Nanoflares and the Heating of the Solar Corona

Arnold O. Benz

Institut of Astronomy, ETH Zurich, CH-8092, Switzerland
email: benz@astro.phys.ethz.ch

Abstract. New observational signatures of the heating process(es) have been revealed by space missions including SoHO, TRACE, Yohkoh, and RHESSI. Evidences for heating in the quiet corona, active region loops, and the solar wind are different and must be distinguished. Prime indications come from line broadening and waves, the distribution of temperature and radiation loss in relation to height, the correlation of magnetic flux and brightness, nanoflares and other fluctuations. This review concentrates on microevents observed at coronal temperatures. The reported nanoflares in quiet regions are about 2 orders of magnitude smaller than microflares reported in active regions and exhibit less radio emission. To estimate the impact of these microevents on the corona, the effects from a localized energy release must be considered, regardless of the energy source (reconnection or waves). In particular, the coupling with the chromosphere, i.e. the back-reaction of the chromosphere on coronal energy release, has an important effect on the corona. A necessary requirement for the heating process(es) is to deposit most of the heat in the low corona, but to heat also the upper corona to even higher temperature. Finally, the heating process must be able to account for the coronae of more active stars showing coronal emissions at levels of more than 3 orders of magnitude higher than the Sun.

1. Introduction

The heating of the solar corona has been a challenge to theory since the discovery of the million degree temperature in the late 1930s. It is no surprise that the number of theoretical studies is far too large to be reviewed here. In the first decades after the discovery, models dominated that sought to interpret it by MHD waves originating from the bubbling photosphere and dissipating their energy in the dilute corona. To honor this work, some few representative names of a very long list may be mentioned such as Biermann (1946), Alfvén (1947), Schwarzschild (1948), Schatrmann (1949), Ulmschneider (1967), Wentzel (1974), Hasegawa & Chen (1974), Stein (1981), Hollweg (1984), Ionson (1984). In the course of time, it has been shown that the initially favored sound waves dissipate their energy before they reach the corona (Athay & White 1978). Alfvén waves propagate energy along field lines, but lose energy only by non-linear effects and by inhomogeneity. They are still viable candidates for heating. Another line of heating models is nearly as old and is based on the fact that the magnetic field
dominates the energy in the corona. Thus only a small fraction of it may be thermalized and be the main source of heating. In the 1940s the dissipation of free magnetic energy was imagined as ohmic heating (such as by the Australian R. G. Giovanelli). However, the resistivity by Coulomb collisions was found to be too low, and a more impulsive energy release like in flares associated with anomalous resistivity was proposed (Gold 1964). The idea became more popular, when Parker (1983) worked out a detailed theory of coronal heating by a large number of small flares below resolution he coined ‘nanoflares’. Later theoretical studies include work by Heyvaerts & Priest (1984), Sturrock (1999), Kopp & Poletto (1993), Cargill (1993), Vekstein, Priest, & Steele (1993), and more recently numerical simulations such as Gudiksen & Nordlund (2002). In both wave and nanoflare heating, the energy originates from convective motions in the solar interior and propagates into the corona as a pointy flux. There are however significant differences how the energy is released, as well as differences in time scales and spatial distribution of energy deposition.

The number of observations relevant for coronal heating rivals the number of theories. Many observations have become common knowledge and trivial for insiders. One of them is the amazing fact that the corona apart from prominences, caused by some thermal instability, seems to have a temperature exceeding about one million degrees everywhere. There seem to be no cool places even at low solar activity, and the filling factor appears to be unity. Contrary to the chromosphere where such properties are hotly disputed, coronal heating seems to be ubiquitous. Other long-known facts include that the coronal is heated more above active regions and that the radiated power is an order of magnitude larger at maximum solar activity. Thus a dominant role of the magnetic field is well accepted.

Coronal observations were limited initially to occasional observations during solar eclipses and low resolution radio measurements. In recent years the pace has considerably changed due to space missions like Yohkoh, SoHO, TRACE, and RHESSI. High-energy observations of the corona have been complemented by high-resolution radio observations of the VLA. This review summarizes the surprising progress in an old field and how it has raised more questions than supplied answers.

The subject of coronal heating has recently been reviewed by Hollweg & Isenberg (2002), but not in full generality since the book by Ulmschneider, Rosner, & Priest (1991). The topic includes varies aspects and can be considered from the perspective of solar wind acceleration, active region loops or low corona quiet regions. This review focuses on the latter.

