A&A 400, 355–367 (2003)
DOI: 10.1051/0004-6361:20021887

Astronomy & Astrophysics

Solar coronal heating by relaxation events

P. K. Browning[1] and R. A. M. Van der Linden[2]

[1] Department of Physics, UMIST, Manchester, UK
e-mail: Philippa.Browning@umist.ac.uk
[2] Royal Observatory of Belgium, Ringlaan 3, 1180 Brussels, Belgium
e-mail: Ronald.VanderLinden@oma.be

Received 11 November 2002 / Accepted 20 December 2002

Abstract. A coronal heating model is proposed which predicts heating by a series of discrete events of various energies, analogous to the observed range of events from large scale flares through various transient brightening phenomena down to the often discussed "nanoflares". We suggest that an energy release event occurs when a field becomes linearly unstable to ideal MHD modes, with dissipation during the nonlinear phase of such an instability due to reconnection in fine-scale structures such as current sheets. The energy release during this complex dynamic period can be evaluated by assuming the field relaxes to a minimum energy state subject to the constraint of helicity conservation. A model problem is studied: a cylindrical coronal loop, with a current profile generated by slow twisting of the photospheric footpoints parameterised by two values of α (the ratio of current density to field strength). Different initial α profiles, corresponding to different footpoint twisting profiles, lead to energy release events of a wide range of magnitudes, but our model predicts an observationally realistic minimum size for these events.

Key words. Sun: corona – Sun: magnetic fields – magnetohydrodynamics (MHD) – plasmas

1. Introduction

Explaining the high temperature of solar coronal plasma remains one of the outstanding unsolved problems for solar physicists (Browning 1991; Jordan 2000; Mandrini et al. 2000; Narain & Ulmschneider 1996; Zirker 1993). While it has been long accepted that the heating mechanism is associated with the strong coronal magnetic field, the details remain uncertain, the essential difficulty being that, in the highly conducting coronal plasma, all effective mechanisms of energy dissipation rely on the existence of small length scale structures which are unobservable. One class of heating mechanisms involves dissipation of MHD waves; the alternative, which we follow here, requires dissipation of magnetic energy associated with quasi-steady currents. Free magnetic energy is stored in the corona in the form of currents, which in the low β corona are field-aligned. These currents may arise as new flux emerges which is already twisted (Leka et al. 1996; Portier-Fozzani et al. 2001) as well as being due to photospheric footpoint motions. Recent observations from SOHO and TRACE (Golub et al. 1999) suggest that the corona is very dynamic, and thus heating may be a result of many transient heating events rather than a smooth ongoing process (Ireland et al. 1999). This is compatible with the picture of heating by "nanoflares" (Parker 1988), whereby the corona is heated by flare-like events with a broad range of energies – flares having the largest energy, ranging down through observable smaller events such as microflares and transient brightenings, down to currently unobservable nanoflares. One key issue is then to determine whether the total energy release by these small events is sufficient to supply the energy losses of the hot coronal plasma. This depends on the relative frequency of such events, and much effort recently has been devoted to determining the occurrence rate as a function of energy (Aschwanden 1999; Berghmans et al. 1998; Katsukawa & Tsuneta 2001; Parnell & Jupp 2000).

A strong candidate for the actual mechanism of energy dissipation is magnetic reconnection (Priest & Forbes 2000), a process which allows conversion of magnetic energy to heat on time scales that can, in the case of fast driven reconnection, be as fast as Alfvénic. Much work on this subject has focused on individual reconnection events in localised current layers or near magnetic neutral points. However, in the complex geometry of a coronal loop, there may be many sites of reconnection, and it is not easy to calculate the dynamics or the energetics in this case.

The energy release due to many localised reconnections can be predicted using relaxation theory (Ortolani & Schnack 1993; Taylor 1974, 1986, 2000) which was first developed to explain the generation of reversed field in laboratory Reverse Field Pinch devices (Taylor 1974, 1986) and later applied to the solar corona (Browning 1988; Browning & Priest 1986; Browning et al. 1986; Dixon et al. 1988; Heyvaerts & Priest 1984; Lothian & Browning2000; Vekstein et al. 1993; Wolfson et al. 1994).

Send offprint requests to: P. K. Browning,
e-mail: Philippa.Browing@umist.ac.uk

In a highly conducting plasma, the magnetic helicity

$$K = \int_V \boldsymbol{A} \cdot \boldsymbol{B} \mathrm{d}V, \tag{1}$$

is conserved. Here $\boldsymbol{B}$ is the magnetic field and $\boldsymbol{A}$ is a vector potential so that $\boldsymbol{B} = \nabla \times \boldsymbol{A}$. Helicity measures the total self-linkage of the magnetic field lines, or the degree of twistedness, and is now widely used in a solar context (Brown et al. 1999). Whereas in a perfectly conducting plasma the helicity of every magnetic flux tube is conserved, in the case of small resistivity (allowing reconnections between flux tubes) only one invariant survives – the global magnetic helicity. When dissipation occurs on small length scales through reconnection, it may be shown that helicity is much less affected by resistive dissipation than magnetic energy (Berger 1984; Browning 1988; Taylor 1986). Then, the plasma tends to relax to a state of minimum magnetic energy whilst conserving its helicity, which is the well-known constant-α or linear force-free field defined by

$$\nabla \times \boldsymbol{B} = \alpha \boldsymbol{B}, \tag{2}$$

(Woltjer 1958). Any energy in the initial field above that of the constant-α state is usually assumed to be dissipated as heat. Relaxation theory bypasses the dynamics by which the minimum energy state is attained – this may involve three-dimensional turbulence with localised reconnections. The microscopic processes by which the energy is thermalised may include Ohmic dissipation within current sheets, viscosity (as the evolving magnetic field generates plasma flows) or even non-MHD processes such as ion cyclotron waves. This tendency to relax to a minimum energy state is well confirmed by experiments in laboratory plasmas (Heidbrink & Dang 2000; Ji et al. 1995; Taylor 1986) as well as by numerical simulations of coronal plasmas (Hu et al. 1997; Kusano et al. 1994).

In the solar corona, the picture is that the coronal magnetic field is stressed by displacements of its photospheric footpoints. If these are slow compared with the travel time of an Alfvén wave, the field evolves into a nonlinear force-free state

$$\boldsymbol{j} \times \boldsymbol{B} = 0 \Longrightarrow \nabla \times \boldsymbol{B} = \alpha(\boldsymbol{r}) \boldsymbol{B}, \tag{3}$$

since the coronal magnetic field is strong and the plasma β is low; this then relaxes, releasing some energy as heat, to a minimum energy state (Heyvaerts & Priest 1984). Depending on the relative timescales of driving and relaxation, there is a competition between the processes so that a fully relaxed state may not be achieved but the field is rather partially relaxed. In the following, we calculate energy release as the difference between the energy of the initial field and that of the constant-α field to which it relaxes, assuming that all this energy is available for heating. Magnetic reconnection itself converts magnetic energy into both thermal energy and plasma kinetic energy. Indeed during the relaxation process it may be expected that there will initially be a build-up of kinetic energy due to reconnection-generated flows, but it is likely that most of this will be subsequently dissipated by viscosity or other processes. This is indeed implicit in relaxation theory, since the final state is static. It should be borne in mind that the relaxation approach gives an upper bound on energy release, since complete relaxation may not be attained in practice.

In the following, we describe a coronal heating model that is based on the scenario of a series of heating episodes. Many coronal heating theories predict a quasi-continuous heating rate, but in this paper we generate heating by a series of discrete events. Such events are triggered by external stressing of the coronal magnetic field by photospheric footpoint motions. The field is perturbed by footpoint motions on a timescale slower than Alfvénic so that the coronal magnetic field has time to adjust to force-free equilibrium. Sometimes this equilibrium may be unstable, and a period of rapid dynamic evolution ensues – during such episodes, energy may be dissipated and the plasma is heated. We suppose that it is ideal MHD instabilities that trigger heating events. This is inspired by 3D numerical simulations which show that fine-scale current structures naturally arise in the nonlinear phase of kink instability (Baty & Heyvaerts 1996; Velli et al. 1997; Lionello et al. 1998; Arber et al. 1999; Baty 2000; Gerrard et al. 2001). Relaxation theory is used to determine the energy release, which could otherwise only be calculated with three-dimensional resistive MHD simulations. This approach is quite advantageous because numerical simulations are complex and computationally intensive, and even so cannot at present (or in the foreseeable future) resolve the fine length scales expected to be involved in magnetic dissipation at the high values of the Lundquist number in the solar corona. Our model predicts ab initio coronal heating by a series of discrete events of varying sizes, with smaller events (which may be identified with nanoflares) being more frequent. A further key prediction of our model is that there is naturally a minimum size of heating event.

The model and its underlying assumptions are explained in the next section. In Sect. 3 the results of the ideal MHD stability calculation are presented, which determines the stability threshold at which a heating event is triggered. The calculation of the energy release, using helicity conservation, is presented in Sect. 4. Finally the results are summarised and discussed in Sect. 5.

2. The model

We are concerned with calculating energy dissipation in a coronal loop that is driven by slow displacements of the photospheric footpoints. There are two key assumptions to our model. The first is that an energy dissipation event is triggered by an ideal MHD instability. This may seem counter-intuitive, since ideal MHD modes intrinsically cannot dissipate magnetic energy. However, the dissipation is supposed to occur in the nonlinear phase of evolution – a disturbance is triggered when the ideal stability threshold is crossed, and as the disturbance grows in magnitude (on Alfvénic timescales) small scale current concentrations are generated where resistivity can rapidly dissipate energy. The flows that are generated as an ideal mode grows also may dissipate viscously – hence there is an indirect energy conversion route, from magnetic through kinetic to thermal. There is strong evidence of the formation of current sheets and the generation of fast reconnection, leading to heating, from 3D magnetohydrodynamic simulations of nonlinear kink instability in line-tied loops (Baty 2000; Baty & Heyvaerts 1996; Gerrard et al. 2001; Lionello et al. 1998;

Velli et al. 1997). Conversely, resistive instabilities such as tearing modes (Furth et al. 1963) are not good candidates for heating events since their timescales are far too slow. For example, in a typical coronal loop of length 20 000 km and radius 2000 km with field strength 0.005 T and density 5×10^{14} m^{-3}, the Alfvén travel time τ_A along the loop is about 4 s. Taking a temperature of 2 million K, the classical resistive diffusion time for fields across the loop (τ_d) is about 10^{13} s, so that the growth rate of tearing modes $\approx (\tau_A \tau_d)^{1/2} \approx 6 \times 10^6$ s.

