Wave heating of the solar atmosphere

Iñigo Arregui\textsuperscript{1,2}

\textsuperscript{1}Instituto de Astrofísica de Canarias, Vía Lactea s/n, La Laguna E-38205, Spain
\textsuperscript{2}Departamento de Astrofísica, Universidad de La Laguna, La Laguna E-38206, Spain

Magnetic waves are a relevant component in the dynamics of the solar atmosphere. Their significance has increased because of their potential as a remote diagnostic tool and their presumed contribution to plasma heating processes. We discuss our current understanding of coronal heating by magnetic waves, based on recent observational evidence and theoretical advances. The discussion starts with a selection of observational discoveries that have brought magnetic waves to the forefront of the coronal heating discussion. Then, our theoretical understanding of the nature and properties of the observed waves and the physical processes that have been proposed to explain observations are described. Particular attention is given to the sequence of processes that link observed wave characteristics with concealed energy transport, dissipation and heat conversion. We conclude with a commentary on how the combination of theory and observations should help us to understand and quantify magnetic wave heating of the solar atmosphere.

1. Introduction

In spite of decades of advances in observations and theory, the solar coronal heating problem remains unsolved. A full understanding of the magnetic and plasma structuring of the corona, the nature of the plasma dynamics and the physical processes that create and maintain a hot atmosphere remains elusive (see [1–8] for reviews).

The coronal heating problem originated more than 70 years ago, when Grotrian [9] and Edlén [10] identified the presence of Fe IX and Ca XIV spectral lines in the light emitted by the corona, indicating the presence...
of fully ionized plasma at multi-million degrees. In a recent historical note, Peter & Dwivedi [11] assign the merit to Alfvén [12], who summarized six arguments that support the corona being heated to such high temperatures. Since then, solar physics research has sought to identify the physical mechanism(s) that might balance the thermal conduction, radiation and solar wind losses, first quantified by Withbroe & Noyes [2]. Researchers in the field agree upon the fact that the origin for such temperatures is of a magnetic nature and that the energy source lies in the solar surface plasma motions. This available energy is then transported to the upper layers of the solar atmosphere where it gets dissipated. The exact physical processes and their contribution to the heating of the plasma remain largely unknown or unquantified [8].

A number of mechanisms are believed to contribute to the heating: the direct dissipation of magnetic energy by processes such as magnetic reconnection [13], current cascades [14], viscous turbulence [15] or magnetic field braiding [16]; the dissipation of magnetic wave energy stored, transported and transferred to small scales by waves [17–21]; or the mass flow cycle between chromosphere and corona [22] are some suggested mechanisms. There is ample evidence for the occurrence of magnetic reconnection processes [23]; magnetic wave dynamics [24]; and mass flows connecting chromospheric and coronal regions [25]. All these processes are under theoretical and observational study and advances on their description have been made so a discussion in terms of mutually exclusive mechanisms is beside the point. It is important however to assess the importance of each mechanism.

Wave energy transport and dissipation were regarded as relevant soon after the coronal heating problem arose. The first proposed theories involved wave-based mechanisms. Biermann [26] and Schwarzschild [27] suggested that acoustic waves, generated in the convection zone, supplied the non-radiative energy needed to heat the chromosphere and corona in the form of mechanical energy transported outwards (see the review by Stein & Leibacher [28]). Magnetic wave heating was proposed by Alfvén [17] soon after the existence of magnetohydrodynamic (MHD) waves was postulated by himself [29]. It was soon realized that the dissipation of magnetic waves could have an important role in coronal heating, because surface Alfvén waves can be transmitted upwards in the atmosphere [30] and heat the corona [31]. These waves are indeed subject to mechanisms, such as resonant absorption [18] and phase mixing [19], that enhance the dissipation of energy by resistive or viscous processes, which would otherwise take too long in the high Reynolds number corona. Theoretical and numerical studies on the dissipation of hydromagnetic waves [32,33] and the resonance absorption of Alfvén waves [18,34–36] were carried out in open and closed magnetic structures. Numerical experiments included the sideways or foot-point driving of the structures, trying to mimic the buffeting of the solar surface [37,38]. The first relations to loop observations were made by Hollweg & Sterling [39], to explain data by Golub [40]. However, proper instrumentation able to detect waves and reliably measure their properties was not available.

