CHROMOSPHERIC AND CORONAL HEATING BY SOUND WAVES

R. GRANT ATHAY AND O. R. WHITE

High Altitude Observatory, National Center for Atmospheric Research

Received 1978 May 5; accepted 1978 June 15

ABSTRACT

Observations of solar oscillations in the middle chromosphere with the University of Colorado extreme-ultraviolet spectrometer on OSO 8 give an estimated energy flux in sound waves of approximately $1 \times 10^4$ ergs cm$^{-2}$ s$^{-1}$. This is interpreted as a lower limit to the true flux resulting from a possible underestimate of the velocity amplitude in the wave. An upper limit to the wave flux is obtained by setting the velocity amplitude equal to $2^{1/2}$ times the nonthermal line broadening velocity interpreted as microrotation. In the middle chromosphere, the maximum wave energy flux falls significantly below the heat input needed for the upper chromosphere and corona. Also, it is shown that the heat input has two maxima within the chromosphere with the intervening minimum input coinciding with the region where the wave energy flux becomes too low to produce the required heating in the higher layers. These results are interpreted to mean that two heating mechanisms are present, and that heating by sound wave dissipation is important only in the low chromosphere. For the average Sun, the situation is no more favorable for Alfvén waves.

Subject headings: radiative transfer — Sun: atmosphere — Sun: chromosphere — Sun: corona

1. INTRODUCTION

The suggestion that sound waves generated in the solar convection zone heat the chromosphere and corona has been popular for many years. From the qualitative point of view, this suggestion is very appealing. Turbulent eddies in the convection zone clearly are a source of sound wave generation, and some of the wave energy clearly must propagate outward and dissipate in the higher layers.

Quantitatively, the theory is less satisfying. The energy flux carried by sound waves in the convection zone is uncertain by a large factor. In addition, the waves must traverse the photosphere where they are subject to strong radiative dissipation, and a substantial portion of the energy flux must pass through the steep temperature and density gradients in the chromosphere and chromosphere-corona transition region before they reach the corona. Reflection of the waves in regions of steep gradients and by the fine structure of the upper chromosphere represents a major obstacle to transmitting sound wave energy into the corona. No comprehensive theory starting with wave generation and following the propagation of the waves through the chromosphere to the corona has been attempted. Stein (1967) and others have discussed the expected form of the energy spectrum for the waves in the convection zone, and Ulmschneider and Kalkofen (1977) and others have considered the propagation through the photosphere into the low chromosphere. Even here there are major uncertainties in the results.

However, it seems clear from these studies that sound waves with periods in the range 10–50 s are good candidates for heating the low chromosphere. For the corona, one needs longer-period waves, but the theory usually starts by postulating the existence of sufficient energy flux at some convenient level within the transition region. There is neither observational nor theoretical justification for the energy fluxes assumed.

Estimates of the flux of mechanical energy required to replace the excess loss of energy by radiation from the chromosphere and corona are discussed by Athay (1976). In the lowest layers of the chromosphere, H$^-$ radiation is the dominant loss mechanism. In the middle chromosphere, the Fraunhofer lines become the dominant radiation with most of the loss occurring in Balmer-alpha and the infrared triplet of Ca II. Radiation losses from these regions are model dependent and are based on computations using the Vernazza, Avrett, and Loeser (1973) model. The losses from H$^-$ radiation are estimated at $2 \times 10^8$ ergs cm$^{-2}$ s$^{-1}$. Balmer and Ca II radiation losses are of similar magnitude, giving a total loss of approximately $4 \times 10^8$ ergs cm$^{-2}$ s$^{-1}$ from the middle and low chromosphere.

In the high chromosphere and corona the dominant radiation losses shift to the EUV spectrum and can be obtained directly from observations. Lyman-alpha carries a flux of about $3 \times 10^7$ ergs cm$^{-2}$ s$^{-1}$ from the upper chromosphere. A flux of similar magnitude is carried from the transition region and low corona by a number of lines formed by the heavier elements. Thus, the upper chromosphere and corona have a total radiation loss of approximately $6 \times 10^7$ ergs cm$^{-2}$ s$^{-1}$.