The second assumption is that, once a magnetic field becomes unstable, after a period of time dependent behaviour, it will ultimately relax towards the appropriate minimum energy state, a constant-α field. The energy release can then be determined straightforwardly through relaxation theory, without recourse to three-dimensional MHD simulations. The principles outlined here could be applied quite generally, but to be specific we develop a model of a cylindrical coronal loop. This allows much of the calculation to be performed analytically. Curvature and expansion are ignored (see Lothian & Browning 2000 for a discussion of the effects of the latter), so the loop is taken to be a cylinder of length L and radius R. All equilibrium field quantities are one-dimensional, depending only on the radial coordinate r in the cylindrical coordinate system (r, θ, z), which we use in the rest of this paper. The magnetic field is perturbed by slow photospheric footpoint displacements, which cause the field to evolve through a sequence of force-free equilibria (3) with a radial profile of α. We parameterise this profile by a two layer model in which the loop has an inner core (radius R_c) and an outer sheath; within each region, α is constant. Thus

$$\begin{aligned} \alpha &= \alpha_1, \qquad r \le R_c, \\ \alpha &= \alpha_2, \qquad R_c < r < R. \end{aligned} \tag{4}$$

In principle the values of α_1 and α_2 are determined by the field line twist profile, since the angle of twist of the field lines from one end of the loop to the other is given by

$$\varphi = \frac{LB_\theta}{RB_z}. \tag{5}$$

The twist profile in turn depends on the history of photospheric footpoint motions as well as on any current pre-existing in the loop when it emerged and on currents remaining after previous relaxation events. This two parameter model, previously used with some extra constraints to model coronal loops (Melrose et al. 1994) and spheromaks (Brennan et al. 1999), represents the wider space of α distributions; as well as smoothly twisted profiles it allows also distributions such as an inner core twisted one way surrounded by a layer of reverse twist. The field components, for $\alpha_1, \alpha_2 > 0$, are given by the well known Bessel function solutions to (2)

$$\begin{aligned} B_z &= B_1 J_0(\alpha_1 r), B_\theta = B_1 J_1(\alpha_1 r), \qquad (r \le R_c), \\ B_z &= B_2 J_0(\alpha_2 r) + C_2 Y_0(\alpha_2 r), \\ B_\theta &= B_2 J_1(\alpha_2 r) + C_2 Y_1(\alpha_2 r), \qquad (r > R_c), \end{aligned} \tag{6}$$

where B_1, B_2 and C_2 are constants. The field must be continuous at the core radius R_c, although the current is in general discontinuous. Matching across this surface, we find the coefficients in the outer layer

$$\begin{aligned} C_2 &= B_1 \frac{J_0(\alpha_1 R_c) J_1(\alpha_2 R_c) - J_1(\alpha_1 R_c) J_0(\alpha_2 R_c)}{\Delta}, \\ B_2 &= B_1 \frac{J_1(\alpha_1 R_c) Y_0(\alpha_2 R_c) - J_0(\alpha_1 R_c) Y_1(\alpha_2 R_c)}{\Delta}, \end{aligned} \tag{7}$$

where

$$\Delta = Y_0(\alpha_2 R_c) J_1(\alpha_2 R_c) - Y_1(\alpha_2 R_c) J_0(\alpha_2 R_c) = \frac{2}{\pi \alpha_2 R_c}, \tag{8}$$

using a Bessel function identity (Abramowitz & Stegun 1965). In the case when the sign of α is reversed throughout, the field is simply a mirror image with the same axial field and reversed azimuthal field. We are interested in the possibility of the inner and outer layers having α with different signs; so for complete generality we consider α_1 always positive but α_2 both positive and negative. The latter case may be dealt with in (6) by replacing α_2 by its modulus and reversing the sign of B_θ; the matching conditions (8) for $\alpha_2 < 0$ become

$$\begin{aligned} C_2 &= B_1 \frac{J_0(\alpha_1 R_c) J_1(|\alpha_2| R_c) + J_1(\alpha_1 R_c) J_0(|\alpha_2| R_c)}{\Delta}, \\ B_2 &= -B_1 \frac{J_1(\alpha_1 R_c) Y_0(|\alpha_2| R_c) + J_0(\alpha_1 R_c) Y_1(|\alpha_2| R_c)}{\Delta}, \end{aligned} \tag{9}$$

where Δ is given by (8) with α_2 replaced by its modulus. Typical fields are shown in Fig. 1. For convenience in the following, we define functions $F_{0,1}$ as

$$F_{0,1}(x) = J_{0,1}(x) + \frac{C_2}{B_2} Y_{0,1}(x), \tag{10}$$

so that $B_\theta = B_2 F_1(\alpha_2 r)$, $B_z = B_2 F_0(\alpha_2 r)$ in the outer layer.

3. Stability results

The first step is to calculate the linear stability of the equilibria described in Sect. 2 to ideal MHD modes. In our case of a force-free twisted loop, the dominant instability mode is a current-driven kink instability with azimuthal mode number $m = 1$ (Van der Linden & Hood 1999). It is well known that the stability of coronal loops is constrained by the line-tying of the photospheric footpoints of the field (Hood 1992), which can be modelled by requiring that the velocity perturbation must vanish at the photospheric boundaries – in our case, $z = 0, L$. We assume a perfectly conducting wall at the outer radius R, which is a common assumption in stability calculations even in solar physics, though this can be rather unrealistic for coronal loops, unless the instability is localised well away from the outer boundary. Stability properties may be altered if the loop is embedded in an external potential field (e.g. Baty 2001) and such models will be considered in future work. The stability threshold depends critically on the aspect ratio L/R, and we take this to be 20, which is representative of typical coronal loops. Again, the effects of varying this will be considered in future work. We also fix the transition from core to outer layer

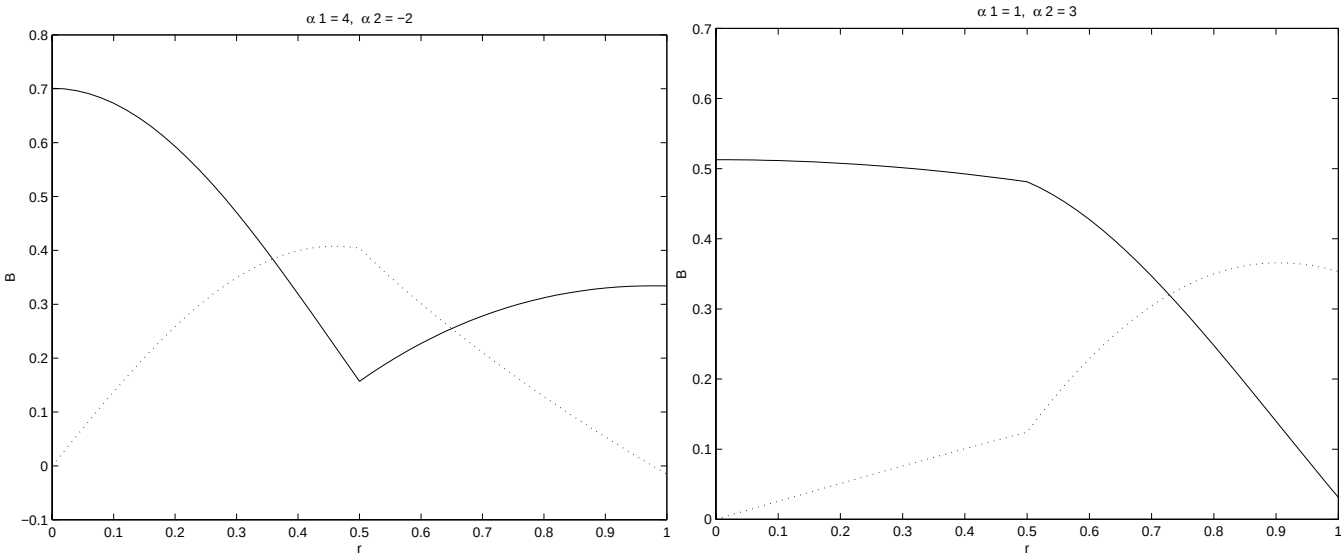

Fig. 1. Profiles of typical equilibrium fields (Eq. (6)) showing axial field (B_z, solid line) and azimuthal field (B_θ, dashed line).

to be at $R_c = 0.5R$ in the current model. In the following numerical results, lengths are scaled with respect to the loop radius by setting $R = 1$ (except where stated otherwise). It is now necessary to determine the marginal stability threshold for kink modes as a function of the current profile – this is a curve in the two-dimensional (α_1, α_2) plane.