This changed near the end of the last century, when routine observations of the solar atmosphere made with instruments onboard the Solar and Heliospheric Observatory (SOHO) and the Transition Region and Coronal Explorer (TRACE) led to a number of discoveries. Remarkable examples are the transverse oscillations in coronal loops [41,42] and their rapid time damping; the compressive waves in short and long coronal loops [43–45]; the quasi-periodic oscillations of the Doppler shift and intensity in emission lines in coronal loops [46,47] and their damping; or the quasi-periodic disturbances in open magnetic regions of the corona [48,49]. Extended reviews on these observations can be found in Nakariakov & Verwichte [24], De Moortel [50], Aschwanden [51] and De Moortel & Nakariakov [52].

The observed wave activity is interpreted in terms of standing and propagating MHD waves. Nowadays, the existence of theoretically predicted MHD wave types in the solar atmosphere is well established [24,53–55]. This has led to the development of solar atmospheric seismology [56,57], which enables the determination of physical parameters in the solar atmosphere from the comparison of observed and theoretical wave properties. The application of this method to the solar atmosphere has been successful, but the wave heating perspective was discontinued for more than a decade.
A renewed interest has now grown on the wave heating question. This was motivated by the evidence for ubiquitous wave activity at different layers of the solar atmosphere, gathered by observations with instruments onboard, for example, the Hinode and Solar Dynamics Observatory (SDO) spacecraft or the High Resolution Coronal Imager (Hi-C) rocket experiment. These high-resolution observations have demonstrated the presence of oscillations of magnetic and plasma structures in solar chromospheric spicules [58]; coronal magnetic loops [59]; prominence plasmas [60,61]; X-ray jets [62]; and extended regions of the corona [63,64]. The observed disturbances display time and spatial damping possibly indicating in situ energy dissipation. In some cases, their energy content seems to be large enough to compensate estimated energy losses. These observations have replaced coronal wave heating at the forefront of the discussion, leading to the need to reassess the significance of wave energy transport and dissipation to the heating problem [8,65].

We discuss the role of magnetic wave energy transport and dissipation processes in the heating of the solar corona, based on our current theoretical understanding and observational evidence. We focus on the following aspects: in §2, we present a selection of wave activity observations by pointing out their relevance to the heating problem; §3 offers a description of theoretical models and physical processes that have been suggested to explain observations; particular emphasis is placed in §4 on the series of physical processes that link wave energy content, propagation, damping and dissipation. Section 5 contains a commentary on how the combination of data and theory, by means of forward modelling of observational signatures and seismic inversion of observed wave properties, should help advance the field in the future. Our conclusions are presented in §6.

2. Observed wave activity

Early observations already pointed to the existence of quasi-periodic perturbations in solar coronal structures (e.g. [66–73]). The detection of these oscillations was based on the measurement of the temporal and spatial variation of spectroscopic properties (intensity, width and Doppler velocity) of coronal emission lines. Some of these observations took advantage of the instrumentation onboard Skylab or Yohkoh spacecraft, but were mostly restricted to time-series analyses with little spatial information (see [74] for a summary of early observations).

This situation changed with the advent of the imaging and spectroscopic instruments onboard SOHO and TRACE spacecraft. Because of the temperature discrimination and spatial resolution capabilities of their extreme ultraviolet (EUV) and soft X-ray telescopes, observations demonstrated the existence of wave-like dynamics in the form of, for example, density fluctuations in plumes and coronal holes [48,49,75], propagating compressional disturbances in coronal loops [43–45,76–78] or transverse coronal loop oscillations [41,42].

The existence of waves and oscillations in magnetic and plasma structures of the solar atmosphere is now beyond question. The spatial, temporal and spectral resolution of imaging and spectroscopic instruments in current ground- and space-based observatories (SST, DST, SOHO, TRACE, CoMP, Hinode, STEREO, SDO, Hi-C and IRIS) have enabled us to directly image and measure motions associated with wave dynamics with increasing precision. Data show evidence for waves in increasingly finer plasma structures belonging to regions of the solar atmosphere with different physical properties, such as coronal loops; prominence plasmas; chromospheric spicules and mottles; coronal holes, plumes, etc. Let us focus on some particular examples and draw attention to the aspects that are relevant to wave heating.