2. Recent Observations

In the past ten years it became routine to observe the low corona from space. Some major discoveries particularly relevant for coronal heating are summarized in the following.

The possibility to image the corona in soft X-rays allows to determine the electron temperature at different altitudes. In active regions, very hot loops are often found at relatively low altitude. Not so in the average corona. Sturrock, Wheatland, & Acton (1996) have integrated many Yohkoh/SXT off-limb images
and found the temperature to increase with altitude in quiet and open regions. This has been confirmed also for ion temperatures using the SUMER instrument on SoHO (see review by Marsch 2004). The temperature increase indicates that coronal heating does take place at high altitudes. However, the same and many other observations indicate that most of the luminosity of the quiet corona in soft X-rays and EUV lines originates in the first few $10^4$ kilometers above the transition region. In addition, transition region radiation indicates that a considerable amount of energy is lost by conduction to the layer below. As the low corona is poorly connected to high altitude (i.e. most loops are small), coronal heating must take place mainly at low altitudes, but continue in height.

An exciting discovery became possible when the temperatures of different ion species were measured with SoHO. The ion temperatures were found to be a decreasing function of the charge to mass ratio using (UVCS; Cranmer et al. 1999). This has been confirmed by SUMER (Marsch review, this volume). In addition it was found that the ion temperatures are enhanced in perpendicular direction. Temperature differences between species equilibrate in the corona within a few tens of seconds. Thus the observations can only be taken as signatures of an ongoing heating process. Interactions of coronal ions with plasma waves through cyclotron and Landau resonance heats them according to their charge-to-mass ratio and to the pitch angle. A natural explanation proposed is the resonant absorption of ion cyclotron waves, the left-circularly polarized high-frequency branch of the Alfvén waves (Marsh & Tu 2001). Ion cyclotron waves are a collisionless phenomenon, thus their life time is limited by the coronal collision time of a few seconds. The origin of these waves, however, is not clear. Axford & MacKencie (1991) suggested that small-scale reconnection events occurring in the supergranular network boundaries, liberating sub-photospheric convective energy, could be the origin of such waves.

The magnetic flux transport through the photosphere can now be measured with higher resolution than before. The MDI instrument on SoHO has revealed that it is much higher than previously assumed (Title & Schrijver 1998). Magnetic flux of both polarities emerges both in active and quiet regions of the photosphere at a rate of the total flux per 40 hours. This phenomenon has become known as "magnetic carpet". The implied energy transport is sufficient to heat the corona.

Parkers papers on nanoflare heating have motivated the search for small flares. The X-ray observations by Lin et al. (1984) in hard X-rays, and Simnett & Dennis (1985), Habbal (1992) in soft X-rays proved that the small energy releases are similar to regular flares. The coincidence of radio and X-ray microflares were reported by Benz et al. (1981), Gopalswamy et al. (1994), White et al. (1995), Gary et al. (1997). The sensitivity of microflare observations has been greatly improved by Shimizu (1995) using Yohkoh/SXT. It became possible for the first time to study the events statistically. Shimizu claimed that the observed microflares did not contain enough energy to account for the heating. He finds that their energy distribution follows a power law with an exponent below 2, thus insufficient to raise hope for a significant contribution of flares below resolution. It is important to notice that the flares reported by all these authors originated in active regions. These events could have contributed, if at all, only to the heating
of their region, but not to the entire corona that exists also in the absence of any active regions.

The absence of flares in quiet regions was the most serious shortcoming of the nanoflare heating hypothesis. It was known for decades that the transition region brightens impulsively in quiet regions (Brueckner & Bartoe 1983). Yet, the nature of these brightenings is not clear, their coincidence with coronal microflares is sporadic and their relation is still not confirmed, thus their contribution to coronal heating remains unknown. Harrison et al. (2003) note that transition region brightenings seem to be more numerous and have shorter duration than coronal microevents.

3. Nanoflare and Subresolution Events in Quiet Regions

Small flares in the quiet corona have first been discovered in deep exposures of Yohkoh/SXT (Krucker et al. 1997a). Initially, the equivalent of 0.33 nanoflares per second over the whole disk was reported. The reality of the flares in soft X-rays was established by coincident radio emission observed at the VLA. The radio emission at 2 cm was polarized, thus suggesting gyrosynchrotron radiation, and indicated even a delay of the soft X-ray emission known as the Neupert effect of regular flares. To distinguish them from microflares in active regions about two orders of magnitude more energetic, they have later been dubbed "nanoflares". All of them occurred above the magnetic network as indicated from chromospheric line observations. From Yohkoh/SXT images integrated in time, Benz et al. (1997) found "that most, but not all, enhanced magnetic field regions are sites of elevated X-ray emission" and therefore of enhanced coronal heating.