Compressive waves are known to exist in the photosphere and chromosphere. At photospheric levels, the wave spectrum extends from periods near
300 s to periods less than 15 s (Deubner 1976). The longer-period waves have been shown to be global P-mode oscillations of high order and to be mainly evanescent (Deubner 1976; Rhodes, Ulrich, and Simon 1977). Propagation studies of the shorter-period waves have not been carried out. However, the acoustic cutoff frequency in the photosphere is near 200 s, and the shorter-period waves are expected to propagate. In addition, some fraction of the energy at longer periods propagates into the higher layers. Deubner (1976) estimates the total energy flux in waves of all periods in the photosphere to be approximately $10^6$ erg s$^{-2}$ s$^{-1}$. This is likely an overestimate, but it is far above the flux of approximately $5 \times 10^5$ ergs cm$^{-2}$ s$^{-1}$ needed to heat the chromosphere and corona.

In recent studies of chromospheric oscillations using the O/SO 8 satellite we have found strong evidence of longitudinal waves with periods from near 300 s to less than 30 s (Athay and White 1979; White and Athay 1978). In this case, the waves with periods between 300 and 100 s are found to propagate energy at about the sound speed. Propagation speeds for the energy in the shorter-period waves could not be determined accurately. However, assuming that all of the chromospheric waves propagate energy at the sound speed, we estimated the total observed energy flux as $1 \times 10^6$ erg s$^{-2}$ s$^{-1}$ at a height in the chromosphere of approximately 1200 km relative to $v_{6000} = 1$. As noted in the preceding discussion, the flux needed to heat the higher layers of the chromosphere and corona is about $6 \times 10^5$ ergs cm$^{-2}$ s$^{-1}$ provided all of the wave flux can be converted to heat and provided it can be properly distributed in height. As we have pointed out (Athay and White 1978), our estimate of the energy flux in the waves could be in error as a result of a variety of effects that tend to obscure the true velocity amplitude. Nevertheless, it is clear that the concept of heating the upper chromosphere and corona by sound waves must be reexamined.

Jordan (1977a, b) recently has argued that sound waves are not the main source of energy for the upper chromosphere. His conclusions are based on attempts to construct model chromospheres by assuming a balance between dissipation of wave energy and radiative losses. Such models appear to fail in the upper chromosphere. However, calculations of this type require that both the spectral energy shape and energy flux in the waves be specified at some boundary level. Also, the results are subject to several additional sources of uncertainty including effects of chromospheric fine structure and unknown propagation characteristics of the waves.

In this paper, we derive an upper limit to the energy flux in sound waves from observational data, and we examine the required heat input to the chromosphere, each as a function of height. We find that the heat input has two maxima and that the flux in sound waves is too low to produce the second maximum. This suggests that the heating mechanism for the upper chromosphere and corona is different from the heating mechanism for the low chromosphere.

II. HEAT INPUT

Energy loss from the middle and upper chromosphere is mainly by radiation in spectral lines, and we equate this loss to the heat input. Thermal energy is converted to radiation via collisional excitation of atomic energy levels followed by radiative decay. The rate of energy loss per unit volume, and hence the required heat input, is given by (cf. Athay 1976)

$$\nabla \cdot F_{\text{in}} = \sum_i \nu_i N_{Li} C_{Li}$$. \hspace{1cm} (1)

where $F_{\text{in}}$ is the heat flux, $N_{Li}$ is the population of the lower level of the transition, and $C_{Li}$ is the collisional excitation rate to the upper level. The sum is over all of the relevant transitions. For simplicity, we will drop the subscript $i$ and consider the summation as being implicit.

In the Vernazza, Avrett, and Loeser (1973) (VAL) model of the chromosphere there are two temperature plateaus. The upper one is near 20,000 K and produces much of the Lyman series emission. Energy loss from this plateau is mainly in $\text{H}_\alpha$, and we will refer to it as the Lyman plateau. Most of the radiation loss from the lower plateau is in $\text{H}_\alpha$, the infrared triplet of $\text{Ca}$ II, the H and K lines of $\text{Ca}$ II, and the resonance doublet of $\text{Mg}$ II (cf. Athay 1976). The strongest single source probably is $\text{H}_\alpha$, but the infrared triplet of $\text{Ca}$ II is a strong competitor. We refer to this lower plateau as the Balmer plateau.

