by particle detectors and long radio wavelength antennae may be available from separate spacecraft.

Much more effort than previously has also been put into scientific planning of observing requirements. This has taken several forms.

(i) NASA's move toward problem-oriented missions may be seen in the many working sessions which have already been held by the Investigator's Working Group on scientific programmes for SMM itself.

(ii) Experience of post-mission workshops on Skylab and OSO-8 has yielded valuable experience for SMM, and the post-SMM workshops will presumably be correspondingly more productive. This should be further expedited by the pre-planning of the Guest Investigator Program and the policy of making SMM data generally available relatively soon after their acquisition.

(iii) The problem of international (especially ground-based instrument) coordination has been taken up by the independent COSPAR body Solar Maximum Year (SMY). Of course the existence of a committee structure behind such an activity in no way guarantees its scientific effectiveness though it greatly expedites organisation of meetings, establishing of the communications network, etc. However, the subsections

13

Table 1. Package of instruments originally proposed for NASA's Solar Maximum Mission due for launch in February 1980.

Experiment	Spectral range	Maximum spectral resolution	Field of view	Maximum angular resolution	Maximum time resolution
γ-ray spectrometer	0·3–9 MeV	7·5% at 0·66 MeV	Full Sun	Full Sun	1 s
Hard x-ray spectrometer	20–300 keV	16 channels	Full Sun	Full Sun	10^{-3}–10^{-1} s
Hard x-ray imaging spectrometer	3·5–30 keV	6 channels	4·5′×4·5′	8″×8″	0·5 s
Soft x-ray polychromator	1·4–22·4 Å	$\lambda/\Delta\lambda \simeq 25\ 000$–1000	6′×6′	10″×10″	0·1 s
XUV spectroheliometer	20–716 Å and 929–1336 Å	0·1 Å		4″×4″	
UV spectrometer and polarimeter	1100–3000 Å	0·02 Å($\pm$50 G)	4′×4′	4″×4″	64×10^{-3}s (per pixel)
White light coronagraph	4435–6583 Å	5–200 Å	(1·5–6·0) $R_\odot$ from solar centre	6″–12″	50 s

of SMY are being scientifically organised in close collaboration with SMM and on the essential problem-oriented approach to detailed formulation of observing programmes. These subsections are: flare build-up study (FBS) (see 1975 *Solar Physics* **47** No 1) dedicated to the evolution of active regions prior to flares, which may help obviate difficulty (*b*) above; study of energy release in flares (SERF) which is specifically oriented to the scientific and organisational problems of the flare flash and decay phases, i.e. to (*d*) and (*e*) above; study of travelling interplanetary phenomena (STIP) which is devoted to the interplanetary flare aftermath and which has already organised several observational alert periods for coordination of spacecraft.

(iv) Provision of the necessary theoretical and diagnostic developments in parallel with obtaining the observations is also underway, at least in some areas, through pre-mission workshops and the Guest Investigator Program. A great deal rests on these activities in that, given the effort and resources involved around SMM, the credibility of solar physics research in space can scarcely be enhanced by anything other than a highly positive outcome of the exercise. Clearly in a competitive funding situation, such factors must affect the future of a subject particularly in countries with a substantial aerospace commitment and a wide range of astrophysical interests. However, there seems to us every reason to expect a very successful outcome from SMM. Looking beyond SMM, there still remains quite enormous scope for development of flare studies on all fronts. Here we can only briefly mention a few of the more striking possibilities, and must entirely neglect peripheral areas such as the subject of solar/weather interrelations currently in vogue.

Probably the most valuable foreseeable ground-based development will be the introduction of vector magnetographs with suitably high time and space resolution, and the further development of large microwave interferometers. A detailed survey of the possibilities for solar space missions after SMM (both free-flying and Shuttle payloads) has been made in a report to NASA by Sturrock *et al* (1977). The chief advantage of the Shuttle will be its capacity to orbit much larger instruments than are involved in SMM, such as a proposal for a large multiple collimator device sensitive enough for arcsecond resolution of hard x-rays up to 100 keV (cf table 1). Other projects may include the large Solar Telescope (ST) and the Grazing Incidence Solar Telescope (GRIST) in the XUV, the latter being a European venture. Further developments could include greatly improved polarimetry in the UV (see above) and in hard x-rays (Tindo and Somov 1977), the latter having a direct bearing on the energetic electron velocity distribution (cf §5).