Using the very intense iron lines in the EUV, EIT on SoHo and later TRACE observed nanoflares in quiet corona with much higher sensitivity. Independently, Clette & Bergmans (1997) and Krucker, Benz & Delaboudinière (1997b) noticed that "even intra-network pixels with minimal emission measure fluctuate". Nevertheless, Benz & Krucker (1998) reported an increase of the amplitude of the events with emission measure of pixels and noted a nearly linear relation between inferred averaged input power and radiative loss per pixel. Krucker & Benz (1998) find an event rate of 306 nanoflares per second extrapolated to the whole surface of the Sun from deep EIT exposures in Fe IX/X and Fe XII lines.

The thermal energy content of coronal events can be estimated from their peak temperature and emission measure. The frequency distribution of nanoflare energy has received great attention, as a power-law index between 2.3 and 2.6 was initially reported by Krucker & Benz (1998), indicating that the integrated energy input is dominated by the smallest events. Criticism arose from the fact that the determined nanoflare energy depends on assumed model parameters that may change with size. A flare depth assumed proportional to the area, for example, reduces the power-law index to values closer to two (Parnell & Jupp 2000). More dramatic, however, are selection effects. As small events occur with a much higher rate close to the noise limit, a slight tightening in the selection criterion reduces the power-law index by a large amount (Aschwanden et al. 2000, Benz & Krucker 2002).
Nanoflares and the Heating of the Solar Corona

2004IAUS..219..461B

Figure 1. Nanoflare frequency distribution as observed by EIT (from Benz & Krucker 2002, solid curve) and TRACE (Parnell & Jupp 2000, dashed curve). The same methods to find events and determine the energy have been applied. The TRACE curve has been corrected (see text).

Figure 1 compares the energy distributions as observed by EIT and TRACE. A trivial arithmetic error in the publication of Parnell & Jupp (2000) has been corrected. The data of the two instruments agree well. The shift to smaller energies of the TRACE curve by about 30% is the result of a higher TRACE threshold, cutting off the rim of events (Benz & Krucker 2002). The power-law index is between 2.0 and 2.3.

Assuming a power-law index of 2.0, the observed frequency distribution \( f \) in flare thermal energy content \( E \) of plasma heated to \( T > T_0 \) (\( \approx 1 \) MK) amounts to

\[
    f(E) \approx f_0 E^{-2} \quad \text{[flares s}^{-1}\text{cm}^{-2}\text{erg}^{-1}] 
\]

and can easily be integrated to the flare energy \( F_{\text{th}} \) deposited per surface area,

\[
    F_{\text{th}} \approx f_0 \int_{E_{\text{min}}}^{E_{\text{max}}} E^{-2} E \, dV = f_0 \ln(E_m / E_{\text{min}}) = f_{\text{obs}} E_{\text{obs}}^2 \ln(E_m / E_{\text{min}}) .
\]

(2)

\( E_{\text{min}} \) and \( E_m x \) are the observed minimum and maximum flare energies, \( f_{\text{obs}} \) is some observed flare frequency at some reference energy \( E_{\text{obs}} \). From Figure 1 we can take the values \( E_{\text{min}} = 7 \times 10^{23} \text{erg} \) and \( E_m x = 5 \times 10^{26} \text{erg} \), and \( f_{\text{obs}} = 10^{-48} \) at \( E_{\text{obs}} = 9 \times 10^{25} \). Equation (2) then yields

\[
    F_{\text{th}} \approx 5.3 \times 10^4 \quad \text{[erg s}^{-1}\text{cm}^{-2}] .
\]

(3)

This may be compared with the observed total coronal radiation output of \( 4.5 \times 10^5 \) \( \text{[erg s}^{-1}\text{cm}^{-2}] \) by the same quiet region (Krucker & Benz 1998).
4. Statistical Analysis

The statistics of single pixels has so far received less attention. It allows investigating properties below the criteria for event selection. There is a non-gaussian, enhanced tail in the distribution of the observed fluctuations of a pixel. Figure 2 shows the difference between maximum and minimum during the time series. As the fitting gaussian distribution may include small fluctuations of solar origin, the curve is an upper limit to the noise contribution. The fraction of the pixels with non-gaussian deviations is 72% (ratio of integrals under dashed curve to full curve in Fig.2). This percentage is a lower limit on the pixels with significant changes of the emission measure within the observing time of 42 minutes.