Estimates of the energy losses from the two plateaus (Athay 1976) yield $2 \times 10^6$ ergs cm$^{-2}$ s$^{-1}$ from the Balmer and $3 \times 10^5$ ergs cm$^{-2}$ s$^{-1}$ from the Lyman. The Balmer plateau has a thickness of about 1000 km whereas the Lyman plateau has a thickness of only about 150 km. Thus, the average loss per unit height is approximately the same in the two cases. This means that the heat input can be monotonic in height only if it is approximately constant.

Since we wish to discuss the height dependence of $F_{\text{in}}$, it is convenient to express each of the factors in equation (1) as exponential functions of height. We note that $N_{Li}$ can be written as

$$N_{Li} = \frac{N_{Li}}{N_i} \frac{N_i}{N_H} \frac{N_H}{N_{Hi}}$$, \hspace{1cm} (2)

where $N_i$ is the total concentration of the ion or atom and $N_{Hi}$ is the hydrogen concentration. We take $N_i/N_{Hi}$ to be constant and write $N_{Hi}$ in the form

$$N_{Hi} \propto \exp (-\beta_i h)$$, \hspace{1cm} (3)

where $h$ is height and $\beta_i$ is the reciprocal scale height. The factor $N_{Li}/N_i$ can usually be expressed as a Boltzmann factor of the form

$$N_{Li}/N_i \propto \exp (-a \chi_{Li}/T)$$, \hspace{1cm} (4)

where $a = 11600$ and $\chi_i$ is the excitation energy of level $i$ in eV. In order to convert this to an exponential function of height, we let

$$T = T_0 (1 + \Delta T/T_0)$$.
and replace $1/T$ with
\[ \frac{1}{T} = \frac{1}{T_0} \left( 1 - \frac{\Delta T}{T_0} \right) = \frac{1}{T_0} (1 - \beta_T h). \quad (5) \]

This is a valid approximation provided $\Delta T/T_0$ is small compared to unity, as it is in the temperature plateaus. (Note that $\beta_T$ is positive for $dT/dh$ positive.) Equations (4) and (5) give
\[ \frac{N_u}{N_i} \propto \exp \left( \frac{a}{T_0} \chi L \beta_T \right) h. \quad (6) \]

The collision rate $C_{LU}$ can be written in the form
\[ C_{LU} \propto N_f f(T) \exp \left( -\frac{aX_{LU}}{T} \right). \quad (7) \]

The quantity $f(T)$ is usually a slow function of $T$, and its height variation can be ignored within a plateau. As before, we replace the exponential temperature term with
\[ \exp \left( -\frac{aX_{LU}}{T} \right) \propto \exp \left( \frac{a}{T_0} \chi L \beta_T h \right), \quad (8) \]

and we replace $N_s$ with
\[ N_s \propto \exp (\beta_e h). \quad (9) \]

Combining all of the above height dependences in equation (1) and setting $\nabla \cdot F_{in} \propto \exp (\beta_{in} h)$, we find
\[ \beta_{in} = \beta_H + \beta_e - \frac{a}{T_0} (\chi L + \chi_{LU}) \beta_T. \quad (10) \]

In this form the total addition to the flux obtained by integrating over height above height $h$ has the same height dependence as $\nabla \cdot F_{in}$. Also, $\chi_L + \chi_{LU}$ is just the excitation energy, $\chi_L$, of the upper level. Thus, we can write equation (10) in the alternative form
\[ \beta_{in} = \beta_H + \beta_e - \frac{a}{T_0} \chi_L \beta_T. \quad (11) \]

Hydrogen ionization increases rapidly across the Balmer plateau, and, as a result, $\beta_e$ within this plateau is coupled with $\beta_T$. To estimate the relationship between $\beta_s$ and $\beta_T$, we write the modified Saha equation for the $N_i$ level of hydrogen in the form
\[ N_i \approx N_H \propto \xi T^{-3/2} \exp (+aX_{LU}/T). \quad (12) \]

In the Balmer plateau in the VAL model $b_i$ and $T^{-3/2}$ are relatively constant. Thus, we may express the height-dependent form of equation (12) as
\[ \beta_H = 2\beta_e + \frac{a}{T_0} \chi L \beta_T \]

or
\[ \beta_e = \frac{1}{2} \beta_H - \frac{a}{2 T_0} \chi L \beta_T. \quad (13) \]