For comparison, it is useful to recall that the infinite-length constant-α field, $\alpha_1 = \alpha_2$, becomes unstable at the critical value $\alpha R \approx 3.176$ (Voslamber & Callebaut 1962). With the stabilising effect of line-tying included, the stability threshold for constant-α fields must of course be higher. For simple current profiles, line-tied loops become unstable when a critical twist is exceeded, $\varphi_{\rm crit} \approx 2.5\pi$ (Hood & Priest 1981). This may be interpreted as increasing current being destabilising, since the free energy driving kink instabilities is due to electrical currents. Intuitively one might then expect that increasing α should be destabilising, but we shall see that this is not always the case. Indeed, in a rather different context, where a similar equilibrium model (with pressure added in the core region) was used to model a sharp screw pinch (Goedbloed 1970), it was shown that replacing a vacuum magnetic field ($\alpha = 0$) in the external region with a constant-α force-free field can enhance stability and complete stabilisation may be obtained. To a certain extent, the stabilisation with increasing α can be associated with the existence of reversals in the axial field B_z in our equilibria (6) when α_1 or α_2 are sufficiently large. It could be argued that such reversals do not exist in real coronal loops, and indeed our results are probably most relevant in the region of parameter space in which $B_z(r)$ is single-valued. On the other hand, the photospheric field consists of many adjacent regions of opposite polarity, and we might consider a "loop" with reversed axial field (a region of positive polarity at one footpoint surrounded by a region of negative polarity) as at least a simplified example of the complex field structures that arise when multiple polarity field regions interconnect. For completeness, therefore, a wide range of parameter space is explored here, with interesting results.

The MHD equations are linearised and normal mode solutions are considered. Due to the axisymmetry of the equilibrium, we may study individual Fourier modes in the angular coordinate θ and write the perturbed quantities f in the form

$$f = \tilde{f}(r, z)\mathrm{e}^{im\theta}\mathrm{e}^{\gamma t}.$$

The line-tying boundary conditions prohibit the use of individual modes of a Fourier expansion in z, because these modes are coupled. We wish to determine the normal mode eigenfunctions and growth rates γ for such linear ideal MHD modes, which will have $m = 1$ for the most unstable modes. This is achieved using a numerical code CILTS, similar to the code MALTS described in Van der Linden & Hood (1999), and which was also previously used to investigate instabilities in spheromaks (Brennan et al. 1999). In this code, a bicubic Hermite finite element method is used to discretise the r and z dependencies. The linearised MHD equations are thus reduced to a matrix eigenvalue problem, with eigenvalue γ^2, which is solved iteratively. Our main interest is in the conditions for onset of instability, so we display only the marginal stability threshold curve (along which $\gamma = 0$).

The results are shown in Fig. 2. The constant-α field, represented by the diagonal $\alpha_1 = \alpha_2$ in Fig. 2, becomes unstable at $\alpha R \approx 4.64$. This is larger than the infinite-length threshold ($\alpha R \approx 3.176$) as expected, showing quite a strong effect of line-tying, discussed further below. An interesting comparison can also be made between the case when $\alpha_1 = \alpha_2$ and the situation when the outer layer is a vacuum (or potential) magnetic field ($\alpha_2 = 0$); from Fig. 2 we see that the latter field is always unstable above about $\alpha_1 R = 5.32$. (Although instability first sets in at about $\alpha_1 R = 2.26$, this is not relevant as there is a stable region, corresponding to a section of the "stable wedge" in Fig. 2, above this.) With a simple rescaling ($\alpha_1 \to \alpha_1/2$, $R \to 2R$) of the latter case, regarding the core radius R_c as the bounding radius, we can transform this into a comparison of the stability of a constant-α field surrounded by a conducting wall and the same field imbedded in a large potential field shell

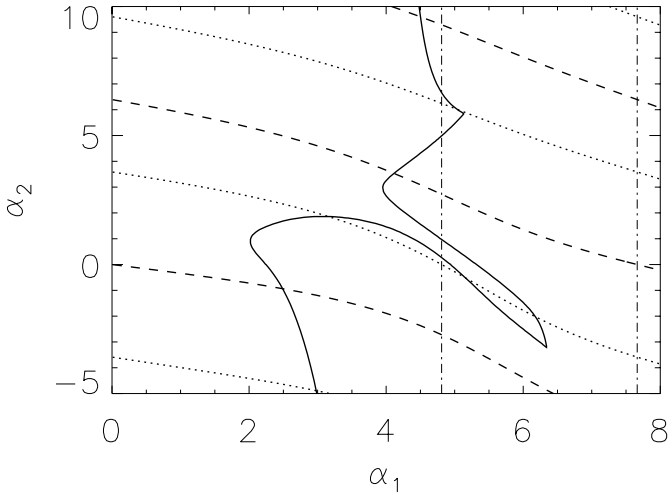

Fig. 2. The marginal stability curve (solid line) for linear $m = 1$ modes in the (α_1, α_2) plane for a line-tied loop with aspect ratio $L/R = 20$. Equilibria to the left of the solid curve are stable, those to the right are unstable. The dotted and dashed curves show the successive zeroes at $r = R$ of B_z and B_θ, respectively, while the vertical dashed lines indicate for which values of α_1 either of the field components are zero at the transition point $r = R_c$. These "nodal" curves and lines indicate how many field reversals of both components are found within the loop.

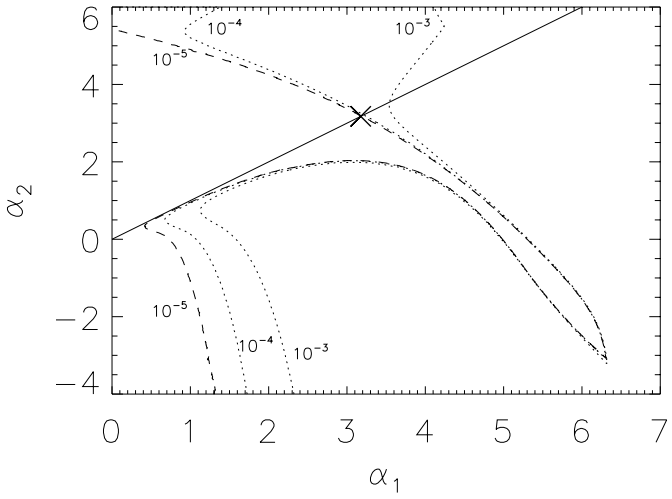

Fig. 3. The stability map in the (α_1, α_2) plane for the same case as in Fig. 2, but for loops of infinite length. Displayed are the σ-stability line for $\sigma = 0.00001$ (dashed curve) and the growth rate contours for $\gamma = 0.001$ and $\gamma = 0.0001$ (dotted lines). The diagonal $\alpha_1 = \alpha_2$ is indicated by the solid line, with the cross indicating the position where the constant α field crosses marginal stability ($\alpha = 3.176$).

with a conducting wall further out. A close fitting conducting wall is well known to have a stabilising effect. Indeed we see that the stability threshold for a field with a conducting shell at R ($\alpha R \approx 4.64$) is greater than that for the transformed field surrounded by a large potential shell ($\alpha R \approx 2.66$), so that the latter field is always unstable whenever the former is, as expected.

Since the kink-like instability modes discussed here are driven by the current, one might intuitively expect that increasing α_2, with fixed α_1, should lead to instability. This is in fact not always the case, and we see from Fig. 2 that when $\alpha_1 = 0$, the field is stable for all values of α_2 in the range investigated (although presumably the field is unstable for even higher values of α_2 but such high values are physically quite meaningless in the present context). This is symptomatic of the rather complex shape of the marginal stability curve, with multiple branches within the range of parameter space presented here. For example, increasing α_2 at a fixed central current to field ratio (α_1) of about $2.5/R$, the field first becomes unstable but then a second stable region is reached. This may be somewhat clarified by noting that the energy integral in the ideal MHD energy principle can be written in such a way that the only destabilising term is proportional to $\alpha B_z B_\theta$ (De Bruyne & Hood 1992). Therefore, only regions where $\alpha B_z B_\theta > 0$ can contribute to driving the instability. Thus, when there are regions of field reversal (as happens at higher values of α), there may be less drive for the instability. For this reason, curves are plotted in Fig. 2 showing the successive zeroes of the field components at the plasma edge. These curves also enable a better visualisation of the characteristics of the equilibria in the map. Note that the rather striking stable "wedge" pointing to the bottom right of the figure coincides rather well with the apparition of the first zero in B_z at the edge $r = R$, creating a non-driving (or damping) shell.

In order to understand further the complex shape of the marginal stability curve, to disentangle the effect of the line-tying, and to validate this calculation, we determine also the stability diagram for a loop of infinite length with the same equilibrium field using the code LEDA (Van der Linden 1991), the results of which are displayed in Fig. 3. In this case, Fourier analysis in the z-coordinate is used, introducing the wave number k by setting $f(r, z) = f(r)\exp(ikz)$. For every choice of equilibrium (α_1, α_2) we then determine the maximum instability growth rate by scanning over the wave number k, including both positive and negative k. The values of α_1 and α_2 are then stepped and the scan repeated to obtain a full stability map in the (α_1, α_2) plane. Close to the marginal stability threshold ($\gamma = 0$), smaller steps are used in the procedure to obtain a sufficiently accurate marginal stability curve. However, a highly complicating factor for the infinite length calculations is the near-singular nature of the eigenmodes as true marginal stability is approached. (In the line-tied case this problem is alleviated due to the coupling of the Fourier modes.) In effect, one or more true singularities always appear when $\gamma = 0$ is reached. To avoid unnecessary cumbersome and lengthy calculations, we employ the concept of σ-stability (Goedbloed & Sakanaka 1974), i.e. we select a (small) threshold value σ for the growth rate that is considered "close enough to marginal stability", and use the contour line of this value as an approximation to the marginal stability curve. In Fig. 3, we show contour lines for three values; from the calculations it is clear that $\sigma = 0.00001$ is a good choice for the σ-stability curve (because near that value the contour lines do not vary strongly anymore). That this value indeed gives a good approximation to the true marginal curve can be deduced from the fact that the diagonal $\alpha_1 = \alpha_2$ crosses the stability threshold at $\alpha_1 = \alpha_2 = 3.176/R$, in full agreement with well-known results (Voslamber & Callebaut 1962). For comparison purposes, also the $\gamma = 0.0001$ and

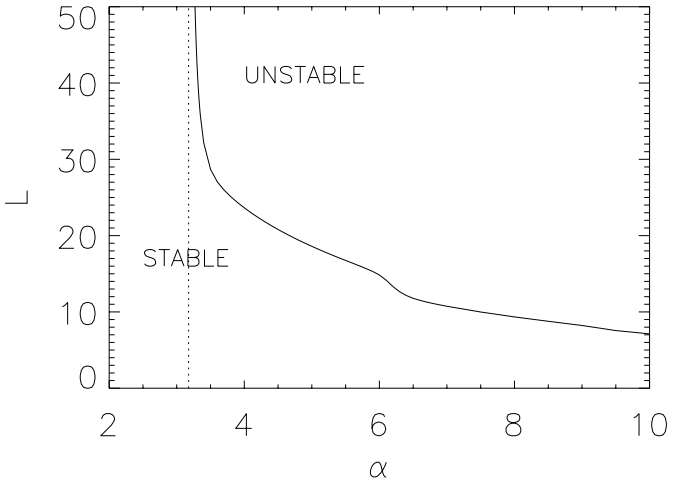

Fig. 4. The critical length for marginal stability of a loop as a function of α for a constant-α field (solid line). The vertical (dashed) line indicates the marginal stability point for infinite loops ($\alpha = 3.176$).

