THE GENERATION OF ION CYCLOTRON TURBULENCE BY THE INTERMITTENT HEAT FLUX IN CORONAL HOLES

S. A. Markovskii and Joseph V. Hollweg

Space Science Center, University of New Hampshire, Durham, NH 03824, USA, Email: Sergei.Markovskii@unh.edu

ABSTRACT

Recently, we suggested that the source of ion heating in solar coronal holes is small-scale reconnection events (microflares) at the coronal base. The microflares launch intermittent heat flux up into the corona exciting ion cyclotron waves through a plasma microinstability. The ions are heated by these waves during the microflare bursts and then evolve with no energy input between the bursts. The overall coronal heating by this mechanism is a summed effect of all microflare bursts during the expansion time of the solar wind and adiabatic cooling between the microflares. In this paper, we calculate the wave energy density of the turbulent fluctuations excited by the heat-flux instability. We show that the time-averaged energy density needed to heat the coronal holes is well within the available observational constraints. The sporadic density is much greater than the time-averaged one but is still small enough for the instability to operate in a marginal regime.

1. INTRODUCTION

The observational data from the UVCs instrument on the SOHO spacecraft provided us with important clues to understanding the coronal heating. These observations strongly suggest that the ions of a coronal hole plasma are heated by waves in the ion cyclotron frequency range. In particular, the data show that the heating increases the ion temperature mostly in the direction perpendicular to the background magnetic field (e.g., Kohl et al., 1998; Dodero et al., 1998; Antonucci et al., 2000) and different ions are heated with different efficiency (e.g., Kohl et al., 1998; 1999). All these and other properties can be explained by ion cyclotron wave-particle interaction (e.g., Hollweg and Isenberg, 2002). It should be mentioned that other possibilities have been discussed too; see Hollweg and Isenberg (2002) for a review. These possibilities cannot be ruled out because there is no direct evidence for the ion cyclotron waves, but the indirect evidence is rather convincing and we will assume that the ion cyclotron waves do exist in the corona.

Then, the next question to ask is, where do these waves come from and how are they replenished to provide the energy source for the corona and the solar wind. This aspect of the coronal heating is much less understood. Several mechanisms of the wave generation have been suggested so far, but all of them seem to have difficulties putting together all elements of the heating process. A recent review of the existing mechanisms can be found, for instance, in Markovskii and Hollweg (2004a). It should be emphasized that the difficulties do not definitively rule out any of the proposed mechanisms of the coronal heating. At the same time, we recently suggested a new approach involving plasma microinstabilities generated by large but intermittent electron heat flux in the corona. The aim of this paper is to develop these ideas further.

2. DESCRIPTION OF THE HEATING MODEL

2.1 Basic concepts

The basic concepts of our model were introduced by Hollweg and Markovskii (2002) and Markovskii and Hollweg (2002b; 2003). Here we briefly repeat them for completeness. The first element of our model is magnetic reconnection resulting in microflares at the coronal base. During the flares a significant portion of the magnetic field energy goes to the thermal energy of electrons. The electrons are heated locally and when the hot particles become collisionless they escape from the heating site generating heat flux going up into the corona. This heat flux can be pictured as electron beams for simplicity, although in fact the distribution function is rather skewed than has a bump-on-tail structure.

At the next step in our model, the heat flux loses energy to protons. This is possible because the heat flux can excite a plasma microinstability and the generated waves are cyclotron resonant with the protons. In our case, the instability has the same nature as a current-driven instability. A simple way to describe it is to consider two electron components: a hot and tenuous heat flux carrying beam and a relatively cold and dense core corresponding to the background electron population. The components of the electron distribution move with respect to each other at the thermal speed of the hot component, and they also move with respect to the protons. This drift of the background electron population with respect to the protons can drive the instability as if there were a current, even though the total current is assumed to be zero. In particular, much like the current-driven instability, the heat-flux instability generates electrostatic waves at lower values of the plasma β and shear Alfvén waves at higher β (e.g., Markovskii and Hollweg, 2002b; 2003). These waves can then interact with ions and heat them.