The most imminent free-flying mission after SMM is the Out-of Ecliptic (OOE) Mission, which will allow the first direct view of solar structure at high latitudes and the possibility of stereo views in conjunction with a second spacecraft. Separate (ecliptic) Stereo Missions would also permit the first direct measurements of directivity in flare radio and x-ray emission, and improved possibilities for vector field measurements. Among the most exotic proposals are the Solar Pinhole Mission and the Solar Plunger (or Solar Probe). The former concept involves orbiting a very large pinhole device remote from a position-sensitive hard x-ray and γ-ray detector intended to give very high spatial resolution (down to 10^{-2} arcsec, cf difficulty (*b*) above) at these wavelengths in combination with high sensitivity. The latter project involves the use of a Jupiter flyby or solar sail propulsion to send a probe directly into the Sun, making *in situ* measurements for as long as it survives (at least down to four solar radii) and so removing the line-of-sight projection problem which dogs all astrophysical observation.

In short, the remainder of this century promises to be a fascinating time of development of solar flare observations, providing excellent opportunities to test current developments in flare plasma physics.

References

Alfven H and Carlquist P 1967 *Solar Phys.* **1** 220
Antiochos S and Sturrock P A 1978 *Astrophys. J.* **220** 1137
Anzer U 1968 *Solar Phys.* **3** 298
Athay R G (ed) 1974 *Chromospheric Fine Structure, IAU Symp. 56* (Dordrecht: Reidel)
Bai T and Ramaty R 1978 *Astrophys. J.* **219** 705
—— 1979 *Astrophys. J.* **227** 1072
van Beek H F, de Feiter L D and de Jager C 1973 *Space Res.* **15**
Benz A O 1977 *Astrophys.J.* **211** 270
Benz A O and Gold T 1971 *Solar Phys.* **21** 157
Biskamp D and Chodura R 1973 *Phys. Fluids* **16** 893
Biswas S and Radhakrishnan B 1973 *Solar Phys.* **28** 211
Bloomberg H W, Davis J and Boris J P 1977 *J. Quant. Spectrosc. Radiat. Transfer* **18** 237
Boris J P, Dawson J M, Orens M H and Roberts K V 1970 *Phys. Rev. Lett.* **25** 706
Brown J C 1971 *Solar Phys.* **18** 489
—— 1972a *Solar Phys.* **25** 158
—— 1972b *Solar Phys.* **26** 441
—— 1973a *Solar Phys.* **31** 143
—— 1973b *Solar Phys.* **29** 421
—— 1974 *Coronal Disturbances, IAU Symp. 57* ed G A Newkirk (Dordrecht: Reidel) p395
—— 1975 *Solar γ- X- and EUV-Radiation, IAU Symp. 68* ed S R Kane (Dordrecht: Reidel) p245
—— 1976 *Phil. Trans. R. Soc.* **281** 473
—— 1977 *Solar Phys.* **53** 263
—— 1978 *Astrophys. J.* **225** 1076
Brown J C, Canfield R C and Robertson M N 1978 *Solar Phys.* **57** 399
Brown J C, Craig I J D and Karpen J T 1980 *Solar Phys.* in press.
Brown J C and Hoyng P 1975 *Astrophys. J.* **200** 734
Brown J C and McClymont A N 1976 *Solar Phys.* **49** 329
Brown J C and Melrose D B 1977 *Solar Phys.* **52** 117
Brown J C, Melrose D B and Spicer D S 1979 *Astrophys. J.* **228** 592
Brown J C and Nakagawa Y 1978 *Astrophys. J. Lett.* **225** L153
Brueckner G E 1976 *Phil. Trans Soc.* A **281** 443
Canfield R C 1974 *Solar Phys.* **34** 339
Canfield R C and Cook J W 1978 *Astrophys. J.* **225** 650
Canfield R C *et al* 1978 *Report of NASA Skylab Workshop on Solar Flares* by P A Sturrock *et al* Team 5 Report
Carmichael H 1964 *AAS-NASA Symp. on Physics of Solar Flares* ed W N Hess (Washington, DC: NASA) p451
Carrington R C 1859 *Mon. Not. R. Astron. Soc.* **20** 13
Cartwright B G and Mogro-Campero A 1973 *High Energy Phenomena on the Sun* ed R Ramaty and R G Stone (Greenbelt: NASA) p393
Cheng C C 1971 *Solar Phys.* **23** 178
Cheng C C and Widing K G 1975 *Astrophys. J.* **201** 735
Chubb T A 1971 *Proc. Leningrad Symp. on Solar–Terrestrial Physics* ed C de Jager, vol 1 (Dordecht: Reidel) p99
Chupp E L, Forrest D J, Highie P R, Suri A N, Tsai C and Dunphy P P 1973 *Nature* **241** 5388
Colgate S A A 1978 *Astrophys. J.* **221** 1068
Colgate S A A, Adouze J and Fowler W A 1978 *Astrophys. J.* **213** 849
Cook J W and Brueckner G E 1979 *Astrophys. J.* **227** 645
Coppi B and Friedland A B 1971 *Astrophys. J.* **169** 379
Coroniti F V and Eviatar A 1977 *Astrophys. J. Suppl.* **33** 189
Cox D P and Tucker W H 1969 *Astrophys. J.* **157** 1157