Transverse loop oscillations offered the first imaging evidence for standing waves in the corona associated with the periodic displacement of these structures [41,42,79,80]. In these events, wave damping can be directly observed and measured. This constitutes important information for seismology diagnostics [81,82] and the reason for considering damping mechanisms such as resonant absorption [83]. However, TRACE data showed that these were occasional events and loops are hot irrespective of the presence or absence of lateral displacements. On the other hand, the Coronal Multi-channel Polarimeter (CoMP) observations indicate that coronal disturbances
are present in extended regions of the corona (1.05–1.35 \( R_{\text{Sun}} \)). The measured Doppler velocity fluctuations, of the order of 0.3 km s\(^{-1}\), do not produce significant intensity variations and are interpreted as Alfvén waves propagating along the coronal magnetic field, with measured phase speeds of the order of 1–4 Mm s\(^{-1}\) \[63,64\]. Importantly, CoMP data show signatures of \textit{in situ} wave damping in the form of a discrepancy in the outward to inward wave power, but energy estimates (approx. 0.01 W m\(^{-1}\)) seem to fall below the amount required to heat the ambient plasma. The SDO has shown that Alfvén waves are common in the transition region and corona \[59\] and that motions visible in both regions share a common origin \[84\]. Energy estimates vary from one region to another. The disturbances seem to be energetic enough to power the quiet Sun corona and coronal hole regions, with an estimated wave power of approximately 100–200 W m\(^{-2}\), but not the active region corona (approx. 100 W m\(^{-2}\) in front of the required 2000 W m\(^{-2}\)). Observations of transverse waves in the active region corona by Morton & McLaughlin \[85\], obtained by combining the highest available resolution instruments onboard Hi-C and SDO/Atmospheric Imaging Assembly (AIA), confirm that wave activity is all-pervasive in the EUV corona. However, this activity is concluded to be of low energy because of the relatively low displacement and velocity amplitudes (less than 50 km and less than 3 km s\(^{-1}\), respectively) in the measured waves.

Chromospheric spicules are excellent candidates for mass and wave energy transport to the corona, as they provide a physical connection between both media. They are present everywhere and at all times and show a complex dynamics with a combination of lateral swaying, upward mass flows and torsional motions \[86,87\]. Recent observations by Morton \[88\] suggest that wave damping is present in these structures. The recent analyses by the \textit{Interface Region Imaging Spectrograph} (IRIS) team will clarify many aspects about their field and plasma structuring and dynamics \[89,90\]. Propagating and standing transverse oscillations are also present in chromospheric mottles belonging to strong photospheric field concentrations. By measuring the amplitude and phase speed variation along these structures, Kuridze \textit{et al.} \[91\] suggest that their sudden change at a given height is an indication of the presence of nonlinear wave processes. Photospheric field concentrations, such as bright points, were proposed as the source of torsional Alfvén waves by Jess \textit{et al.} \[92\]. The increasing evidence for surface vortex motions \[93,94\] strongly suggests that this type of dynamics can be important in the generation and wave energy propagation to upper layers \[95\]. The chromosphere seems to be a vast reservoir of MHD wave energy with observations showing a complex combination of intensity and width variations of chromospheric structures together with their transverse displacement \[96\]. The evidence points to the photospheric excitation of the dynamics observed at the upper layers \[97\].

Waves in polar coronal hole regions are gaining ample attention. The measurement of wave characteristics, such as the amplitude, periodicity and phase speed provide plasma and field diagnostics and constraints on theoretical models of coronal heating and solar wind acceleration \[98\]. Recent analyses have enabled the measurement of the spatial variation of the non-thermal velocity for Alfvén waves and the electron density along the propagation direction. Banerjee \textit{et al.} \[99\] find that the non-thermal velocity is inversely proportional to the quadratic root of the electron density, in agreement with what is predicted for undamped radially propagating linear Alfvén waves. Hahn \textit{et al.} \[100\] report that the widths of spectral lines, which are assumed to be proportional to the wave amplitude, decrease at relatively low heights. This is interpreted as an indication of wave damping. The energy dissipated in between 1 and 1.3 solar radii is estimated to account for up to 70% of the amount required to heat the polar coronal hole and accelerate the solar wind. Similar analyses have quantified the energy carried out and dissipated by Alfvén waves in these structures \[101,102\]. By combining spectroscopic measurements with a magnetic field model, Hahn & Savin \[103\] were able to trace the variation of the non-thermal velocity along the magnetic field above the limb of a quiet Sun region. They find that the waves are dissipated over a region centred on the top of the loops. The position along the loop where the damping begins is strongly correlated with the length of the loop, implying that the damping mechanism depends on the global loop properties rather than on local collisional dissipation. Transverse
waves seem to be present also in solar plumes and Thurgood et al. [104] have measured the occurrence rate of wave properties, such as periods and displacement and velocity amplitudes, in great detail.