It is important that in our model the heating has to be intermittent. When the microflare occurs, the heat flux only exists in the flux tube connected to the flare site, because the particles do not move across the magnetic field. The next flare at the same site does not happen immediately. There is some waiting time between the flares. Therefore, the beams launched by the microflares have a finite extent along the background magnetic field and these localized
beams are separated by periods of much smaller heat flux. Furthermore, it turns out that not only is the intermittency inevitable, but this mechanism can only work if the heat flux is intermittent. If the heat flux were constant and large enough to exceed the threshold of the instability, then its contribution to the total energy flux carried by the solar wind would be much greater than the total energy flux observed at 1 AU, which is impossible. To avoid this contradiction, we assume that sporadically the heat flux is large enough to excite the instability, but because these sporadic bursts are rare, the total time-averaged heat flux is much smaller and it gives the correct energy flux to the solar wind. In fact, this is how we can estimate the degree of intermittency, i.e., the ratio of the time between the bursts to the burst time. We need to compare the energy budget of the solar wind and the threshold of the instability. At the coronal base the degree of intermittency turns out to be around several hundred (Markovskii and Hollweg, 2004a).

Markovskii and Hollweg (2004a) have developed a simple method of calculating the proton heating during the heat flux burst. They have shown that it is sufficient to assume a marginal instability slightly above the threshold, so the calculation can be done in the quasilinear limit. The reason is that the characteristic time of the global, macroscopic, expansion of the solar wind is much greater than any characteristic time of the microscopic plasma processes, except perhaps the collision time. For instance, the proton gyroperiod is about 6 orders of magnitude smaller than the expansion time at 1.5 solar radii. Therefore, even if the instability results only in slow heating on the microscopic scale, it can still be fast enough on the global scale to account for the observed heating of the solar wind.

The heating during the heat flux bursts contributes to the overall heating, which in this model is a summed effect of a burst and of what happens until the next burst. The solar wind behavior between the bursts is relatively simple. There is no energy input and the solar wind just cools down adiabatically because of its expansion. This kind of behavior is illustrated in Fig. 1. If we mark a plasma parcel and follow its motion with the solar wind, the temperature profile of this parcel will look like sawteeth, instead of a smooth line. By comparing the heating during the burst and the adiabatic cooling between the bursts, we can determine whether this mechanism maintains the temperature on the level suggested by the observations.

One of the problems associated with intermittent heating is that the proton distribution function at the beginning of the interval $\Delta R$ in Fig. 1 is not the same as it is at the end of the interval. It gets distorted. There are several sources of the distortion. One of them is the heating itself. The wave particle interaction due to the heat flux instability operates only in a certain region of the particles phase space. As a result, the distribution deviates significantly from its initial structure during the burst. Another source of the distortion is the mirror force, which is different for particles with different velocities across the magnetic field. Therefore, in general, for every cycle $\Delta R$ we would need to calculate the heating for a different distribution. It turns out, however, that there is always a process that works against the distortion of the distribution function. This process can recycle the distribution, which will then have qualitatively the same structure by the end of every cycle.

The situation is especially simple close to the Sun because for reasonable parameters of the heat flux the time between the bursts can be greater than the time of proton-proton collisions. As a result, the distortion from the heating is completely absorbed and the distribution is isotropized by the collisions by the time of the next burst. The distortion due to the mirror force also does not occur. This means that the variation of the macroscopic solar wind parameters can be calculated simply from fluid equations. Note that the particle diffusion within the burst is not affected by the collisions because the bursts are short enough.

2.2 Plasma heating in the collision-dominated region of the solar wind

Below, we will focus on the collision-dominated region of the solar wind. A discussion of the recycling mechanisms at greater distances can be found in Markovskii and Hollweg (2004a). A simple way to describe the global evolution of the solar wind in the collision-dominated region is to use the fluid equations with a heating function derived from the microscopic kinetic theory. In the fluid models, the heating function determines the amount of energy put into the solar wind by some unspecifed mechanism at a given heliocentric distance and it plays the role of a source term in the MHD equation of energy. To establish the analogy between the fluid models and our heating mechanism, we need to assume that the time between the heat flux bursts is much smaller than the solar wind expansion time. In that case, the variation of the macroscopic solar wind parameters will be determined by a time-averaged effect of many bursts and each burst will give only small variations of the solar wind parameters.
Since the bursts are very short and the evolution of the solar wind between the bursts is controlled by the collisions, as discussed above, our mechanism will then give the same results as the fluid model if the heating function is the same.