Craig I J D 1978 *Astron. Astrophys.* **61** 575
Craig I J D and Brown J C 1976a *Astron. Astrophys.* **49** 239
—— 1976b *Nature* **264** 340
Craig I J D and McClymont A N 1976 *Solar Phys.* **50** 133
Craig I J D, McClymont A N and Underwood J H 1978 *Astron. Astrophys.* **70** 1
Crannell C J, Frost K J, Mätzler C, Ohki K and Saba J L 1978 *Astrophys. J.* **223** 620
Crawford H J, Price P B, Cartwright B G and Sullivan J D 1975 *Astrophys. J.* **195** 213
Culhane J L, Vesecky J F and Phillips K J H 1970 *Solar Phys.* **15** 396
Datlowe D W, Elcan M J and Hudson H S 1974a *Solar Phys.* **39** 155
Datlowe D W, Hudson H S and Peterson L F 1974b *Solar Phys.* **35** 193
Datlowe D W and Lin R P 1973 *Solar Phys.* **32** 459
Datlowe D W, O'Dell S L, Peterson L E and Elcan M J 1977 *Astrophys. J.* **212** 561
Davis J, Kepple P C and Strickland D J 1977 *J. Quant. Spectrosc. Radiat. Transfer* **17** 711
Davis J and Rogerson J E 1977 *Solar Phys.* **51** 185
Davis L K 1956 *Phys. Rev.* **101** 351
Davis W D 1977 *Solar Phys.* **54** 139
Dere K P 1978 *Astron. Astrophys.* **70** 439
Dere K P, Horan D M and Kreplin R W 1974 *Solar Phys.* **36** 459
—— 1977 *Astrophys. J.* **217** 976
Dizer M 1969 *Solar Phys.* **10** 416
Donnelly R F 1970 *ESSA Tech. Rep.* ERL 169–5DL14
Donnelly R F and Kane S R 1978 *Astrophys. J.* **222** 1043
Drake J F and Lee Y C 1977 *Phys. Fluids* **20** 1341
Dubov F E 1963 *Sov. Phys. Dokl.* **8** 543
Dulk G A 1973 *Solar Phys.* **32** 491
Dungey J W 1953 *Phil. Mag.* **44** 725
Eastwood J W 1972 *Planet. Space Sci.* **20** 1555
—— 1975 *Planet. Space Sci.* **23** 1
Elgaroy F O 1977 *Solar Noise Storms* (Oxford: Pergamon)
Elwert G and Haug E 1970 *Solar Phys.* **15** 234
—— 1971 *Solar Phys.* **20** 413
Emslie A G 1979 *Astrophys. J.* submitted
Emslie A G, Brown J C and Donnelly R F 1978 *Solar Phys.* **57** 175
Emslie A G and Noyes R W 1978 *Solar Phys.* **57** 373
Fainberg J, Evans L G and Stone R G 1972 *Science* **178** 743
Falciani R, Landini M, Righini A and Rigutti M 1968 *Structure and Development of Solar Active Regions, IAU Symp. 35* ed K O Keipenheuer (Dordrecht: Reidel) p451
de Feiter L D 1971 *Solar Phys.* **19** 207
Fermi E 1949 *Phys. Rev.* **75** 1169
Fisk L H and Axford W I 1969 *Solar Phys.* **7** 486
Forbush S E 1946 *Phys. Rev.* **70** 771
Freier P S and Webber W R 1963 *J. Geophys. Res.* **68** 1605
Friedman M 1969 *Phys. Rev.* **182** 1408
Frost K J 1969 *Astrophys. J. Lett.* **158** L159
Frost K J and Dennis B R 1971 *Astrophys J.* **165** 655
Frost K J, Dennis B R and Lencho R J 1971 *New Techniques in Space Astronomy* ed N R Labrum and R Lüst (Dordrecht: Reidel) p185
Furth H P, Killeen J and Rosenbluth M N 1963 *Phys. Fluids* **6** 459
Furth H P, Rutherford P H and Selberg H 1973 *Phys. Fluids* **16** 1054
Gabriel A H and Jordan C 1969 *Mon. Not. R. Astron. Soc.* **145** 241
Giovanelli R G 1946 *Nature* **158** 81
Gold T and Hoyle F 1960 *Mon. Not. R. Astron. Soc.* **120** 89
Gotwols B L 1972 *Solar Phys.* **24** 232
Gurnett D A and Anderson R R 1976 *Science* **194** 1159
Guseinov R B 1973 *Solar Phys.* **31** 401
Hamburger S M and Jancarik J 1972 *Phys. Fluids* **15** 825
Hasegawa A 1975 *Plasma Instabilities and Nonlinear Effects* (Berlin: Springer-Verlag).
Haug E 1972 *Solar Phys.* **25** 425
—— 1976 *Solar Phys.* **45** 453