$\gamma = 0.001$ contours are plotted in Fig. 3. It is interesting to see how well the $\gamma = 0.001$ contour corresponds with the marginal stability curve for the line-tied case depicted in Fig. 2. The calculation performed for Fig. 3 gives also some interesting information about the nature of the eigenmodes in the different regions. Thus it is seen, for example, that there exist in fact several branches of instabilities. Having selected $m = 1$, the instability to the left of the stable wedge in Fig. 3 consists of modes with $k < 0$ only, while near the lower half of the stability boundary to the right of the wedge only $k > 0$ instabilities exist. These two branches of instability overlap near the tip of the wedge. In the upper half of the map, both $k < 0$ and $k > 0$ modes exist. We point out that this behaviour agrees perfectly with previous calculations in a similar model for the special cases $\alpha_1 = \alpha_2$ (Voslamber & Callebaut 1962) and $\alpha_2 = 0$ (Chiuderi et al. 1980). In the line-tied calculations presented in Fig. 2, the separation in different values of k no longer exists, but corresponding changes in the eigenfunction structure are found.

Comparing Figs. 2 and 3, it can be seen that as expected, the unstable region in the (α_1, α_2) plane for the line-tied loop is entirely contained within that for the infinitely long loop. Furthermore, the shape of the unstable regions is similar, giving confidence in the (at first surprising) shape of this region found with CILTS. It has also been checked, by varying the length L, that the stability threshold for a constant-α field smoothly approaches that for the infinite length loop as the length increases. This is shown in Fig. 4, which displays the critical loop length for stability for the constant-α field as a function of α. It is interesting to note that rather long loops are needed in order to approach the un-line-tied situation; for example our loop with $L/R = 20$ is unstable only if $\alpha \approx 4.64/R$, significantly higher than the limit $\alpha = 3.176/R$ for an infinite length loop. The at first sight suspicious "kink" in the curve near $\alpha = 6/R$ corresponds to the point where instability is triggered in a second independent subinterval (Giachetti et al. 1977); a similar feature is seen if one plots the growth rate for this equilibrium in an infinite cylinder (Goedbloed & Hagebeuk 1972).

4. Calculation of energy release

In order to calculate the relaxed state for a given initial α profile, we must determine the energy and helicity as functions of α_1 and α_2. A preliminary investigation of these and a relaxation calculation has been carried out by Sutcliffe (2002). First, we calculate the helicity. The loop has field lines crossing its surface, at the photospheric ends ($z = 0, L$) and thus helicity as defined by (1) is not gauge invariant. Relative helicity (Berger & Field 1984) must therefore be used, which essentially is the difference between the helicity of the actual field and that of a potential field with the same normal flux distribution; this is a unique (gauge invariant) and physically meaningful quantity. We use the expression (Finn & Antonsen 1985),

$$K = \int_V (\boldsymbol{A} + \boldsymbol{A'}) \cdot (\boldsymbol{B} - \boldsymbol{B'}) \mathrm{d}V, \tag{11}$$

where $\boldsymbol{B'}$ is a vacuum field whose normal field component matches $\boldsymbol{B}$ on the boundary, with corresponding vector potential $\boldsymbol{A'}$. Following Finn & Antonsen (1985), this can be simplified to a convenient form for axisymmetric fields by writing the field in the form $\boldsymbol{B} = \frac{1}{2\pi r}\nabla\psi \times \hat{\theta} + \frac{I}{r}\hat{\theta}$, where $\psi(r) = \int_0^r 2\pi r B_z \mathrm{d}r$ is the axial magnetic flux and I/μ_0 is the axial current within radius r. The vector potential can be expressed as $\boldsymbol{A} = \frac{1}{2\pi r}\psi\hat{\theta} + \boldsymbol{A}_\perp$ where $\boldsymbol{A}_\perp$ has only r and z components; expressions for the vacuum field and potential are similar except that for our case of a simply-connected volume the vacuum field has no azimuthal component so $I' = 0$. Substituting into (11) gives

$$K = \int_V \frac{I\psi}{2\pi r^2}\mathrm{d}V + \int_V \frac{1}{2\pi}(\boldsymbol{A}_\perp + \boldsymbol{A'}_\perp) \cdot \nabla \times [(\psi - \psi')\hat{\theta}]\mathrm{d}V; \tag{12}$$

after a little manipulation and use of the boundary conditions $\boldsymbol{B} \cdot \hat{\boldsymbol{n}} = \boldsymbol{B'} \cdot \hat{\boldsymbol{n}}$, we obtain

$$K = 2\int_V \frac{I\psi}{2\pi r^2}\mathrm{d}V = 2L\int_0^R \frac{I\psi}{r}\mathrm{d}r, \tag{13}$$

equivalent to the spheromak formula in Finn & Antonsen (1985), simplified for our cylindrical geometry.

For the field given by (6), we find

$$\begin{aligned} I &= rB_1 J_1(\alpha_1 r), && r \le R_\mathrm{c}, \\ I &= r[B_2 J_1(\alpha_2 r) + C_2 Y_1(\alpha_2 r)], && r > R_\mathrm{c}, \end{aligned} \tag{14}$$

and

$$\begin{aligned} \psi &= \frac{2\pi r}{\alpha_1} B_1 J_1(\alpha_1 r), && (r \le R_\mathrm{c}); \\ \psi &= \frac{2\pi R_\mathrm{c}}{\alpha_1} B_1 J_1(\alpha_1 R_\mathrm{c}) + \frac{2\pi}{\alpha_2} B_2 [rF_1(\alpha_2 r) - R_\mathrm{c} F_1(\alpha_2 R_\mathrm{c})], \\ & && (r > R_\mathrm{c}). \end{aligned} \tag{15}$$

(For clarity, throughout this section we only give expressions for the case $\alpha_1, \alpha_2 > 0$; various special cases are given in the Appendix.) Splitting the integral into the ranges over the core and outer layers,

$$\begin{aligned} K = {} & \frac{4\pi L B_1^2}{\alpha_1}\int_0^{R_\mathrm{c}} r J_1^2(\alpha_1 r)\mathrm{d}r + \frac{4\pi L B_2^2}{\alpha_2}\int_{R_\mathrm{c}}^R r F_1^2(\alpha_2 r)\mathrm{d}r \\ & + 4\pi L B_1 B_2 R_\mathrm{c} J_1(\alpha_1 R_\mathrm{c})\left(\frac{1}{\alpha_1} - \frac{1}{\alpha_2}\right)\int_{R_\mathrm{c}}^R F_1(\alpha_2 r)\mathrm{d}r. \end{aligned}$$

The integrations may be performed analytically using standard Bessel function integrals (Abramowitz & Stegun 1965), bearing in mind that F_1 is a linear combination of Bessel functions of the first and second kind, to yield (for $\alpha_1, \alpha_2 > 0$)

$$\begin{aligned} K &= \frac{2\pi L B_1^2}{\alpha_1}\left[R_c^2 J_0^2(\alpha_1 R_c) + R_c^2 J_1^2(\alpha_1 R_c)\right. \\ &\quad \left. -2\frac{R_c}{\alpha_1} J_0(\alpha_1 R_c) J_1(\alpha_1 R_c)\right] \\ &\quad + \frac{2\pi L B_2^2}{\alpha_2}\left[R^2 F_0^2(\alpha_2 R) + R^2 F_1^2(\alpha_2 R)\right. \\ &\quad \left. -2\frac{R}{\alpha_2} F_0(\alpha_2 R) F_1(\alpha_2 R)\right] \\ &\quad - \frac{2\pi L B_2^2}{\alpha_2}\left[R_c^2 F_0^2(\alpha_2 R_c) + R_c^2 F_1^2(\alpha_2 R_c)\right. \\ &\quad \left. -2\frac{R_c}{\alpha_2} F_0(\alpha_2 R_c) F_1(\alpha_2 R_c)\right] \\ &\quad + \frac{4\pi L B_1 B_2}{\alpha_2}\left[F_0(\alpha_2 R_c) - F_0(\alpha_2 R)\right] R_c J_1(\alpha_1 R_c)\left(\frac{1}{\alpha_1} - \frac{1}{\alpha_2}\right). \end{aligned} \tag{16}$$

It has been shown that the same expression is obtained using the gauge-corrected expression for helicity in a torus (Bevir et al. 1985; Bevir & Gray 1980), since the fields at the two ends of the loop are identical so we may regard it as a straightened torus of major radius $L/2\pi$ with an arbitrary cut at some toroidal location. It is easy to see that (16) reduces to the conventional expression (Taylor 1974) for a constant-α cylinder (or an infinite-aspect-ratio torus) with radius $R = R_c$ and $\alpha = \alpha_1$

$$K = \frac{2\pi L B^2}{\alpha}\left[R^2 J_0^2(\alpha R) + R^2 J_1^2(\alpha R) - 2\frac{R}{\alpha} J_0(\alpha R) J_1(\alpha R)\right], \tag{17}$$

if we consider only the first term (corresponding to the core). This expression will be used later for the helicity of a fully relaxed state. Also if we take $\alpha_2 = \alpha_1$, it can be shown that we again retrieve the helicity (17) of a constant-α cylinder, now with radius R. The first term of (16) evidently corresponds to the helicity of the core, and the next two to that of the outer layer; the fourth term, which arises as a "gauge correction", reminds us that helicity of two volumes is not additive, because of the interlinkage of the flux between the two regions (in this case the axial flux of the outer layer interlinks the azimuthal flux of the core, and vice versa). The expressions for negative α_2, and for the cases when inner or outer layers have zero current density, are given in the Appendix.