The solar wind expansion time, which we define as the ratio of the scale of the density inhomogeneity to the solar wind speed, varies with the heliocentric distance. The minimum value of the expansion time is approximately 1000 s at 1.5 solar radii. Following Markovskii and Hollweg (2004a), we assume that the time interval between the heat flux bursts is \( \Delta t = V(R)\Delta R \) is much smaller and equals to 100 s. Here \( V(R) \) is the solar wind speed. Once this time is specified, we can estimate the maximum heliocentric distance \( R_{coll} \) at which the distribution between the bursts can be recycled and isotropized by the collisions by comparing \( \Delta t \) with the proton-proton collision time. As shown by Markovskii and Hollweg (2004a), \( R_{coll} \) cannot exceed approximately 1.5\( R_{sun} \).

To illustrate the heating and acceleration of the solar wind, we will use the following numerical example. We start from the steady-state fluid equations for the protons with gravity

\[
n(R)V(R)A(R) = \text{const},
\]

\[
V(R)\frac{dV(R)}{dR} = -\frac{k}{m_p n(R)} \frac{d[n(R)(T_p(R) + T_e)]}{dR} - \frac{GM_{sun}}{R^2},
\]

\[
V(R)\frac{dT_p(R)}{dR} = -\frac{2}{3} \frac{T_p(R)}{A(R)} \frac{d[A(R)]}{dR} + \frac{2Q(R)}{3nk},
\]

where \( R \) is the heliocentric distance, \( Q \) is the heating function, and \( A \) is the flow tube cross section. Recall that the quantities entering into these equations are time-averaged ones as displayed in Fig. 1 by a dashed line. For simplicity, we assume that the electron temperature \( T_e \) is constant.

In general, the heat flux instability results in momentum addition as well as heating, because the particles acquire a bulk velocity along the magnetic field, but in the present case the momentum addition is insignificant (Markovskii and Hollweg, 2004a). Usually Eqs. (1)–(3) are solved with respect to the unknown functions \( n(R) \), \( V(R) \), and \( T_p(R) \). Our goal, however, is to check whether the heating rate associated with the heat flux instability can provide reasonable parameters of the solar wind. Therefore, we assume that the density \( n(R) \) is a known function that can be derived from observations, and consider \( Q(R) \) the third unknown function instead.

We supplement the system of Eqs. (1)–(3) by the area expansion function \( A(R) \) in the form

\[
A(R) = \frac{f_{max}(R/R_{sun})^2}{1 + (f_{max} - 1)(1 + ((R/R_{sun}) - 1)^p)}.
\]

Here \( f_{max} \) is the overall expansion factor of the flow tube, which we set equal to 5. The constants \( b \) and \( c \) determine how the overall expansion is distributed as a function of the heliocentric distance. To illustrate how our mechanism operates in the vicinity of the proton temperature maximum, we put \( b = 2 \) and \( c = 4 \). With these parameters, the temperature maximum of the solution obtained below is achieved within 1.5 solar radii. This condition is somewhat arbitrary but it is not inconsistent with the observed temperature limits discussed by Esser et al. (1999).

We then solve Eqs. (1)–(3) with the boundary condition \( T_p(1.5R_{sun}) = 1.5 \cdot 10^6 \text{ K} \), which is again consistent with the observations reported by Esser et al. (1999). The constant in Eq. (1) is equal to the proton flux density at 1 AU \( 3 \cdot 10^6 \text{ cm}^{-2} \text{s}^{-1} \). The density is given by the formula

\[
n = 3.2 \cdot 10^8 r^{-15.6} + 2.5 \cdot 10^8 r^{-3.76} \text{ cm}^{-3}
\]

derived from the observational data of Feldman et al. (1997), where \( r \) is the heliocentric distance in solar radii. In addition, we set \( T_e = 7.5 \cdot 10^5 \text{ K} \) in agreement with the observational data showing that the electrons are cooler than the protons in the corona (Wilhelm et al., 1998). The numerical solutions for \( T_p(R) \) and \( Q(R) \) are displayed in Fig. 2. We do not extend these solutions beyond 1.5\( R_{sun} \) because the isotropic fluid equations are not applicable to our model at greater distances. At the same time, within 1.5\( R_{sun} \) these solutions give the temperature values close to the observed ones, and the heating rate comparable to the one used, for instance, by Esser et al. (1997) in their fluid model of the solar wind.