The relationship between wave signatures at different heights of the solar atmosphere is reported in many studies (e.g. [105–107]). They provide information about the phase speed of propagating waves from the time delay of signatures detected at different heights in the solar atmosphere [108]. The high spatial and temporal resolution of SDO/AIA is enabling additional analyses of the spatial distribution of the frequency, time-series analyses and correlations between the signals at different atmospheric heights [109,110]. However, obtaining a causal connection between wave dynamics at different regions of the atmosphere is still problematic. Understanding how the waves are generated and behave as a function of the line formation temperature and the magnetic field structure is essential [111] and observations should be complemented with numerical modelling [112,113].

3. Theoretical models and physical processes

The observed wave activity is interpreted as the manifestation of the presence of MHD waves. There are three basic MHD wave-types in a uniform plasma of infinite extent and they have properties that make them suitable for wave energy transport and eventual plasma heating. Fast waves transport energy across magnetic surfaces. They resonantly couple to Alfvén and slow waves in non-uniform plasmas, which leads to the transfer of energy from large to small spatial scales. Slow waves are basic in the generation of wave dynamics in the lower solar atmosphere, where the plasma still plays the relevant role in front of the magnetic forces. They are crucial in nonlinear wave transformation processes leading to the development of shocks and their eventual dissipation. Alfvén waves can connect and carry energy between remote regions of the atmosphere. They propagate to large distances along the magnetic field lines. Their energy is not easily dissipated, but their dynamics is characterized by large cross-field gradients in non-uniform plasmas, which leads to the enhancement of viscous and resistive dissipative processes.

None of the observed waves is expected to happen with the properties described in textbooks [114–116]. Pure MHD waves do not exist in the solar atmosphere. The classic description in terms of fast, slow and Alfvén waves treated separately is not accurate enough for the description of the observed dynamics. The highly non-uniform and dynamic nature of the plasma in the solar atmosphere leads to complex wave properties. We identify three levels of complexity in the theoretical and numerical modelling of MHD waves in the solar atmosphere, as follows.

(i) The analysis of linear waves in simple magnetic configurations enables us to study basic properties of wave trapping and propagation and to understand in detail the workings of individual physical mechanisms. They offer local information on plasma and field properties from seismology analyses, but observational signatures with little potential to be compared with real data. (ii) The analysis of nonlinear waves in structured and dynamic plasmas enables us to study wave transformation processes, wave–flow interactions and the modelling of relevant processes such as shock waves, nonlinear Alfvén waves and Alfvén wave turbulence. They produce useful synthetic data to extract conclusions from comparison with observations. (iii) Global numerical models use prescribed or observed boundaries and drivers as an input on large-scale numerical simulations. They make it possible to include almost any physical ingredient, to perform global seismology and to obtain useful observational signatures and predictions.

The three approaches offer valuable information and are highly complementary. Results from simple models cannot be directly used to extract solid conclusions, but offer detailed understanding about physical processes that might occur under real Sun conditions. Their benefit is that the mechanism(s) of interest can be studied in isolation and under controlled conditions. At the other extreme, sophisticated simulations enable direct comparison with observed data, but detailed understanding is lost when many physical processes are included in the simulations simultaneously. The level of detailed understanding will decrease as we increase applicability and vice versa.
Resonant damping of oscillations is a well-studied mechanism for energy transport from large-scale transverse motions to localized small-scale motions and is due to the inhomogeneity of the medium in the cross-field direction (see [83] for a review). The mechanism has been mainly studied in flux tube models [117,118], but a cylindrically symmetric flux tube is not necessary for the process to operate (e.g. [119]). Resonant absorption is robust in front of model complications, such as nonlinear evolution [120] or multi-stranded structure [121,122], and works for both standing [123,124] and propagating [125] waves, producing time and spatial damping of the wave amplitude, respectively. This is an ideal physical process, independent of resistivity in the limit of high Reynolds numbers [126]. Once in the resonant layer, energy can further scale down to smaller spatial scales by the process of phase mixing [19,127]. The stronger the cross-field inhomogeneity, the quicker is the damping, but the slower the small-scale creation [128]. When a given length scale is reached, viscous and/or resistive dissipation processes turn on, thus enabling the heating of the plasma.