Henoux J C and Nakagawa Y 1977 *Astron. Astrophys.* **57** 105
—— 1978 *Astron. Astrophys.* **66** 385
Heyvaerts J, Priest E R and Rust D M 1977 *Astrophys. J.* **216** 123
Hirosa A and Skarsgard H M 1976 *Phys. Rev. Lett.* **36** 252
Holt S S and Cline T L 1968 *Astrophys. J.* **154** 1027
Holt S S and Ramaty R 1969 *Solar Phys.* **8** 119
van Hoven G and Cross M A 1973 *Phys. Rev.* A **7** 1347
Howard R (ed) 1971 *Small Scale Solar Magnetic Fields, IAU Symp. 43* (Dordrecht: Reidel)
Hoyng P 1977a *Astron. Astrophys.* **55** 23
—— 1977b *Astron. Astrophys.* **55** 31
Hoyng P, Brown J C and van Beek H F 1976 *Solar Phys.* **48** 197
Hoyng P, Knight J W and Spicer D S 1978 *Solar Phys.* **58** 139
Hoyng P and Melrose D B 1977 *Astrophys. J.* **218** 866
Hudson H S 1972 *Solar Phys.* **24** 414
—— 1973 *Symp. on High Energy Phenomena on the Sun, NASA GSFC Preprint* X-693-63-193, ed R Ramaty and G E Stone
—— 1978 *Astrophys. J.* **224** 235
Hundhausen A J 1972 *Coronal Expansion and Solar Wind* (Berlin: Springer-Verlag)
Hutchinson I H 1976 *Phys. Rev. Lett.* **37** 338
de Jager C 1968 *Mass Motions in Solar Flares and Related Phenomena* ed Y Öhman (New York: Wiley) p171
Jaggi R K 1964 *AAS-NASA Symp. on Physics of Solar Flares* ed W M Hess (Washington, DC: NASA) p419
Jefferies J T and Orrall F Q 1965 *Astrophys. J.* **141** 519
Jokipii J R 1971 *Rev. Geophys. Space Sci.* **9** 27
Kadomtsev V V 1966 *Rev. Plasma Phys.* **2** 153
Kahler S W 1971a *Astrophys. J.* **164** 365
—— 1971b *Astrophys. J.* **168** 319
—— 1975 *Solar γ-, X-, and EUV Radiation, IAU Symp. 68* ed S R Kane (Dordrecht: Reidel) p211
Kahler S W, Krieger A S and Vaiana G S 1975 *Astrophys. J. Lett.* **199** L57
Kane S R 1974 *Coronal Disturbances, IAU Symp. 57* ed G A Newkirk Jr (Dordrecht: Reidel) p105
Kane S R (ed) 1975 *Solar γ-, X- and EUV Radiation, IAU Symp. 68* (Dordrecht: Reidel)
Kane S R and Anderson K A 1970 *Astrophys. J.* **162** 1003
Kane S R and Donnelly R F 1971 *Astrophys. J.* **164** 151
Kane S R *et al* 1978 *Report of NASA Skylab Workshop on Solar Flares* by P A Sturrock *et al* Team 4 Report
Kantrowitz A and Petschek H E 1966 *Plasma Physics in Theory and Application* ed W B Kunkel (New York: McGraw-Hill) p148
Kaplan S A, Pikelner S B and Tsytovich V N 1974 *Phys. Rep.* **15C** 1
Kaplan S A and Tsytovich V N 1973 *Plamsa Astrophysics* (Oxford: Pergamon)
Kaufmann P, Piazza I R and Raffaelli J C 1977 *Solar Phys.* **54** 179
Kennel C F, Lanzerotti L J and Parker E N (ed) 1978 *Solar Flares, Solar System and Plasma Physics* in press
Kennel C F and Petschek H E 1966 *J. Geophys. Res.* **71** 1
Knight J W and Sturrock P A 1977 *Astrophys. J.* **218** 306
Korchak A A 1971 *Solar Phys.* **18** 284
—— 1974 *Comm. Astrophys. Space Phys.* **6** 57
Korobeinikov V P 1969 *Solar Phys.* **7** 463
Kostjuk N D 1976a *Sov. Astron. Astrophys. J.* **19** 458
—— 1976b *Sov. Astron. Astrophys. J.* **20** 206
Kostjuk N D and Pikelner S B 1975 *Sov. Astron. Astrophys. J.* **18** 590
Krall N A and Trivelpiece A W 1973 *Principles of Plasma Physics* (New York: McGraw-Hill) p105
Kruskal M and Tuck J L 1958 *Proc. R. Soc.* A **245** 222
Kulsrud R M and Ferrara A 1971 *Astrophys. Space Sci.,* **12** 302
Kundu M R 1965 *Solar Radio Astronomy* (New York: Wiley Interscience)
Kuperus M (ed) 1977 *Report CECAM/Workshop on Solar Flares* (Paris: CECAM)