3. HEATING RATE PRODUCED BY THE INSTABILITY

We now postulate that the heating function shown in Fig. 2 is produced by the heat flux exciting the plasma instability. The heat flux is controlled by three parameters of the electron beam, its bulk speed, temperature, and density. For an illustration, we assume that the beam electrons are heated up to the temperature \( T_b = 8 \cdot 10^6 \text{ K} \) at the coronal base and the beam is launched at the bulk speed equal to its thermal speed \( V_{th} = \sqrt{2kT_b/m_e} \). Following Markovskii and Hollweg (2004a), we will also
assume that the beam speed and temperature remain constant as it propagates to higher altitudes. Then, the beam density will be the only varying parameter. As shown by Markovskii and Hollweg (2004b), this is a reasonable approximation at least in our region of interest, within 1.5Rsun.

To proceed, we need to relate the heat flux instability with the heating function. The heating function plays the role of a source term in the energy equation. Consequently, it is equal to the energy input during the heat flux burst divided by the time between the bursts

$$Q \approx \frac{3}{2} nkT_p / \Delta t,$$

where $T_p$ is the proton temperature. Recall that there is no energy input between the bursts and the interval $\Delta t = 100$ s includes the time of the burst itself $\Delta t$ and the period of the zero heat flux until the beginning of the next burst.

The temperature increase $\Delta T_p$ for given background plasma parameters and given heat flux above the threshold value can be calculated using the method developed by Markovskii and Hollweg (2004a). If the proton distribution function before the heat flux burst is Maxwellian then after the burst the distribution takes the following form. Outside of a certain resonance interval $-\nu_{\text{max}} < \nu_z < -\nu_{\text{min}}$ of the particles velocity along the background magnetic field $B_0 = (0, 0, B_{\phi})$, the distribution function $f_p$ preserves its initial Maxwellian structure. Inside the resonance interval, $f_p$ reads

$$f_p(v_z, \nu_z) = \frac{n_p(v_{\text{p}}^2, \pi)^{-3/2} 2\pi v_z F_1(v_z^2 - \alpha v_z^2)}{v_{\text{max}} - \sqrt{\nu_z^2 - \alpha v_z^2}}$$

if $0 < \nu_z < \sqrt{\alpha v_z^2 - \alpha v_z^2_{\text{min}}}$ and

$$f_p(v_z, \nu_z) = \frac{n_p(v_{\text{p}}^2, \pi)^{-3/2} 2\pi v_z F_2(v_z^2 - \alpha v_z^2)}{\nu_{\text{max}} - \nu_{\text{min}}}$$

if $\sqrt{\alpha v_z^2 - \alpha v_z^2_{\text{min}}} < \nu_z < \infty$, where

$$F_1(\eta) = e^{-\eta/\sqrt{2}} \int_{-\nu_{\text{max}}}^{-\alpha \nu_{\text{max}}} e^{-(\alpha+1)\eta s^{2}/2} ds,$$

$$F_2(\eta) = e^{-\eta/\sqrt{2}} \int_{\nu_{\text{max}}}^{-\nu_{\text{max}}} e^{-(\alpha+1)\eta s^{2}/2} ds.$$  

Here $n_p$ and $V_{\text{TP}}$ are the proton number density and thermal speed, $\alpha = \Omega_p / (\Omega_{\text{max}} - \Omega_p)$, where $\Omega_p$ is the proton gyrofrequency and $\Omega_{\text{max}}$ is the frequency of the unstable waves at the maximum growth rate, $\nu_{\text{max}} = \sqrt{\nu_z^2 + \nu_{\text{r}}^2}$ is the particle velocity perpendicular to the background magnetic field, and the negative $\nu_z$ is in the sunward direction. The distribution is affected by the waves only in the region $-\nu_{\text{max}} < \nu_z < -\nu_{\text{min}}$ because the unstable waves are excited only in a certain interval of frequencies $\omega$ and parallel wavenumbers $k_z$, where the growth rate of the instability calculated from the linear dispersion relation (see Eq. (12) below) is positive. As a result, the waves can be cyclotron resonant only with particles that have certain parallel velocities. The boundary values $-\nu_{\text{max}}$ and $-\nu_{\text{min}}$ are related to the boundaries of the instability region in the $(\omega, k_z)$ space by the resonance condition

$$\omega - \Omega_p = k_z v_z$$

(11)

At typical distribution function obtained this way is shown in Fig. 3 as a 3D logarithmic plot.