To date, two arguments support the mechanism of resonant absorption being operative in the observed waves. First, theory predicts damping time and spatial scales that are compatible with those observed. Second, resonant absorption is a frequency selective process [129], with low-frequency waves being less damped in front of high-frequency waves. A comparison between the outward and the inward power ratio measurements by Tomczyk et al. [63] using CoMP and the theoretical predictions by Verth et al. [130] strongly supports the idea that resonant absorption might be producing the observed in situ energy loss.

Another mechanism studied in detail is Alfvén wave turbulence [131–133]. According to the models, photospheric foot-point motions transported along loops get amplified and reflected at the transition region boundary, thus producing a pattern of counter-propagating perturbations. Their complex interactions lead to the creation of small scales and dissipation. Recent numerical experiments show that, if that is the case, the obtained dissipation rates are able to reproduce chromospheric and coronal heating requirements. The heating rate scales with the magnetic field strength and with the loop length. The modelling of the lower atmosphere is also an important factor. Realistic lower atmosphere modelling favours AC heating in front of DC dissipation.

Observations are now starting to offer support in favour of this process. De Moortel et al. [134] analyse CoMP Doppler shifts and their frequency distribution at both sides of a large trans-equatorial loop system and find that an excess of high-frequency power is present at the apex. This is interpreted as being due to low- and mid-frequency wave energy cascading down because of Alfvén wave turbulence. A recent comparison between observed non-thermal velocities and predictions from theoretical models in coronal loops by Asgari-Targhi et al. [135] leads to the conclusion that Alfvén wave turbulence is indeed a strong candidate for explaining how the observed loops are heated.

Significant efforts are being done in the design and application of numerical models, such as the Alfvén Wave Solar Model (AWSoM) [136,137], that address the coronal heating problem and the solar wind acceleration from a global perspective. These models incorporate Alfvén wave turbulence. The result is a remarkable capacity to reproduce the observed EUV emission and produce solar wind predictions.

Physical processes discussed above are mostly restricted to MHD models. However, the MHD approximation is possibly part of the difficulty to advance the wave heating question, as the actual heating processes must occur at kinetic scales and hybrid or fully kinetic analyses need to be further developed. The ideal MHD model should fail when applied to the lower solar atmosphere where partial ionization effects are important [138] and effects such as ambipolar diffusion, collisions and non-magnetization become non-negligible. For instance, Vranjes & Poedts [139] have pointed out that, because of the presence of non-magnetized ions in the photosphere, Alfvén waves cannot be efficiently generated in, nor travel through, this region. More importantly, the wave energy flux through the photosphere becomes orders of magnitude smaller, compared with the ideal case, when the effects of partial ionization and collisions are consistently taken into account [140]. On the other hand, Song & Vasyliunas [141] and Tu & Song [142] have studied Alfvén wave propagation and heating through plasma-neutral collisions and find that they can
generate sufficient heat, with most of the heat deposited as required at lower altitudes. To clarify these aspects is crucial in view of the recent observations of the chromospheric dynamics with IRIS [143].

4. Wave energy content, damping and dissipation

Our estimates for the energy content of the observed waves are based on simple expressions for bulk Alfvén waves in homogeneous plasmas. Goossens et al. [144] have shown that the energy flux computed with the well-known expression for bulk Alfvén waves could overestimate the real flux by a factor in the range 10–50. Van Doorsselaere et al. [145] provide approximations to the energy propagated by kink modes in an ensemble of flux tubes finding a correction factor for the energy in kink waves, compared with the bulk Alfvén waves, in terms of the density filling factor.