Figure 3 displays a comparison of the fluctuations in various wavelengths with the 195 Å observations of EIT. The far best cross-correlation is found between the two coronal lines (top). The cooler lines, Fe IX-X, have been found to peak latest of all emissions observed. They are delayed by 23 seconds relative to Fe XII. The other cross-correlations have smaller peaks. It may interpreted by less correlation, but also by a displacement of the sources at different wavelength, such as footpoints vs. loop tops. The radio emissions precede the coronal brightenings, but scatter between a few seconds and more than one minute. Preceding radio emission in regular flares is well known to be caused by the Neupert effect, but most of the radio emission observed in the quiet Sun does not originate from gyrosynchrotron emission. The transition region line O V having a maximum ion formation temperature of 2.3·10⁵K precedes the coronal variations by a surprising 5 minutes. In the average it is the earliest emission of all studied here. The He I line originates mostly from plasma below 4·10⁵K, thus from the transition region and the chromosphere. In regular flares, O V enhancements have been interpreted by energy deposition of the precipitating electron beam.
of flare particles heating the upper chromosphere to some $10^5 \text{K}$ and exciting O V line emission by collisions. He I may be excited in a similar way (e.g. Porter et al. 1989).

![Cross-correlation graphs](image)

Figure 3. Cross-correlation of the temporal variations of various emissions in relation to the flux of the coronal Fe XII line. Pixels were adjusted to the same size and time interval. Then the time profile of each pixel was cross-correlated. The result was averaged over all pixels in common. Negative lag indicates that Fe XII is delayed (Benz & Krucker 1999).

The delay between Fe XII and Fe XI-X in flares is generally interpreted as a cooling of impulsively heated plasma. As convection seems to play a role in this process, longer loops are cooling slower. The observed delay in single pixel cross-correlation suggests that loops of the order of 2000 km dominate, the intersection of the observed delay and the dashed curve in Fig. 4. More importantly, it is pointed out here that a nearly linear relation between delay and loop length extending to nanoflares emphasizes the importance of conductive cooling. This in turn suggests that conductive losses dominate radiative losses thus requiring an even larger energy input for coronal heating.
5. Flare Energy Partition and Heating

It may be worth pointing out here that the reported energy in nanoflares may not be the primary energy released by the events. The enhanced emission measure indicates freshly heated material from the chromosphere. Energy may have been transported there most likely by particle beams. If nanoflares are similar to regular flares, the energy is primarily released into particle acceleration, or, equivalently, into heating coronal plasma to several keV energies. In addition, some energy is possibly converted into waves and mass motion. The initial partitioning of the released energy is not well known, thus the nanoflare energies are dubious.

With the new generation of solar instruments in EUV lines, soft and hard X-rays, a more accurate determination of the various forms of energy becomes feasible for large flares. Recently, Saint-Hilaire & Benz (2002) have presented a C9 event observed by TRACE at 195 Å and RHESSI. The TRACE observations, showing a blend of relatively cool material in Fe XXII and possibly hot plasma in Fe XXIV, display an ejection that does not leave the Sun, but apparently transfers its kinetic bulk energy to the corona. If the ejection is interpreted as a reconnection jet, as it starts simultaneously with the HXR emission, the flare energy initially went into electron (and possibly ion) acceleration, plasma motion (reconnection jet), and heating of the ejected material. The kinetic energy of the precipitating electrons observed in HXR is assumed to thermalize in the chromosphere and is estimated from a thick-target model. Plasma motion is observed and measured in EUV (TRACE observations). The thermal energy of the ejecta is also estimated from EUV observations. Figure 5 summarizes the observed and inferred energies. The observed total energy release amounts to $3.7 \times 10^{30}$ erg. The second line in Fig. 5 displays the primary energy partition and includes the best-estimate percentages. The energy in the possible reconnection jet motion is only 3% of the total, thus apparently an order of magnitude less
than the thermal energy of the hot plasma. Note that the partitions into the various forms of energy are tentative, particularly for the part that stays in the corona, namely the energy of the bulk motion of a possible reconnection jet and the thermal energy content of the heated material that was in the corona before and after the flare. They must be considered order-of-magnitude estimates, rounded such that each line adds up to 100%.