In the low-$\beta$ region of the corona close to the Sun the dominant mode of the heat-flux instability is the electrostatic one. The dispersion relation of the heat flux instability in the frame moving at the bulk speed of the protons is given by the usual formula:

$$1 + \varepsilon_p + \varepsilon_c + \varepsilon_b = 0,$$

(12)

$$\varepsilon_p = \frac{2\omega_p^2}{k^2 V_{\text{TP}}^2} \left[ 1 + \frac{i \sqrt{\pi} (\omega + i \nu_c) \sum_{n=-\infty}^{\infty} W \left( \frac{\omega + i \nu_c - n \Omega_p}{k_z V_{\text{TP}}} \right) \right] \int \left( \frac{k_z^2 \nu_c^2}{2 \Omega_p^2} \right) I_n \left( \frac{k_z^2 \nu_c^2}{2 \Omega_p^2} \right) \left( \frac{k_z^2 \nu_c^2}{2 \Omega_p^2} \right),$$

(13)

$$\varepsilon_c = \left(2\omega_p^2/v_{\text{c}}^2 V_{\text{TP}}^2\right) W \left( (\omega + i \nu_c - k_z V_{\text{TP}}) \right) \sum_{n=-\infty}^{\infty} W \left( (\omega + i \nu_c - n \Omega_p)/k_z V_{\text{TP}} \right) \right] \times \exp(-k_z^2 \nu_c^2/2 \Omega_p^2) I_n \left( k_z^2 \nu_c^2/2 \Omega_p^2 \right),$$

(14)

$$\varepsilon_b = \left(2\omega_b^2/k^2 V_{\text{TP}}^2\right) W \left( (\omega + i \nu_b - k_z V_{\text{TP}}) \right) \sum_{n=-\infty}^{\infty} W \left( (\omega + i \nu_b - n \Omega_p)/k_z V_{\text{TP}} \right) \right] \times \exp(-k_z^2 \nu_b^2/2 \Omega_p^2) I_n \left( k_z^2 \nu_b^2/2 \Omega_p^2 \right),$$

(14)
into the heating function (6)

\[ \delta T_p = \frac{m_p}{3n_0 k} \int_{-\infty}^{\infty} dv_z \int_{0}^{\infty} (v_z^2 + v_Tp^2) f_p - f_0) dv_p, \]  

(16)

where \( f_0 \) is the initial Maxwellian distribution. Note that the distribution function \( f_p \) is anisotropic. Therefore, in principle, we have to distinguish between the variations of the perpendicular and parallel temperatures. However, for an order-of-magnitude estimate it is sufficient to calculate an overall temperature variation, which we define in Eq. (16) proceeding from the variation of the total thermal energy per particle.

Based on these calculations we have derived the following numerical formula for the growth rate of the instability required to produce a given \( \delta T_p / T_{p0} \) as a function of \( \delta T_p / T_{p0} \) and \( T_{p0} / T_c \) in our region of interest \( 0.8 < T_p / T_c < 2.15 \) and \( 0 < \delta T_p / T_{p0} < 0.1 \):

\[
\frac{\gamma}{\Omega_p} = \left[ 0.597 - 0.470 \frac{T_{p0}}{T_c} + 0.106 \left( \frac{T_{p0}}{T_c} \right)^2 \right] \frac{\delta T_p}{T_{p0}}
\]

\[ + \left[ -4.77 + 3.61 \frac{T_{p0}}{T_c} - 0.780 \left( \frac{T_{p0}}{T_c} \right)^2 \right] \left( \frac{\delta T_p}{T_{p0}} \right)^2
\]

\[ + \left[ 14.2 - 4.36 \frac{T_{p0}}{T_c} \left( \frac{\delta T_p}{T_{p0}} \right) - 19.3 \left( \frac{\delta T_p}{T_{p0}} \right)^4 \right],
\]  

(17)

where \( T_{p0} \) is the temperature of the initial Maxwellian distribution. The function \( \gamma / \Omega_p \) is plotted in Fig. 4. Equation (17) relates \( \delta T_p \), and thus the macroscopic heating rate given by Eq. (6), to the growth rate of the microscopic instability, which is determined by how much the heat flux is greater than the threshold value. We will use formula (17) in the next section to compute the energy density of the unstable waves.