The full energy balance will be given by the Poynting theorem, which relates the amount of electromagnetic energy variation, the fraction of that energy that propagates with the wave and the remaining energy that is dissipated. The assessment of these quantities requires the measurement of the spatial and temporal variation of both velocity and magnetic field perturbations. This being a challenge, a first step is the detailed analysis of the flow of energy and its spatial distribution in the waveguide. For resonance absorption in a flux tube model, this has been done by Arregui et al. [146], in the prominence context. The results show that ideal damping is produced by the jump of the radial component of the Poynting vector, which produces the energy flow into the dissipative layer from both sides of the resonance. Once in the resonant layer, energy flows along the field and resistivity, wave fields, and the created currents determine the energy dissipation and its spatial distribution. Quantification and comparison with observed field-aligned brightness variations remain to be done.

In numerical experiments, all magnetic and plasma perturbations are available. The kinetic, magnetic and total energy and their spatial and temporal evolution can readily be analysed. This is not the case with observations, where the energy contained in the magnetic field is hidden and has to be inferred. The numerical simulations by De Moortel & Pascoe [147], for transverse wave propagation in a multi-stranded waveguide model, have shown that line of sight integration effects can be important when evaluating the energy content of a wave from only the measured kinetic energy. They have also shown that a relevant amount of energy that is present in the form of magnetic perturbations in the simulation would be out of observational reach. The ‘visible’ energy would only be in the 3–10% range of the total energy.

As first noted by Lee & Roberts [148], wave damping is not a synonym for wave dissipation. The two processes might operate at different time and spatial scales and, for example, viscosity should not immediately dissipate the enhanced local oscillations. The matching of the two time scales is important. Consider for instance resonant damping, phase mixing and resistive dissipation. A calculation by J. Terradas (2011, personal communication) shows that, under coronal conditions (magnetic Reynolds number of the order of 10^{12}), the observational consequence, the attenuation of the motion, might give little information on the concealed physical process, the dissipation of the wave energy. For instance, for a typical coronal loop subject to resonant damping because of a transverse inhomogeneity length scale of \( l/R = 0.1 \) (in units of the tube radius \( R \)), the damping time per period is \( \tau_d/P \sim 13 \). The time at which dissipation becomes important, once phase mixing has sufficiently decreased the characteristic transverse spatial scales, is \( \tau_{ra}/P \sim 170 \). The thicker the non-uniform layer, the faster is the damping, \( \tau_d/P \sim 3 \) for \( l/R = 0.5 \), but the slower the small-scale creation due to phase mixing. We need therefore to wait longer for resistivity to become important, \( \tau_{ra}/P \sim 500 \). This means that energy dissipation would only be efficient at very large times, and there would be no heating during the observed oscillations.

The numerical experiment on resonant absorption by Ofman et al. [149] analysed the plasma response and the creation of small-scale structure because of the nonlinear response of the density, but the assumed energy balance was not entirely self-consistent. More recent numerical
experiments by Terradas et al. [120,121] have analysed the first two stages (damping and phase mixing), in complex models including the nonlinearity, but no heating was computed. A detailed numerical study of the full process of wave damping, phase mixing, resistive or viscous dissipation and plasma heating, including the forward modelling of the synthetic plasma response, has not yet been undertaken. Such an experiment would prove/disprove the viability of heating by resonant absorption and phase mixing, by producing spatial and temporal heating profiles to be compared with observations.

5. Confronting theory and observations

Most of the efforts in the area have focused on the design and use of better instrumentation for data acquisition and analysis and the improvement of theoretical and numerical modelling. The two developments being important, further advance will only come from the development of informative tools for the combination of data and theory to assess the physical conditions and processes that operate in the solar atmosphere. The combination of observed data and theoretical results to infer physical properties and mechanisms is not an easy task. One needs to solve two problems simultaneously. In the forward problem, we prescribe theoretical models and parameters (the causes) and analyse the theoretical wave properties (the consequences). In the inverse problem, we try to infer the causes (the unknown physical conditions/mechanisms) from the consequences (the observed wave properties). The discovery of wave dynamics pervading the solar atmosphere has led to considerable efforts in both forward and inverse modelling.