Equations (11) and (13) then give
\[ \beta_{in} = (3/2) \beta_H - \frac{a}{T_0} \left( \chi_L + \chi_L \right) \beta_T. \quad (14) \]

The quantity $(a/T_0)(\chi_L + \chi_L/2)$ is not particularly sensitive to individual chromospheric models. For the Balmer plateau $a/T_0 \approx 1.9$. Also, $\chi_L/2 = 6.8$ and $\chi_L = 10$ for $\text{H} \alpha$ and 3.2 for both the infrared triplet and the H and K lines of $\text{Ca} \ \Pi$. Thus, the coefficient of $\beta_T$ in equation (14) is 31.9 for $\text{H} \alpha$ and 19 for $\text{Ca} \ \Pi$. Also, $\beta_T$ is not very model dependent and has a value near 5.5 Mm$^{-1}$ (Mm = megameter). The quantity $\beta_T$, of course, is very model dependent, and it follows that $\beta_{in}$ is model dependent. In the VAL model $\beta_T \approx 0.11$ Mm$^{-1}$, and we find
\[ \beta_{in} = 0.86 \beta_H, \quad \text{H} \alpha, \]

and
\[ \beta_{in} = 1.12 \beta_H, \quad \text{Ca} \ \Pi. \]

In each of these cases $\nabla \cdot F_{in}$ decreases by a large factor across the Balmer plateau. Since the average value of $\nabla \cdot F_{in}$ in the same in the Balmer and Lyman plateaus, the decrease across the Balmer plateau must be followed by an increase in the Lyman plateau. This means that $\nabla \cdot F_{in}$ has at least two maxima, one in the low chromosphere and a second in the Lyman plateau or above. Such a distribution suggests, but does not necessarily require, two different input mechanisms.

Since $\beta_{in}$ is model dependent, we can alter the model to reduce $\beta_{in}$. A large decrease in $\beta_{in}$ can be accomplished only by increasing $\beta_T$. On the other hand, there are limits on the amount by which $\beta_T$ can be increased and still yield an acceptable model. For example, a wealth of observations of line and continuum intensities in the chromosphere seen at the limb at the time of total eclipse requires that $\beta_e$ be positive, i.e., that $N_e$ decrease outward.

The condition for $\beta_e > 0$, according to equation (13), is
\[ \beta_T < \frac{T_0}{aX_L} \beta_H = 0.039 \beta_H. \]

With this limit on $\beta_T$, equation (14) gives
\[ \beta_{in} > 0.26 \beta_H, \quad \text{H} \alpha, \]

and
\[ \beta_{in} > 0.76 \beta_H, \quad \text{Ca} \ \Pi. \]

For even the smaller of these values, $\nabla \cdot F_{in}$ decreases by over a factor of 4 across the Balmer plateau. It seems highly improbable, therefore, that any acceptable revision of the model will satisfy the dual conditions $\beta_e > 0$ and $\beta_{in} \approx 0$. Thus, we cannot escape the conclusion that $\nabla \cdot F_{in}$ has at least two maxima separated by a minimum located near the top of the Balmer plateau.

In the next section, we show that the energy flux in sound waves falls below the required heat input near the same height where this minimum in the heat input occurs.
III. SOUND WAVES

Although the observed energy flux in sound waves in the middle chromosphere as inferred from the measured velocity amplitude is much too low to heat the upper chromosphere and corona, there are a number of reasons for believing that the measured velocity amplitude is lower than the true velocity amplitude, possibly by a large factor (Athay and White 1978). However, the effects of any velocities that are present cannot be completely hidden. Any radiative transfer, chromospheric, or instrumental effects that obscure the true velocity amplitude, as measured by line shifts, must broaden the lines. Thus, the true velocity amplitude must be contained in the combination of line broadening and line shifting.