In order to consider variations of helicity, and other quantities, with the current profile, we must choose what is kept constant. The simplest choice would be to fix B_1, the axial field strength on axis, in which case (16) may be considered directly as a function of α_1 and α_2, but there seems to be little physical justification for this. We choose instead to normalise with respect to the total axial flux

$$\psi_0 = \frac{2\pi B_2}{\alpha_2} R F_1(\alpha_2 R) + 2\pi R_c B_1 J_1(\alpha_1 R_c)\left(\frac{1}{\alpha_1} - \frac{1}{\alpha_2}\right), \tag{18}$$

scaling our results so that $\psi_0 = 1$ (hence B_1 depends on the α profile). This is equivalent to evaluating K/ψ_0^2, as is common (Taylor 1974).

The magnetic energy, using the expressions (6–8) for the field components, is

$$\begin{aligned} W &= \frac{1}{2\mu_0}\int_V B^2 \mathrm{d}V \\ &= \frac{L\pi}{\mu_0} B_1^2 \int_0^{R_c} r\left[J_0^2(\alpha_1 r) + J_1^2(\alpha_1 r)\right]\mathrm{d}r \\ &\quad + \frac{L\pi}{\mu_0} B_2^2 \int_{R_c}^{R} r\left[F_0^2(\alpha_2 r) + F_1^2(\alpha_2 r)\right]\mathrm{d}r. \end{aligned} \tag{19}$$

Making use of further Bessel function integrals (Abramowitz & Stegun 1965), this may be shown to give

$$\begin{aligned} W &= \frac{L\pi}{\mu_0}\left(B_1^2\left\{R_c^2\left[J_0^2(\alpha_1 R_c) + J_1^2(\alpha_1 R_c)\right]\right.\right. \\ &\quad \left.\left. -\frac{R_c}{\alpha_1} J_0(\alpha_1 R_c) J_1(\alpha_1 R_c)\right\}\right. \\ &\quad + B_2^2\left\{R^2\left[F_0^2(\alpha_2 R) + F_1^2(\alpha_2 R)\right] - \frac{R}{\alpha_2} F_0(\alpha_2 R) F_1(\alpha_2 R)\right. \\ &\quad \left.\left. - R_c^2\left[F_0^2(\alpha_2 R_c) + F_1^2(\alpha_2 R_c)\right] + \frac{R_c}{\alpha_2} F_0(\alpha_2 R_c) F_1(\alpha_2 R_c)\right\}\right), \\ &\qquad (\alpha_1, \alpha_2 > 0). \end{aligned} \tag{20}$$

Again, expressions for the various special cases are given in the Appendix.

Armed with these expressions, we can now calculate the energy release during a single relaxation event. A heating event occurs when the footpoint motions cause the field to cross the threshold for ideal MHD instability – the field subsequently evolves dynamically towards the appropriate minimum energy state, releasing any excess magnetic energy as heat. We suppose that the field is in an equilibrium with given α_1 and α_2, which lie somewhere on the marginal stability curve for ideal MHD modes as calculated in Sect. 3. The field then relaxes to a constant-α state with the same helicity, with $\alpha = \alpha(\alpha_1, \alpha_2)$ thus determined by finding the root of

$$K(\alpha) - K(\alpha_1, \alpha_2) = 0. \tag{21}$$

Since all helicities are normalised with respect to the axial flux (18), it is effectively imposed that axial flux is conserved during the relaxation, as expected (this is equal to the photospheric flux at the loop footprints). We do allow the normal field distribution at the ends of the loop to change during relaxation, however. It might be more realistic to fix this, being matched in principle to the photospheric normal field, so that the relaxed state would be two-dimensional (dependent on r and z). However, we note that in such a case, all z variations in field components are confined to narrow boundary layers near the footpoints (Browning & Hood 1989) and the loop structure is predominantly 1D; furthermore, we may assume that reconnections during the non-ideal relaxation phase, may break the connections with photospheric field. Hence, the final relaxed constant-α field in this cylindrical geometry is taken to be the well known Bessel function field (see the first equation in (6))

$$B_z = B J_0(\alpha r), \qquad B_\theta = B J_1(\alpha r), \tag{22}$$

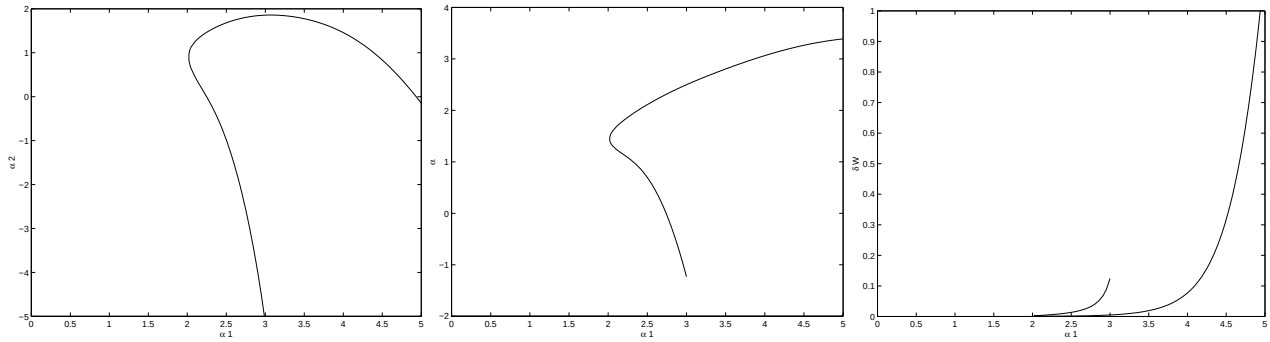

Fig. 5. The effects of relaxation to an axisymmetric constant-α state for the first branch of the marginal stability curve, normalised to $\psi_0 = 1$, $\mu_0 = 1$ and $R = 1$: **a)** the section of the marginal stability curve showing the dependence of α_2 on α_1; **b)** the value of α obtained by helicity-conserving relaxation following the marginal stability curve in **a)**; **c)** the energy release (δW) per unit length due to the relaxation.

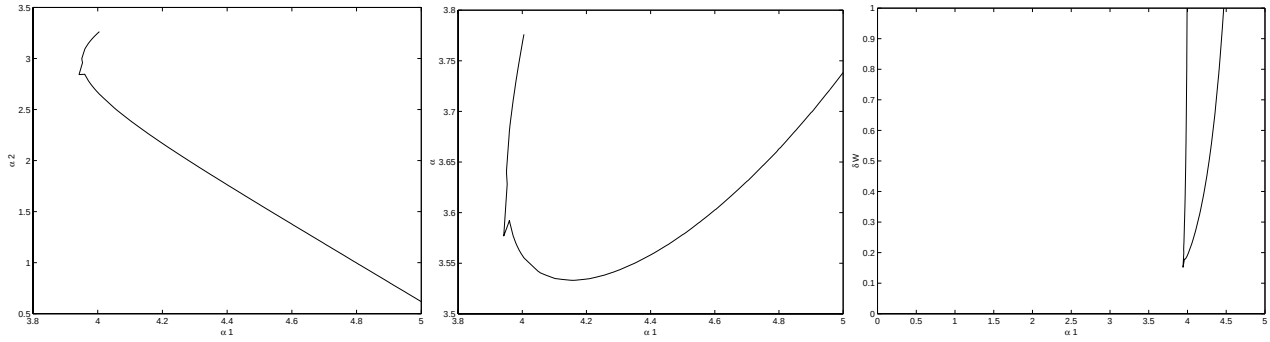

Fig. 6. The effects of relaxation to an axisymmetric constant-α state for the second branch of the marginal stability curve, non-dimensionalised to $\psi_0 = 1$, $\mu_0 = 1$ and $R = 1$: **a)** the section of the marginal stability curve showing the dependence of α_2 on α_1, cut off where quantities normalised to the axial flux diverge; **b)** the value of α obtained by helicity-conserving relaxation following the marginal stability curve in **a)**; **c)** the energy release (δW) per unit length due to the relaxation.

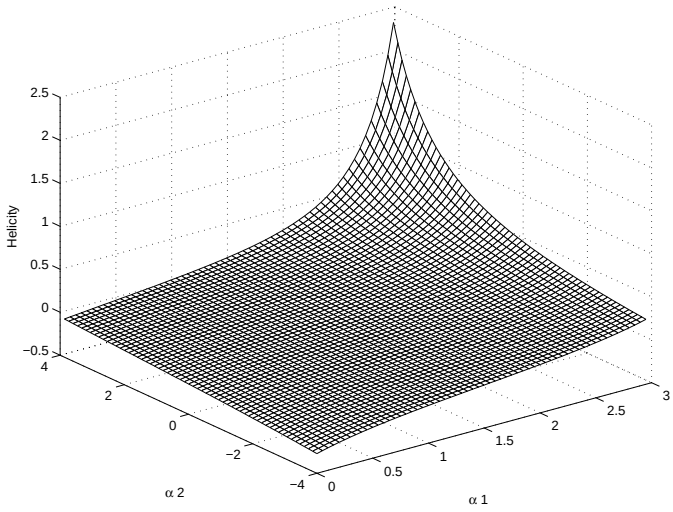

Fig. 7. A surface plot showing the dependence of helicity, normalised with respect to the axial flux ψ_0, on the current profile (α_1, α_2). Note that the helicity diverges along a curve towards the upper right hand corner.

with helicity (17) (some modification to this is considered later).