(a) Forward modelling of observational signatures

The creation of synthetic imaging and spectral observational signatures for direct comparison with data is starting to be used as a means to extract conclusions about the goodness of our wave models (e.g. [150–154]).

Forward modelling is essential to solve wave-type identification issues. Van Doorsselaere et al. [155] raised the concern as to whether the observed transverse motions should be interpreted as fast kink waves or Alfvén waves. The issue was clarified by Goossens et al. [156], who pointed out that the observed dynamics is the result of Alfvén waves modified by the radial structuring of the plasma density. Their nature being highly Alfvénic, these authors refer to them as surface Alfvén waves. Kink and torsional Alfvén waves in magnetic flux tubes produce distinct observational signatures [157]. Besides the naming, which is of secondary importance, it is crucial to consider an adequate description of the structures in which the waves propagate, in particular when it comes to seismology and energy budget calculations using the observed waves [158]. Observed tilted features in spectrograms have been interpreted as signatures of torsional Alfvén waves [87]. The forward modelling by Goossens et al. [157] suggests that tilted features may also arise under the kink wave interpretation, depending, among other factors, on the line of sight integration.

Other applications of forward modelling techniques include the fitting of wave properties, such as wave periods and amplitudes, from the comparison of alternative data realizations and observed data [59]; the assessment of the wave or flow nature of observed quasi-periodic disturbances from Doppler shifted spectra [159–163]; or the analysis of the nature of shock distributions along coronal loops as being due to either nonlinear Alfvén waves or nanoflare heating [164].

The potential of forward modelling should still be further exploited to compare alternative wave models producing distinct observational consequences and alternative heating mechanisms. Different heating models are known to produce distinct heating profiles along waveguides. For instance, viscous heating is determined by the velocity perturbation profile along waveguides, leading to apex heating, whereas resistive dissipation is determined by magnetic field perturbations, thus producing foot-point heating [165]. The recent application of forward modelling techniques to waves by Antolin & Van Doorsselaere [166], Antolin et al. [167] and De Moortel et al. [168] shows great promise.
(b) Seismology inversion from observed wave properties

In parallel with theoretical and observational advances, the field of solar atmospheric seismology has emerged. The term refers to the study of the physical conditions in solar atmospheric magnetic and plasma structures from the analysis of the observed wave properties. It was first suggested by Uchida [56] and Roberts et al. [57] in the coronal context, and by Roberts & Joarder [169] in the prominence context. The aim is to increase our knowledge about the complicated structure and dynamics of the solar atmosphere and is based on the fact that the properties of the observed oscillations are determined by the plasma and magnetic field properties.

Solar atmospheric seismology has experienced a great advancement (see [24,50,52,53,170,171] for reviews). This was made possible by the increase in the quantity and quality of wave activity observations and the refinement of theoretical MHD wave models.

Local seismology, based on the use of simplified models for the magnetic and plasma structuring of tube-like waveguides and observations of the period and damping of standing and propagating waves, enabled us to obtain information on the magnetic field strength [172,173], the coronal density scale height [174,175], the magnetic field expansion [176], the Alfvén speed [81,177] or the longitudinal and cross-field magnetic field and density structuring [178–181]. This information is crucial to evaluate the time and spatial scales for wave damping and dissipation processes ($\S$4). The simplicity of the models considered so far imposes a limited applicability to real Sun conditions. As pointed out by De Moortel & Pascoe [182], current inversion results need to be re-examined in light of the outcome from more involved numerical models, in order to determine their accuracy. As the number of observed events increases, statistical analyses—both frequentists [183] and Bayesian [184]—offer valuable wide-ranging information.

Seismology diagnostics will play a key role when it comes to quantifying the energy carried out and dissipated by MHD waves, but the area needs to mature by considering models more akin to the structure and dynamics of solar plasmas. Aiming at the inference of magnetic field and plasma properties in spatially extended regions of the atmosphere is another important topic. The available maps with measured wave characteristics by, for example, Tomczyk & McIntosh [64] offer a way to perform global seismology, but few global inversion techniques have been devised so far (e.g. [185,186]).