In the chromosphere the line broadening velocities in excess of the thermal velocity are much larger than the velocity amplitudes of the oscillations. It has been suggested a number of times that the nonthermal broadening of chromospheric lines is due, in fact, to unresolved wave motions (cf. Boland et al. 1973). Regardless of whether or not this is true, we can obtain an upper limit to the energy flux in waves by equating the wave amplitude to the maximum amplitude that is consistent with the nonthermal component of line broadening. Also, the maximum propagation speed for the wave energy is the sound velocity $V_s$. Thus, we write the maximum wave energy flux as

$$F_w^{\text{max}} = \frac{1}{2} \rho V_{\text{max}}^2 V_s,$$  

where $\rho$ is the matter density, $V_{\text{max}}$ is maximum velocity amplitude consistent with nonthermal broadening velocity, and $V_s$ is the sound velocity. Both $\rho$ and $V_s$ are well known, so the only freedom is with $V_{\text{max}}$.

The minimum effect on line broadening by nonthermal motions occurs when the motions act in the sense of microturbulence, i.e., to raise the effective “temperature” of the small-scale motions. Thus, we maximize $V_{\text{max}}$ by treating all of the excess line broadening as microturbulence. We further maximize $V_{\text{max}}$ by assuming that the microturbulence is given by the average of the wave velocity over a full oscillation. This is equivalent to assuming that the broadening is produced by coherent, plane waves, and gives $V_{\text{max}}^2 = 2 \xi^2$, where $\xi$ is the microturbulent velocity. Hence, we obtain

$$F_w^{\text{max}} = \rho \xi^2 V_s.$$  

Values for the microturbulent velocity are given in the VAL model. These have been modified recently by Tripp, Athay, and Peterson (1978) as a result of spectrum synthesis of the EUV emission lines of $\text{Si II}$ and $\text{Si III}$. These lines are formed in a temperature range that spans both the Balmer and Lyman plateaus, and they provide a sensitive measure of $\xi$. The mean values of $\xi$ found for the Balmer and Lyman plateaus by Tripp et al. are close to the mean values in the VAL models. However, the gradients are different.

Tripp et al. find the best fit to the $\text{Si II}$ and $\text{Si III}$ lines when $\xi = 0.09 T^{1/2}$ at all heights above 1000 km. Also, we may replace $V_s$ with $(\gamma k T / \mu m_p)^{1/2}$, where $\gamma$ is the ratio of specific heats and $\mu$ is the mean molecular weight. These substitutions give

$$F_w^{\text{max}} = 1.2 \times 10^{-12} N_\text{He} T^{3/2} (\gamma / \mu)^{1/2}.$$  

Within the Balmer plateau, $T^{3/2}$ varies by less than 20% and $\gamma / \mu$ is a slowly varying quantity of order unity. The reason for this latter condition is that hydrogen is mainly neutral but is ionizing rapidly as the temperature increases. For mostly neutral hydrogen, $\mu \approx 1.3$; and for a strong change in ionization associated with a small change in temperature, $\gamma$ close to unity. It follows from equation (18) that $F_w^{\text{max}}$ decreases with height approximately in proportion to $N_\text{He}$, i.e.,

$$\beta_w \approx \beta_\text{He}.$$  

Values of the parameters in equations (17) and (18) and resultant values of $F_w^{\text{max}}$ are summarized in Table 1 for five heights in the chromosphere and lower transition region. $T$ and $N_\text{He}$ are taken from the VAL model for $h \leq 2300$ km and from an extrapolation of the model at constant gas pressure for $T = 10^8$ K. Values of $\gamma$ are estimated somewhat crudely as $4/3$ in the Balmer plateau and $5/3$ in the Lyman plateau, and values of $\mu$ is determined from the model. These values are not important to the results. The values of $\xi$ are taken from Tripp, Athay, and Peterson (1978). They differ from $V_s$ by a factor $(\gamma / \mu)^{1/2}$. The last column of Table 1 contains estimated values for

$$\int_h^\infty \nabla \cdot F_w \, dh$$

taken from Athay (1976).