The root α for (21) is found numerically, using the expressions (16) and (17) for the initial and final helicities (with special cases as appropriate), normalised such that $\psi_0 = 1$. The energy release during the heating event is then given by the difference between the energies of the initial unstable state and the final constant-α field, using the expression (20) for the energy,

$$\delta W(\alpha_1, \alpha_2) = W(\alpha_1, \alpha_2) - W(\alpha(\alpha_1, \alpha_2)). \tag{23}$$

This must always be positive (a useful numerical check) since the constant-α field is a minimum energy state – this indeed turns out to be the case. This process is repeated for all points along the marginal stability curve, so that we obtain the energy release for a set of values of α_1 and α_2. We initially limit this calculation to what seems to be a reasonably realistic range of initial α values such that $|\alpha_{1,2}R| < 5$; in this region the marginal curve (Fig. 2) splits into two disjoint branches, the results for which are plotted separately in Figs. 5 and 6. For convenience, the relaxed state α and the energy released as heat δW are displayed as functions of α_1 only, but there is an implicit dependence on α_2 which can be obtained from the marginal curve Fig. 2 (the appropriate portions of which are shown again in Figs. 5a and 6a for clarity). It should be noted that the curve for energy release in Fig. 5c does not actually touch the α_1 axis but rather the non-dimensional energy release per unit length has a minimum value of about 0.001 corresponding to the nose of the marginal stability curve.

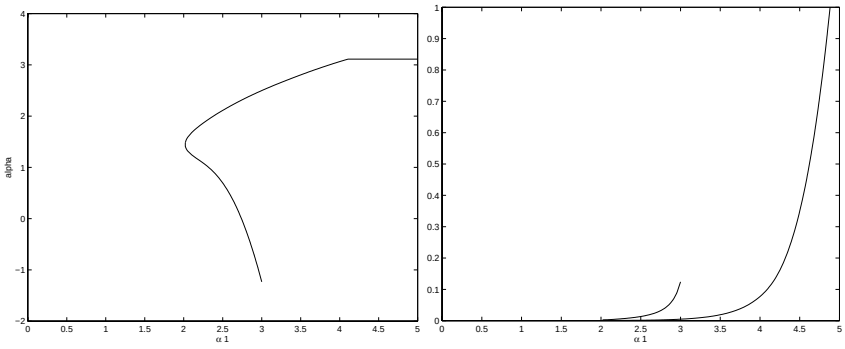

Fig. 8. The effects of relaxation to a minimum energy state which is helical for sufficiently large helicity, for the first branch of the marginal stability curve, non-dimensionalised to $\psi_0 = 1$, $\mu_0 = 1$ and $R = 1$: **a)** the value of α obtained by helicity-conserving relaxation following the marginal stability curve in Fig. 5a; **b)** the energy release (δW) per unit length due to the relaxation.

Figures 5c and 6c have been terminated for visualisation purposes at a rather arbitrary upper limit for the dimensionless energy release per unit length of $\delta W = 1$, corresponding to $2\pi W_{\rm p}$, where $W_{\rm p}$ is the energy of a potential field with the same axial flux. In fact, if we move further along the marginal stability curve, δW actually diverges to infinity. This is because our normalisation relative to axial flux (ψ_0) breaks down for initial profiles that have (for finite B_1) $\psi_0 = 0$. Such profiles have reversed axial field, so that the flux carried by the layer of positive B_z exactly cancels that in the negative B_z region. This defines a curve in the (α_1, α_2) plane, along which the normalised axial field strength on axis diverges, as do all corresponding field-related quantities (energy, helicity etc.). For example, a surface plot of normalised helicity, shown in Fig. 7, clearly diverges along a curve in the upper right hand corner. Our results should therefore be restricted to within this region; this actually affects only the "upper branch" of the marginal stability curve (Fig. 2) – this cuts the contour on which the field quantities diverge at about $(\alpha_1 R, \alpha_2 R) = (4.01, 3.26)$. The relaxation energy must diverge at this cut-off, although this is not clear from Fig. 6c due to the very sharp nature of this divergence. Evidently the region in which our model is appropriate must be cut off to exclude this divergence, and to forbid profiles with unreasonably large stored energies. One option is, in accordance with the plots Figs. 5c and 6c, to propose that stored energies greater than the potential field energy (dimensionless $\delta W = 1/2\pi \approx 0.16$) are unrealistic; alternatively, one might exclude profiles with axial field reversals, thus removing the upper branch (Fig. 6) completely and giving a maximum dimensionless energy release corresponding to the crossover of the first branch (Fig. 6a) with the reversal curve of B_z (Fig. 2) of about $\delta W = 0.2$. These two options in fact give very similar magnitudes for the upper bound to the energy release.

We have assumed so far that the minimum energy state is the simple Bessel function field (21). However, although the initial loop twisted by the footpoint motions is axisymmetric by assumption, there is no intrinsic reason why the relaxed state following the nonlinear evolution of the 3D kink instability should be symmetric: in fact, it can be shown that the minimum energy state is helical for fields with large currents (Taylor 1974, 1986). Helically distorted structures also are found in numerical simulations (e.g. Gerrard et al. 2001) and are observed in experiment (Duck et al. 1997). The minimum energy state is the cylindrical Bessel function field if $\alpha R < 3.11$ but if the normalised helicity exceeds some critical value, then the field with lowest energy has the fixed value $\alpha R \approx 3.11$ and is a combination of the axisymmetric and first helical modes

$$\boldsymbol{B} = a_0 \boldsymbol{B}^{00} + a_1 \boldsymbol{B}^{1k}, \tag{24}$$

with axial wave number $kR \approx 1.25$ (Taylor 1974, 1986). The critical helicity $K_{\rm crit}$ above which the minimum energy state is helical is thus given by $K_{\rm crit} = K(3.11/R) \approx 8.21 L \psi_0^2 / 2\pi R$. The components of the m, k equilibrium field modes are

$$\begin{aligned} B_r^{mk} &= -\frac{1}{\sqrt{\alpha^2 - k^2}} \left[k \frac{{\rm d}J_m(y)}{{\rm d}y} + \frac{m\alpha}{y} J_m(y) \right] \sin(m\theta + kz), \\ B_\theta^{mk} &= -\frac{1}{\sqrt{\alpha^2 - k^2}} \left[\alpha \frac{{\rm d}J_m(y)}{{\rm d}y} + \frac{mk}{y} J_m(y) \right] \cos(m\theta + kz), \\ B_z^{mk} &= J_m(y) \cos(m\theta + kz), \end{aligned} \tag{25}$$

where $y = r\sqrt{\alpha^2 - k^2}$ (Taylor 1986). For these values of α and k, the radial field vanishes on the wall $r = R$. The helicity then determines the amplitude of the asymmetric component (the ratio a_1/a_0), rather than the value of α as for the symmetric case, according to

$$K = \frac{L}{2\pi} \left(8.21 + 4.49 \frac{a_1^2}{a_0^2} \right), \tag{26}$$

(Taylor 1974, 1986); here, as earlier, we have normalised with respect to the flux and scaled the lengths with respect to the loop radius. The axisymmetric amplitude a_0 is determined by the axial flux ($\psi_0 = 1$) and is independent of K since the asymmetric field components have zero net flux. Evidently, the larger the helicity, the stronger is the helical deformation. In order to calculate the energy release we use the expression

$$W = \frac{\alpha}{2\mu_0} \left(K + \frac{L}{2\pi R} \frac{J_0(\alpha R)}{J_1(\alpha R)} \psi_0^2 \right), \tag{27}$$

(Reiman 1980) for the energy of a helical or axisymmetric constant-α field, with $\alpha = 3.11$. Thus, we do not need to explicitly calculate the helical amplitude a_1 (from (26)) or the field components (from (25)).

The results for energy release allowing helical minimum energy states where appropriate are shown in Fig. 8 (first branch) and Fig. 9 (second branch). The effect is to increase the heating, but only for limited range of initial field profiles;

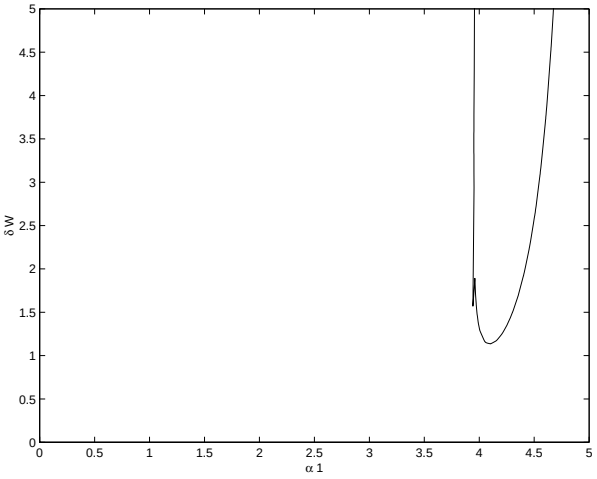

Fig. 9. The energy release (δW) per unit length due to the relaxation to a minimum energy state which is helical for sufficiently large helicity, for the second branch of the marginal stability curve, non-dimensionalised to $\psi_0 = 1$, $\mu_0 = 1$ and $R = 1$. The relaxed state $\alpha = 3.11/R$ throughout. The value of α_2 is from the portion of the marginal stability curve shown in Fig. 6a.

the range in which a helical state is obtained is quite narrow, since the contour of normalised helicity $K \approx 8.21L/2\pi R$ in the (α_1, α_2) plane is not very far from the $K = \infty$ contour (see Fig. 7). The first branch of the stability curve (compare Figs. 8a, b with Figs. 5b, c) is not much affected by the inclusion of helical relaxed states; some "saturation" of the relaxed state α can be seen in Fig. 8a, and the energy release is greater within this parameter range. However, for the second branch of the stability curve, the minimum energy state is always helical, since the value of α calculated assuming a symmetric relaxed state is always greater than $3.11/R$ (Fig. 6b), and the energy release is always more than predicted by the axisymmetric model (Fig. 9). In this case, the field always relaxes to a constant-α state with $\alpha R \approx 3.11$. This is because coincidentally (perhaps), this portion of the marginal stability curve closely follows a constant normalised helicity contour (see Fig. 7); thus $K_{\rm crit} < K < \infty$ (although at the "nose" of this branch the curve takes a sharp turn and cuts across the $K = \infty$ curve).