Landini M and Fossi B C 1975 *Astron. Astrophys.* **42** 213

Langer S H and Petrosian V 1977 *Astrophys. J.* **215** 666

Lanzerotti L J 1973 *High Energy Phenomena on the Sun* ed R. Ramaty and R G Stone (Greenbelt: NASA) p427

Lin R P 1974a *Space Sci. Rev.* **16** 189

—— 1974b *Coronal Disturbances, IAU Symp. 57* ed G A Newkirk (Dordrecht: Reidel) p201

Lin R P and Hudson H S 1976 *Solar Phys.* **50** 153

Lites B W and Cook J W 1979 *Astrophys. J.* **228** 598

McClymont A N 1976 *PhD Thesis* University of Glasgow

Machado M E, Emslie A G and Brown J C 1978 *Solar Phys.* **58** 363

Machado M E and Linsky J L 1975 *Solar Phys.* **42** 395

Machado M E and Noyes R W 1978 *Solar Phys.* **59** 129

McKenzie D C 1975 *Solar Phys.* **40** 183

McLean D J, Sheridan K V, Stewart R T and Wild J P 1971 *Nature* **234** 140

MacQueen R M, Gosling J T, Hildner E, Munro R H, Poland A I and Ross C L 1976 *Phil. Trans. R. Soc.* A **281** 295