(c) Bayesian analysis

Because any inversion process has to be done under conditions in which information is incomplete and uncertain, probabilistic inference is the next step. The solution to a probabilistic inversion problem is given in the form of statements about the plausibility of parameters/models [187]. Bayesian analysis then enables us to quantify the degree of belief in these statements, by measuring to what extent they are supported by information on observed data. This is done by application of Bayes’ rule [188], a mathematical theorem that teaches us how to combine prior knowledge with the likelihood of obtaining a data realization as a function of the unknown parameters/models to compute the so-called posterior, which accounts for what can be said about a parameter or model, conditional on data.

The concept can be applied to the problems of parameter inference and model comparison. In parameter inference, the posterior is computed for different combinations of parameters and then one marginalizes the full posterior to obtain information about the parameter of interest. In model comparison, the ratio of posteriors for alternative models is computed to assess which one better explains observed data. Bayesian analysis is producing successful results in the physical sciences (e.g. [189]) and other areas of space science and astrophysics research, such as cosmology or exoplanet detection [190–192]. Applications to solar physics are scarce, but the few applications to the analysis of solar oscillations [193], to the comparison of heating profiles along loops [194] or to inference and model comparison problems in solar atmospheric seismology [195,196] show great promise.

The main advantages of the Bayesian methodology are that inference is made using self-consistently all the available information in prior knowledge, observed data and modelling,
additionally obtaining the correct propagation of uncertainty. In model comparisons, the method enables us to assess quantitatively which model among competing alternative explanations better explains observed data. The development and application of Bayesian techniques can shed light on the problems of wave-type identification, wave versus flow interpretation and the comparison between alternative wave damping and heating scenarios.

6. Conclusion

Physical explanations based on waves have been at the forefront of coronal heating discussions since the problem came into existence. At that time, evidence about the presence of waves and oscillations was absent and the role magnetic fields could play was largely unknown. The solar atmosphere is now one of the best observed and studied astrophysical systems and the presence of waves and oscillations is beyond question. Waves and oscillations are found and analysed in structures with very different physical conditions. Multi-wavelength and multi-scale imaging and spectroscopic observations enable us to measure their properties with increasing precision.

Wave heating theories remain plausible, but most current models are still too simple to be applicable to the real Sun. Future models should consider the highly structured and dynamic nature of the coronal plasma. The observed waves act as energy carriers and can transport energy to small scales, where the heating processes occur concealed from observational scrutiny. The theoretical analysis of such kinetic processes should be pursued to derive their large-scale observational consequences. Similarly, the consequences of partial ionization and collisions for wave properties and their energy dissipation need to be quantified. These observational consequences need to be fully developed, by forward modelling of synthetic data and comparison with observations. As alternative explanations for the same phenomena arise, there is a need to devise model comparison tools to assess the performance of alternative wave energy transport and dissipation mechanisms in explaining data. Similar methods need to be formulated to confront wave-based mechanisms with other explanations, such as mass flows or nanoflare heating.

Wave energy seems to fill the solar atmosphere. Enough energy is available at the photospheric level and part of it is transmitted above. We do not know exactly how and to what extent. Future observations should concentrate on tracking the flow of energy across different regions of the atmosphere. The amount of energy in magnetic perturbations is hidden. Obtaining estimates of the energy in small scales and dealing with line of sight integration effects represent a challenge for data analysis. The solution is to combine the appropriate modelling with the application of forward modelling and inference tools.

Advances in both theory and observations have again placed waves at the forefront of the discussion. The design and application of tools for comparison between theory and observations is essential. Forward modelling is in an early development stage and seismology is in its infancy. The development of a self-consistent methodology to combine information from theory and data, such as Bayesian analysis, is still waiting to be developed and applied to this area. This would enable us to establish what can be plausibly said about the physical conditions and processes operating in the observed wave dynamics, by using all the available information.

Acknowledgements. I am grateful to Ineke De Moortel and Philippa Browning for the invitation to speak at the Royal Society New Approaches in Coronal Heating Discussion Meeting. Many thoughts here presented are the result of an exchange of ideas with Jaume Terradas, Ineke De Moortel, José Luis Ballester, Ramón Oliver, Marcel Goossens, Toshifumi Shimizu and Andrés Asensio Ramos, over recent years. I received substantial support from Agurtzane Arregi.

Funding statement. This work was supported by a Ramón y Cajal Fellowship and project AYA2011-22846 from the Spanish Ministry of Economy and Competitiveness (MINECO).

References