One consequence of the existence of helical relaxed states is that when the constant-α field ($\alpha_1 = \alpha_2$) goes unstable, it does relax to the lower energy helical state, even for an infinitely long (un-line-tied) loop, since the ideal instability onset ($\alpha R \approx 3.176$) occurs above the limit $\alpha R \approx 3.11$ for the appearance of helical distortion. Why is the cylindrical constant-α state ideally stable (albeit for a very small range of α in the case of an infinite length loop) when a lower energy state exists? This is because the helical field is only accessible through resistive evolution, consistent with our view that the ideal instability triggers reconnections – indeed, the critical value $\alpha R \approx 3.11$ is the stability threshold for resistive tearing modes which differs (albeit slightly for an infinite length loop) from the ideal stability threshold $\alpha R \approx 3.176$. Thus energy release is finite even when the loop evolves to a constant-α state. In fact for our choice of aspect ratio (L/R), instability of a constant-α field is not obtained in the viable range, since the instability onset at $\alpha R \approx 4.64$ is in the "forbidden" region outside the helicity divergence contour (where the profiles have zero axial flux for finite field magnitude). Nevertheless, it is still the case, as expected, that the more non-uniform the α profile (the further from the diagonal in the (α_1, α_2) plane), the greater the heating.

The interpretation of these results is as follows. The loop current profile will evolve in a quasi-random manner as it is perturbed by slow photospheric footpoint motions. From time to time, the marginal stability threshold will be crossed and a heating event occurs, whose magnitude is as shown in Figs. 8 and 9. The loop has now relaxed to a stable constant-α state, represented by a point somewhere along the diagonal within the stable region of Fig. 2, and the process then repeats, with the current profile evolving again until another relaxation event occurs. (This becomes beyond the scope of our model to calculate once a helical field is attained, but this is likely to be a fairly rare occurrence, and even then, the general physical principles of our model remain valid.) Since the size of possible heating events varies considerably, from a minimum of about 0.001 in dimensionless units up to order 1, sometimes small and sometimes large heating events will occur. The relative frequency of events of different magnitudes clearly depends on the spectrum of driving motions, albeit in a rather complex manner. It is clear that not all crossing points of the marginal stability curve are equally likely – intuitively we might expect that the curve will be most often reached near the nose of the first branch (i.e. around $(\alpha_1 R, \alpha_2 R) = (2.02, 0.84)$) and hence small energy releasing events ("nanoflares") will be much more common. Future work will consider the relationship between the driving motions and the heating event frequency spectrum in more detail.

5. Conclusions

We have presented a model for coronal heating which predicts ab initio a range of heating events of different sizes, reminiscent of the flare-like events (observed and inferred) which have been predicted to heat the corona. The basic assumptions of our model are two-fold: first, a heating event is triggered by an ideal current-driven instability, second, that the field then relaxes to a minimum energy state, conserving its total helicity. The free magnetic energy liberated during the relaxation is assumed to be converted into thermal energy, although during the relaxation phase there may be a build up of kinetic energy, which is subsequently dissipated. Coronal heating is then a consequence of a sequence of such heating events, with low energy events being more common. We have worked out this model quantitatively in a simple geometry consisting of a straight cylinder representative of a single loop. This is essentially a proof of principle calculation, and the model could be applied in future to a variety of geometries.

The results presented so far have been scaled to a loop with unit radius and flux, giving a minimum (dimensionless) energy

release of about 0.001. The dimensional energy release is given by

$$\delta W = \frac{1}{\mu_0}\psi_0^2 \frac{L}{R^2}\delta W^* = \frac{\pi^2 L R^2 B^2}{\mu_0}\delta W^*, \quad (28)$$

where δW^* is the dimensionless energy per unit length (as plotted in Figs. 5c, 6c, 8b and 9) and B is the mean axial field strength (the magnitude of the potential field). Taking typical coronal loop parameters L = 20 Mm, R = 1 Mm, and $\psi_0 = 10^{10}$ Wb (corresponding to B = 0.0032 T) we find the energies must be scaled by 1.6×10^{21} J (corresponding to a dimensionless energy of 1 in Figs. 5, 6, 8, 9). The range of heating events is thus around 10^{18}–3×10^{20} J. This may be usefully summarised in the form

$$\delta W = 1.6 \times 10^{25} \left(\frac{B}{32\ \mathrm{G}}\right)^2 \left(\frac{R}{1\ \mathrm{Mm}}\right)^2 \left(\frac{L}{20\ \mathrm{Mm}}\right) \left(\frac{\delta W*}{0.001}\right) \text{ ergs}, \quad (29)$$

expressed here in the still widely-used cgs units. As discussed in the previous section, the maximum energy release allowable is rather flexible but we have taken here a maximum dimensionless energy release per unit length ≈0.2; in any case, high energy events will be rare. However, the model does clearly predict a minimum size of heating event (for a loop of given dimensions), with these small events being much more frequent; this lower end of the energy range (10^{18} J) is somewhat large for a nanoflare, but smaller energies would be obtained in weaker field regions or smaller loops. It is an interesting feature of our model that we predict this minimum magnitude of a heating event. Observational evidence of the existence of such a minimum energy cut-off of around 10^{17} J has been proposed by Aschwanden (1999), although this author seeks to explain this on different grounds (chromospheric pressure balance). Of course, the range of heating events will be different for loops of different sizes and field strengths, with the scaling clear from (28), although note that there is also a dependence on L/R implicit in δW^*.

There are of course many complexities in the real solar corona that must be taken into account when considering the application of our simple model. We have considered only a single loop, which (ignoring curvature and axial variations) is taken to be a cylinder. The length, between the photospheric ends, and the radius must be chosen. One problem, which indeed faces most theoretical coronal heating models, including those based on wave dissipation, is how to identify an individual loop from observed structures. The length is relatively easy to identify, although of course the range of loop lengths is large. In the context of our model, what radius should be chosen? The observed radius from soft X-ray emission is a function of the temperature and density structure, which might not correspond directly to the underlying heating; this quantity is not clearly defined anyway (since higher resolution tends to reveal narrower structures) and the range of observed values is large. It might be more realistic to define a loop in terms of the underlying magnetic field, as consisting of the field lines emerging from a single photospheric flux source, or perhaps a set of sources. The size of the loop thus introduces effectively another dimension to the parameter space of our model, for which here we have taken only a single value. It is likely that in reality there will be spectra of heating events, in loops of different sizes, which superpose. Sometimes a small bundle of field lines within a larger structure may become unstable and generate some heat (within a range of magnitudes as predicted by our model), sometime a larger set of field lines may trigger a heating event. Obviously there is the possibility to release more free energy in larger structures, so for example a full-scale flare must involve the instability of a fairly large loop (see (28)).

The issue of helical loops is interesting. One interpretation of the frequently observed sigmoidal structures is that they correspond to such helical fields. Here, our model is rather limited by the unrealistic (albeit commonly used) boundary condition of a conducting cylindrical wall at $r = R$. The helical fields presented in Sect. 4 have been constructed, following earlier work (Taylor 1974) to fit into such a cylindrical container, so that the asymmetric perturbation vanishes at the outer wall. But a solar loop is of course free to alter its shape, being constrained only weakly by surrounding field lines rather than a solid wall. We might thus expect helical structures to exist at lower currents than the limit presented here. Use of a more realistic outer boundary condition will affect the linear ideal stability, and this will be considered in future; the effects on the energetics in the nonlinear phase may be even stronger, though at present there are few (if any) models in the literature of "free surface" helical fields.

Unlike most coronal heating theories, which predict a smooth continuous heating rate, our model predicts heating by a series of discrete heating events. In particular, we predict a minimum size of energy release event for a loop of given dimensions. It is intuitively clear from our model that low energy (nanoflare like) events will be far more common than large events (flares) but to predict the actual frequency spectrum of events would require some knowledge or model of the driving photospheric motions and further calculation to determine the exact relationship between these and the coronal field equilibrium (here modelled by a simple two parameter α profile).

Appendix: Expressions for helicity and energy in special cases

Special cases for the expressions for helicity and energy (the main text for simplicity deals only with $\alpha_1, \alpha_2 > 0$) are obtained by performing the integrals (13) and (19) with appropriate expressions for the fields. These are given by Eq. (9) for $\alpha_2 < 0$. When $\alpha_1 = 0$, the field in the core is purely axial ($B_z = B_1$) with the outer field determined by (6) with matching conditions giving

$$B_2 = -\frac{\pi}{2} B_1 \alpha_2 R_c Y_1(\alpha_2 R_c),$$
$$C_2 = \frac{\pi}{2} B_1 \alpha_2 R_c J_1(\alpha_2 R_c).$$

When $\alpha_2 = 0$, the core field is as usual and the outer potential field consists of a constant axial field and the azimuthal component

$$B_\theta = \frac{B_1 J_1(\alpha_1 R_c) R_c}{r}.$$

As discussed in Sect. 2, the case $\alpha_1 < 0$ is covered by symmetry.