McWhirter R W P, Thonemann P C and Wilson R 1975 *Astron. Astrophys.* **40** 63

Magelssen G R and Smith D F 1977 *Solar Phys.* **55** 211

Massey H, Sweet P A and Gabriel A H (ed) 1976 *A Discussion on the physics of the solar atmosphere. Phil. Trans. R. Soc.* **281**

Matsushita S 1962 *J. Geophys. Rev.* **67** 3753

Mätzler C 1976 *Solar Phys.* **49** 117

Mätzler C, Bai T, Crannell C J and Frost K J 1978 *Astrophys. J.* **223** 1058

Melrose D B 1974 *Solar Phys.* **37** 353

—— 1980 *Plasma Astrophysics* (New York: Gordon and Breach)

Melrose D B and Brown J C 1976 *Mon. Not. R. Astron. Soc.* **176** 15

Moore *et al* 1978 *Report of NASA Skylab Workshop on Solar Flares* by P A Sturrock *et al* Team 7 Report

Moreton G E 1964 *Astrophys. J.* **69** 145

Morrison P 1961 *Handb. Phys.* **46** 1

Mouschovious T C and Poland A I 1978 *Astrophys. J.* **220** 675

Munro R H, Depree A K and Withbroe G L 1971 *Solar Phys.* **19** 347

Najita K and Orrall F Q 1970 *Solar Phys.* **15** 176

Nakada M P, Neupert W M and Thomas R J 1974 *Solar Phys.* **37** 429

Nakagawa Y and Rust D M (ed) 1974 *Flare Related Magnetic Field Dynamics* (Boulder: High Altitude Observatory, Boulder)

Nakagawa Y, Wu S T and Han S M 1973 *Solar Phys.* **30** 111

Nakagawa Y, Wu S T and Tandberg-Hanssen E 1975 *Solar Phys.* **41** 387

Neupert W M, Thomas R J and Chapman R D 1975 *Solar Phys.* **34** 349

Newkirk G A Jr (ed) 1974 *Coronal Disturbances, IAU Symp. 57* (Dordrecht: Reidel)

Obashev S O 1966 *Sov. Astron. Astrophys. J.* **9** 784

Orrall F Q and Zirker J 1978 *Astrophys. J.* **208** 618

Parker E N 1963 *Astrophys. J. Suppl.* **8** 177

Parker E N and Tidman D A 1958 *Phys. Rev.* **111** 1206

Parkinson J H, Kestenbaum H L, Long K S and Novick R 1977 *Proc. of November 1977 OSO-8 Workshop* (Boulder: University of Colorado) p299

Parks G K and Winckler J R 1971 *Solar Phys.* **16** 186

Petrosian V 1973 *Astrophys. J.* **186** 291

Petschek H E 1964 *AAS-NASA Symp. on Physics of Solar Flares* ed W N Hess (Greenbelt: NASA) p425

Pikel'ner S B and Tsytovich V N 1976 *Sov. Astron. A J* **19** 450

Pneuman G W 1969 *Solar Phys.* **2** 462

Pottasch S R 1964 *Bull. Astron. Inst. Netherlands* **18** 7

Pustil'nik L A 1975 *Sov. Astron. A J* **19** 195

Quon B H and Wong A Y 1976 *Phys. Rev. Lett.* **37** 1393

Ramaty R 1969 *Astrophys. J.* **158** 753

Ramaty R, Kozlovsky B and Lingenfelter R E 1975 *Space Sci Rev.* **18** 341

Ramaty R and Stone G E (ed) 1973 *Symp. on High Energy Phenomena in the Sun* NASA SP-342

Riddle A C 1974 *Solar Phys.* **34** 181

Robinson R D 1974 *Proc. Astron. Soc. Aust.* **2** 258

Rosenberg H 1976 *Phil. Trans. R. Soc.* A **281** 461

Rossner R, Tucker W H and Vaiana G S 1978 *Astrophys. J.* **220** 643

Roy J-R and Datlowe D W 1975 *Solar Phys.* **40** 165

Rust D M 1976 *Phil. Trans R. Soc.* A **281** 427

Rust D M and Hegwer F 1975 *Solar Phys.* **40** 141

Rust D M *et al* 1978 *Report of NASA Skylab Workshop on Solar Flares* by P A Sturrock *et al*
     Team 6 Report