4. ENERGY DENSITY OF THE UNSTABLE WAVES

An important property of the electrostatic ion cyclotron waves is that they are compressional and they can be, in principle, detected by radio scintillation measurements (Hollweg, 2000). At present, the observational data are available only for regions as close to the Sun as 1.5R_{Sun}. Therefore, the data cannot be directly compared to our calculations within 1.5R_{Sun}. Nevertheless, it is useful to estimate the density fluctuation level associated with the instability to make sure that it does not get unreasonably high.

The wave energy density at the saturation phase of the instability \( W \) is related to the proton heating rate by the following approximate formula:

\[ Q = nk (dT_p / dt) \sim \gamma W. \]  

(18)

This relation is obtained from the general diffusion equation for the particles neglecting the derivatives with respect to the particle velocity parallel to the magnetic field in the case of highly oblique waves (e.g., Lysak et al., 1980). Using Eqs. (6), (17), and (18) along with the solutions of the fluid equations for \( Q(R) \) and \( T_p(R) \) shown in
The relative wave energy density $W/nkT_p$ of electrostatic waves is of the same order as the observed quantity $(\delta n/n)^2$, where $\delta n$ is the density fluctuation. The observational data in coronal holes at $5R_{\odot}$, obtained by Coles and Harmon (1989), have been analyzed, in particular, by Markovskii and Hollweg (2002a). It was shown that the square of the relative amplitude $(\delta n/n)^2$ associated with the density fluctuations in the entire spectrum (from the large-scale energy containing region to the small-scale dissipation region) is $2 \cdot 10^{-5}$ and $(\delta n/n)^2$ associated with the dissipation region alone is $3 \cdot 10^{-4}$. As we see from Fig. 5 the density fluctuations occurring in our model are well below the observed values. It should be mentioned that the observed spectrum does not extend beyond length scales of the order of the ion inertial length, while the electrostatic waves have much smaller wavelengths, comparable to the proton gyroradius. However, at $5R_{\odot}$ the plasma $\beta$ is higher than at $1.5R_{\odot}$ and the shear Alfvén branch of the heat flux instability becomes the dominant mode. The wavelength of the shear Alfvén instability is greater than the proton gyroradius and can be comparable to the proton inertial length.

Finally, we note that the wave energy density and the heating rates in Eqs. (6) and (18) are time-averaged ones, because the time interval $\Delta t$ in Eq. (6) includes the duration of the heat flux burst itself $\Delta t$ and the period of zero heat flux and no energy input between the bursts. The sporadic wave energy density $(\Delta t/\Delta t)W$ and heating rate $(\Delta t/\Delta t)Q$ are much greater. Nevertheless, it can be shown that $(\Delta t/\Delta t)W$ is of the order of $10^{-4}$ in the entire region under consideration. Indeed, at the coronal base the degree of intermittency $\Delta t/\Delta t$ is around several hundred. As the heliocentric distance increases, the quantity $W$ increases (Fig. 5), but $\Delta t/\Delta t$ decreases to about 10 at $1.5R_{\odot}$ (Markovskii and Hollweg, 2004a; b). The reason why the heat flux becomes less intermittent is that the beams have a finite spatial extent and they are expanding in both directions along the magnetic field roughly at the mean square speed of the collisionless beam particles. As a result, the time interval between successive beams $\Delta t - \Delta t$ is decreasing, while $\Delta t$ stays constant. The low level of the sporadic wave energy density $(\Delta t/\Delta t)W$ assures that the instability is operating in a marginal regime in agreement with our assumption.

Thus, the parameters of the turbulent fluctuations needed to heat the coronal holes by our mechanism are consistent with both observational and theoretical constraints.

ACKNOWLEDGMENTS

This work is supported by NSF under grant ATM-0211402 and by NASA’s Sun-Earth Connection Theory program under grant NAG5-11797, Solar and Heliospheric Physics SR&T program under grant NAG5-10988, and Living With a Star program under grant NAG5-10835 to the University of New Hampshire.

REFERENCES