We note from the results in Table 1 that

$$F_w^{\text{max}} \ll \int_h^\infty \nabla \cdot F_w \, dh$$

for $h \leq 2000$ km, i.e., near the top of the Balmer plateau. With such an inequality it is clear that sound waves are not the heating source for the upper chromosphere and corona. We note that the region where $F_w^{\text{max}}$ falls below the required heat input is just the region where $\nabla \cdot F_w$ has a minimum. It is noteworthy, also, that hydrogen changes from being mainly neutral to being mainly ionized near this same height. Furthermore, the value of $F_w^{\text{max}}$ at this height is obtained with an assumed velocity amplitude in the waves of 1.4 times the sound speed. It is highly unlikely that the velocity amplitude could be much larger than this. Thus, the upper limit set for $F_w$ appears to be firmly established. The estimates of

$$\int_h^\infty \nabla \cdot F_w \, dh$$

for $h \geq 2000$ km come almost entirely from EUV emission lines where one simply needs to sum the observed line intensities. About half of the EUV
TABLE 1

<table>
<thead>
<tr>
<th>Height (km)</th>
<th>T(K)</th>
<th>$N_s$</th>
<th>$\gamma/\mu$</th>
<th>$V_e$</th>
<th>$\xi$</th>
<th>$F_{w}^{\text{max}}$</th>
<th>$\int \nabla \cdot F_{w} , dh$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>6100</td>
<td>$2.5 \times 10^{12}$</td>
<td>1</td>
<td>7</td>
<td>7</td>
<td>$1.4 \times 10^9$</td>
<td>$5 \times 10^6$</td>
</tr>
<tr>
<td>2000</td>
<td>6800</td>
<td>$1 \times 10^{11}$</td>
<td>1</td>
<td>7.4</td>
<td>7.4</td>
<td>$6.7 \times 10^6$</td>
<td>$6 \times 10^6$</td>
</tr>
<tr>
<td>2300</td>
<td>20300</td>
<td>$2 \times 10^{10}$</td>
<td>2.5</td>
<td>20.5</td>
<td>12.8</td>
<td>$1.1 \times 10^6$</td>
<td>$3 \times 10^6$</td>
</tr>
<tr>
<td>2500</td>
<td>100000</td>
<td>$4 \times 10^9$</td>
<td>2.5</td>
<td>45</td>
<td>28.4</td>
<td>$2.0 \times 10^6$</td>
<td>$2 \times 10^6$</td>
</tr>
</tbody>
</table>

* Velocity units are km s$^{-1}$, and flux units are ergs cm$^{-2}$ s$^{-1}$.

emission is in $\lambda_a$ and the remainder is mainly in lines formed in the transition region. The observed EUV flux fluctuates with solar activity by about a factor of 2 but never approaches the low value of $F_{w}^{\text{max}}$ at 2000 km.

If we allow for the fact that much of the energy flux in waves will be reflected and refracted in the upper chromosphere and transition region, we can accept the wave heating hypothesis only when $F_{w}^{\text{max}} \gg \int \nabla \cdot F_{w} \, dh$. This condition is satisfied only in the low chromosphere. We conclude, therefore, that the Balmer plateau may be heated by sound waves but that the Lyman plateau, the transition region, and the corona are not heated by this mechanism.

IV. ALFVÉN WAVES

There is a possibility that the microturbulence derived from line broadening is produced by Alfvén waves rather than sound waves. If this is true, then the energy flux propagates at the Alfvén speed rather than the sound speed. The ratio of the Alfvén speed to the sound speed is given by

$$\frac{V_A}{V_s} = B \left( \frac{1}{4\pi \mu k T \sigma} \right)^{1/2}$$

At a height of 2000 km, the parameters in Table 1 give $V_A/V_s \approx B$, the magnetic field strength in gauss. Therefore, we can increase $F_{w}^{\text{max}}$ at 2000 km by whatever value of $B$ we are willing to consider. Since the wave flux is linear in $B$, the spatially averaged flux is proportional to the spatially averaged $B$. Estimates of the spatially averaged $B$ in the photosphere give values of 2–3 gauss, and it is therefore not very helpful to replace $V_A$ by $V_s$ for the average Sun.

In active regions and in network features where $B$ may exceed 10 gauss in the upper parts of the Balmer plateau, heating of the overlying corona and upper chromosphere by Alfvén waves is still a possibility. This does not remove the overall problem for the average Sun, however, and heating of the upper chromosphere and corona by Alfvén waves appears to be just as unacceptable as heating by sound waves.

REFERENCES


R. GRANT ATHAY and O. R. WHITE: High Altitude Observatory, P.O. Box 3000, Boulder, CO 80307