Rutherford P H 1973 *Phys. Fluids* **16** 1903

Santangelo N, Horstman H and Horstman-Moretti E 1973 *Solar Phys.* **39** 143

Schnack D and Killeen J 1978 *Theoretical and Computational Plasma Physics* (Vienna: IAEA)
     p337

Sedov L I 1959 *Similarity and Dimensional Methods in Mechanics* (New York: Academic)

Shafranov V D 1956 *Sov. J. Atom. Energy* **1** 709

Shapiro P and Knight J W 1978 *Astrophys. J.* **224** 1028

Shapiro P R and Moore R T 1977 *Astrophys. J.* **217** 621

Shmeleva O P and Syrovatskii S I 1973 *Solar Phys.* **33** 341

Simnett G M 1974 *Space Sci. Rev.* **16** 257

Smith D F 1970 *Adv. Astron. Astrophys.* **7** 147

—— 1974a *Space Sci. Rev.* **16** 91

—— 1974b *Coronal Disturbances, IAU Symp. 57* ed G A Newkirk (Dordrecht: Reidel) p253

—— 1976a *J. Geophys. Res.* **82** 704

—— 1976b *Astrophys. Space Sci.* **42** 261

—— 1977a *J. Geophys. Res.* **82** 704

—— 1977b *Astrophys. J.* **212** 291

—— 1977c *Astrophys. J.* **217** 644

—— 1977d *Astrophys. J.* **216** L53

Smith D F and Lilliequist C G 1979 *Astrophys. J.* **232** 582

Smith D F and Nicholson D R 1979 *Wave Instabilities in Space Plasmas* eds P J Palmadeso
     and K Papadopoulos (Dordrecht: Reidel) p225

Smith D F and Priest E R 1972 *Astrophys. J.* **176** 487

Smith D F and Riddle A C 1975 *Solar Phys.* **44** 471

Smith H J and Smith E P 1963 *Solar Flares* (New York: Macmillan)

Somov B V 1975 *Solar Phys.* **42** 235

Somov B V and Syrovatskii S I 1976 *Sov. Phys.–Usp.* **19** 813

Soward A M and Priest E R 1977 *Phil. Trans. R. Soc.* A **284** 369

Spicer D S 1976 *NRL Rep.* 8036

—— 1977 *Solar Phys.* **53** 305

Spicer D S and Davis J 1975 *Solar Phys.* **43** 107

Spitzer L 1962 *Physics of Fully Ionized Gases* (New York: Wiley Interscience)

Steinolfson R S, Wu S T, Dryer M and Tandberg-Hanssen E 1977 *Proc. Symp. on Study of
     Travelling Interplanetary Disturbances* ed M A Shea (Dordrecht: Reidel)

Stewart R T 1974 *Coronal Disturbances, IAU Symp. 57* ed G A Newkirk (Dordrecht: Reidel)
     p161

Strauss F M and Papagianis M D 1971 *Astrophys. J.* **164** 369

Stringer T E 1964 *Plasma Phys.* **6** 267

Sturrock P A 1966 *Nature* **211** 695

—— 1968 *Structure and Development of Solar Active Regions, IAU Symp. 35* ed K O Kipen-
     heuer (Dordrecht: Reidel) p471

—— 1973 *High Energy Phenomena on the Sun* ed R Ramaty and R G Stone (Greenbelt: NASA)
     p3

Sturrock P A *et al* 1977 *Possible Space Missions for Solar Research after Solar Maximum
     Mission. SUIPR Rep. No. 706* (Stanford: SUIPR (NASA))

Sturrock P A *et al* 1978 *Report of NASA Skylab Workshop on Solar Flares*

Svestka Z 1970 *Solar Phys.* **13** 471

—— 1973 *Ann. Rev. Astron. Astrophys.* **10** 1

—— 1975 *Solar Phys.* **47** No 1

—— 1976 *Solar Flares* (Dordrecht: Reidel)