The secondary forms of energy are listed in the third line of Fig. 5. The thermal energy content of the flare kernel was measured in the 10–15 keV range by a fit to the RHESSI spectrum at low-energy peak time. The thermal energy content can be compared to the thermal radiation emitted by this plasma over the entire event, calculated from the measured flux in the observed band and extrapolated to the entire wavelength range using the observed temperature and emission measure. The radiated energy amounts to only 5% of the thermal energy content. It suggests that most of the thermal energy is lost in other ways, such as by heat conduction and emissions at lower temperature.

The third line of Fig. 5 also includes other forms of energy inferred from the thermal energy content measured at peak soft X-rays (SXR, < 15 keV). The other forms have been estimated for nanoflares in the quiet corona by Benz & Krucker (2002), who suggested that (i) the total energy content must be doubled, as roughly the same energy is lost before that peak as afterwards, and that (ii) the expansion of the heated plasma into the corona absorbs about three times as much energy. The rest of the energy deposited by the electron beam heats low temperature plasma in the transition region and chromosphere. For the flares considered here, the energy loss before the peak is less as regular SXR flares are usually asymmetric having a longer decay time. On the other hand flare loops are larger and require more energy for the expansion. The different energy partitions between nanoflares and regular flares are probably within the accuracy of these estimates.

Keeping in mind the large uncertainties given in Fig. 5, we note nevertheless that relatively little of the initial energy in accelerated electrons is actually observed in soft X-ray and coronal EUV emission. Even if this value is corrected for energy that is lost by radiation and conduction before being measured at SXR peak time, and by the energy needed for expansion into the corona, less than 10% of the total flare energy seem to end up in the hot flare plasma of the kernel. The rest, more than 60%, gets apparently lost in the chromosphere and transition region. Thus the conversion of non-thermal beam energy into hot thermal plasma seems to be less than 20% efficient in the observed flare.

The bottom line of Fig. 5 displays where the energy ends up. Roughly 40% of the released flare energy appears in thermal form of a plasma at coronal temperature. Some of the heated plasma ultimately cools to temperatures below the observing range, where its remaining energy is lost from the point of view of coronal observations. Most important, the RHESSI/TRACE observation suggests that the energy deposited into the corona may be an order of magnitude larger than the published estimates on the energy content of nanoflares.
Figure 5. Schematic distribution and low of energy in a C9-flare. Approximate (!) energy partition in percentage of the total is indicated as derived from the flare well observed by RHESSI and TRACE (from Saint-Hilaire & Benz, 2002). Encircled is the part measured as thermal energy content in coronal brightenings.

6. Conclusions

The popularity of nanoflare heating of the corona comes in waves. It reached a first peak in the 1980s with microflares discovered in active regions, a second peak was reached with nanoflares discovered in the quiet corona. In both cases it was later realized that the observed energy was too low to account for all of the heating, luring some people to conclude that nanoflares cannot heat the corona. It is argued in this review that this conclusion is premature.

It is clear from the experience of regular flares that soft X-ray and EUV brightenings are not direct evidence of the heating process, but represent secondary reactions of the chromosphere resulting from a primary energy input possibly at a different site. The conversion of the primary energy, possibly precipitating energetic particles, into thermal plasma at high temperature is just one of several possible channels of coronal energy input. Nevertheless the observed impulsive energy content manifest in nanoflares of the quiet corona amounts to more than 10% of the energy required for the observed coronal radiation. The rest of the energy input appears in the form of a quasi-constant base. Its origin is unclear and may have several reasons: Unresolved nanoflares at lower energy may mimic a smooth input. Energy released directly into the corona by the observed microevents in the form of waves may distribute energy in a non-impulsive way. More information on flare energy partition is necessary to decide whether there is room for some entirely different process to release additional energy.
References

Biermann, L. 1946, Naturwiss., 33, 228
Hollweg, J. V. & Isenberg, P. A. 2002, JGR, 107, SSH 12-1
Marsch, E. & Tu, C.-Y. 2001, JGR, 106, 227
Marsch, E. 2004, this volume
Schatzmann, E. 1949, Ann. d’Astrophys. 12, 203
Ulmschneider, P. 1967, ZAp, 67, 193