Svestka Z (ed) 1977 *Solar Phys.* **53** 217

Sweet P A 1958 *Electromagnetic Phenomena in Cosmical Physics, IAU Symp. 6* ed B Lehnert (Dordrecht: Reidel) p123
—— 1969 *Ann. Rev. Astron. Astrophys.* **7** 149
Syrovatskii S I 1966 *Sov. Astron.* **10** 270
—— 1969 *Solar Flares and Space Research* ed C de Jager (Dordrecht: Reidel) p346
—— 1972 *Comm. Astrophys. Space Sci.* **4** 65
Takakura T 1972 *Solar Phys.* **26** 151
—— 1973 *High Energy Phenomena on the Sun* ed R Ramaty and R G Stone (Greenbelt: NASA) p179
Takakura T and Kai K 1966 *Publ. Astron. Soc. Japan* **18** 57
Takakura T, Ohki K, Shibunya N, Fujii M, Matsuoka M, Miyamoto S, Nishimura J, Oda M, Ogawara Y and Ota S 1971 *Solar Phys.* **16** 454
Tanaka K and Nakagawa Y 1973 *Solar Phys.* **33** 187
Tandberg-Hanssen E 1967 *Solar Activity* (Waltham, Mass.: Blaisdell) chap 7
Thompson W B 1955 *Proc. R. Soc.* A **233** 402
Tindo I P, Ivanov V D, Mandelstam S L and Shurygin A I 1970 *Solar Phys.* **14** 204
—— 1972a *Solar Phys.* **24** 429
Tindo I P, Ivanov V D, Valnicek B, and Livshits M A 1972b *Solar Phys.* **27** 26
Tindo I P, Mandelstam S L and Shurygin A I 1973 *Solar Phys.* **32** 469
Tindo I P and Somov B V 1977 *Lebedev Physical Institute Preprint* 149
Tomblin F F 1972 *Astrophys. J.* **171** 377
Trubnikov B A 1965 *Rev. Plasma Phys.* **1** 105
Tsytovich V N 1970 *Non-Linear Effects in a Plasma* (New York: Plenum)
Tsytovich V N, Stenflo L and Wilhelmsson H 1975 *Phys. Scr.* **11** 251
Ulmschneider P, Kalkofen W, Nowak T and Bohn U 1977 *Astron. Astrophys.* **54** 61
Underwood J H, Antiochos S K, Feldman U and Dere K P 1978 *Astrophys. J.* **224** 1017
Vasyliunas V M 1975 *Rev. Geophys. Space Phys.* **13** 303
Vernazza J E, Avrett E H and Loeser R 1973 *Astrophys. J.* **184** 605
Vesecky J F and Meadows A J 1969 *Solar Phys.* **6** 80
Vorpahl J, Gibson E G, Landecker P B, McKenzie D L and Underwood J H 1975 *Solar Phys.* **45** 199
Wang H T and Ramaty R 1974 *Astrophys. J.* **202** 532
Widing K and Cheng C C 1974 *Astrophys. J. Lett.* **194** L111
Wild J P and Smerd S F 1972 *Ann. Rev. Astron. Astrophys.* **10** 159
Wild J P, Smerd S F and Weiss A A 1963 *Ann. Rev. Astron Astrophys.* **1** 291
Wilson D C 1977 *NCAR Cooperative Thesis No 40* University of Colorado
Wood A T, Noyes R W, Dupree A K, Huber M C E, Parkinson W H and Withbroe G L 1972 *Solar Phys.* **24** 169
Wu S T, Dryer M and McIntosh P S 1975 *Solar Phys.* **44** 117
Zakharov V E 1972 *Sov. Phys.–JETP* **35** 908
Zaumen W T and Acton L W 1974 *Solar Phys.* **36** 139
Zheleznyakov V V 1970 *Radio Emission of the Sun and Planets* (Oxford: Pergamon)
Zirin H 1978 *Solar Phys.* **58** 95
Zirin H and Tanaka T 1973 *Solar Phys.* **32** 173