For the case $\alpha_2 < 0$ (current is reversed in the outer layer), the helicity is

$$\begin{aligned} K &= \frac{2\pi L B_1^2}{\alpha_1}\left[R_c^2 J_0^2(\alpha_1 R_c) + R_c^2 J_1^2(\alpha_1 R_c) \right. \\ &\left. -2\frac{R_c}{\alpha_1} J_0(\alpha_1 R_c) J_1(\alpha_1 R_c)\right] \\ &-\frac{2\pi L B_2^2}{|\alpha_2|}\left[R^2 F_0^2(|\alpha_2|R) + R^2 F_1^2(|\alpha_2|R)\right. \\ &\left. -2\frac{R}{|\alpha_2|} F_0(|\alpha_2|R) F_1(|\alpha_2|R)\right] \\ &+\frac{2\pi L B_2^2}{|\alpha_2|}\left[R_c^2 F_0^2(|\alpha_2|R_c) + R_c^2 F_1^2(|\alpha_2|R_c)\right. \\ &\left. -2\frac{R_c}{|\alpha_2|} F_0(|\alpha_2|R_c) F_1(|\alpha_2|R_c)\right] \\ &-\frac{4\pi L B_1 B_2}{|\alpha_2|}\left[F_0(|\alpha_2|R_c) - F_0(|\alpha_2|R)\right] R_c J_1(\alpha_1 R_c) \\ &\times\left(\frac{1}{\alpha_1} - \frac{1}{\alpha_2}\right), \qquad (\alpha_1 > 0, \alpha_2 < 0), \end{aligned}$$

Expressions for helicity when one of the layers has zero current density are:

$$\begin{aligned} K &= \frac{2\pi L B_1^2}{\alpha_1}\left[R_c^2 J_0^2(\alpha_1 R_c) + R_c^2 J_1^2(\alpha_1 R_c) \right. \\ &\left. -2\frac{R_c}{\alpha_1} J_0(\alpha_1 R_c) J_1(\alpha_1 R_c)\right] \\ &+2\pi L B_1^2\left[\frac{2R_c^2}{\alpha_1} J_1^2(\alpha_1 R_c)\log(R/R_c)\right. \\ &-R_c^3 J_0(\alpha_1 R_c) J_1(\alpha_1 R_c)\log(R/R_c) \\ &\left. +\frac{1}{2} R_c J_0(\alpha_1 R_c) J_1(\alpha_1 R_c)\left(R^2 - R_c^2\right)\right], \\ &\qquad (\alpha_1 > 0, \alpha_2 = 0), \end{aligned}$$

and

$$\begin{aligned} K &= \frac{2\pi L B_2^2}{\alpha_2}\left[R^2 F_0^2(\alpha_2 R) + R^2 F_1^2(\alpha_2 R)\right. \\ &\left. -2\frac{R}{\alpha_2} F_0(\alpha_2 R) F_1(\alpha_2 R)\right] \\ &-\frac{2\pi L B_2^2}{\alpha_2}\left[R_c^2 F_0^2(\alpha_2 R_c) + R_c^2 F_1^2(\alpha_2 R_c)\right. \\ &\left. -2\frac{R_c}{\alpha_2} F_0(\alpha_2 R_c) F_1(\alpha_2 R_c)\right] \\ &+\frac{2\pi L B_1 B_2 R_c^2}{\alpha_2}\left[F_0(\alpha_2 R_c) - F_0(\alpha_2 R)\right], \\ &(\alpha_1 = 0, \alpha_2 > 0). \end{aligned}$$

Note that some care is needed in interpreting helicity in these cases, since helicity is a non-local quantity. For example, even though it might be naively thought that current-free fields have no helicity, in the case when $\alpha_2 = 0$, there is a non-zero contribution to helicity due to the field in the outer layer. This is because of the interlinkage of the azimuthal field in this layer with the axial field in the core and vice versa.

Corresponding expressions for energy are

$$\begin{aligned} W &= \frac{L\pi}{\mu_0} B_1^2 \left\{R_c^2\left[J_0^2(\alpha_1 R_c) + J_1^2(\alpha_1 R_c)\right] - \frac{R_c}{\alpha_1} J_0(\alpha_1 R_c) J_1(\alpha_1 R_c)\right. \\ &\left. + \frac{1}{2} J_0^2(\alpha_1 R_c)\left(R^2 - R_c^2\right) + R_c^2 J_1^2(\alpha_1 R_c)\log(R/R_c)\right\}, \\ &\qquad (\alpha_1 > 0, \alpha_2 = 0), \end{aligned}$$

and

$$\begin{aligned} W &= \frac{L\pi}{\mu_0}\left(B_2^2\left\{R^2\left[F_0^2(\alpha_2 R) + F_1^2(\alpha_2 R)\right]\right.\right. \\ &\left.\left. -\frac{R}{\alpha_2} F_0(\alpha_2 R) F_1(\alpha_2 R)\right\} - \frac{B_1^2 R_c^2}{2}\right), \qquad (\alpha_1 = 0, \alpha_2 > 0). \end{aligned}$$

In the case $\alpha_2 < 0$, the energy is given by (20) with α_2 replaced by $|\alpha_2|$. Finally, in order to normalise properly, expressions for the total axial flux are needed:

$$\begin{aligned} \psi_0 &= \frac{2\pi B_2}{|\alpha_2|} R F_1(|\alpha_2|R) + 2\pi R_c B_1 J_1(\alpha_1 R_c)\left(\frac{1}{\alpha_1} - \frac{1}{\alpha_2}\right), \\ &\qquad (\alpha_1 > 0, \alpha_2 < 0), \\ \psi_0 &= \pi B_1\left[R_c^2 - \pi R_c R Y_1(\alpha_2 R_c) F_1(\alpha_2 R)\right], \\ &\qquad (\alpha_1 = 0, \alpha_2 > 0), \\ \psi_0 &= \frac{2\pi R_c}{\alpha_1} B_1 J_1(\alpha_1 R_c) + \pi(R^2 - R_c^2) B_1 J_0(\alpha_1 R_c), \\ &\qquad (\alpha_1 > 0, \alpha_2 = 0). \end{aligned}$$

References

Abramowitz, M., & Stegun, A. 1965, Handbook of Mathematical Functions (Dover, USA)

Arber, T. D., Longbottom, A. W., & Van der Linden, R. A. M. 1999, ApJ, 517, 990

Aschwanden, M. J. 1999, Sol. Phys., 190, 233

Baty, H. 2000, A&A, 360, 345

Baty, H. 2001, A&A, 367, 321

Baty, H., & Heyvaerts, J. 1996, A&A, 308, 935

Berger, M. A. 1984, Geophys. Astrophys. Fluid Dyn., 30, 79

Berger, M. A., & Field, G. B. 1984, J. Fluid Mech., 147, 133

Berghmans, D., Clette, F., & Moses, D. 1998, A&A, 336, 1039

Bevir, M. K., Gimblett, C. G., & Miller, G. 1985, Phys. Fluids, 28, 1826

Bevir, M. K., & Gray, J. 1980. in Reversed Field Pinch Theory Workshop, Relaxation, flux consumption and quasi-steady-state pinches, ed. H. R. Lewis, & R. A. Gerwin (Los Alamos), 176

Brennan, D., Browning, P. K., Van der Linden, R. A. M., Hood, A. W., & Woodruff, S. 1999, Phys. Plasmas, 6, 4248

Brown, M. R., Canfield, R. C., & Pevstov, A. A. 1999, Magnetic helicity in space and laboratory plasmas, vol. Geophys. Monograph 111 (Washington DC: American Geophysical Union)

Browning, P. K. 1988, J. Plasma Phys., 40, 263

Browning, P. K. 1991, Plasma Physics and Controlled Fusion, 33, 539

Browning, P. K., & Hood, A. W. 1989, Sol. Phys., 124, 271

Browning, P. K., & Priest, E. R. 1986, A&A, 159, 129

Browning, P. K., Sakurai, T., & Priest, E. R. 1986, A&A, 158, 217

Chiuderi, C., Einaudi, G., Ma, S. S., & Van Hoven, G. 1980, J. Plasma Phys., 24, 39
De Bruyne, P., & Hood, A. W. 1992, Sol. Phys., 142, 87
Dixon, A. M., Browning, P. K., & Priest, E. R. 1988, Geophys. Astrophys. Fluid Dyn., 40, 293
Duck, R. C., Browning, P. K., Cunningham, G., et al. 1997, Plasma Physics and Controlled Fusion, 39, 715
Finn, J. M., & Antonsen, T. M. 1985, Communications in Plasma Physics and Controlled Fusion, 9, 111
Furth, H. P., Killeen, J., & Rosenbluth, M. N. 1963, Phys. Fluids, 6, 459
Gerrard, C. L., Arber, T. D., Hood, A. W., & Van der Linden, R. A. M. 2001, A&A, 373, 1089
Giachetti, R., Van Hoven, G., & Chiuderi, C. 1977, Sol. Phys., 55, 371
Goedbloed, J. P. 1970, Phys. Rev. Lett., 24, 253
Goedbloed, J. P., & Hagebeuk, H. J. L. 1972, Phys. Fluids, 15, 1090
Goedbloed, J. P., & Sakanaka, P. H. 1974, Phys. Fluids, 17, 908
Golub, L., Bookbinder, J., DeLuca, E., et al. 1999, Phys. Plasmas, 6, 2205
Heidbrink, W. W., & Dang, T. H. 2000, Plasma Physics and Controlled Fusion, 42, L31
Heyvaerts, J., & Priest, E. R. 1984, A&A, 137, 63
Hood, A.W. 1992, Plasma Physics and Controlled Fusion, 34, 411
Hood, A.W., & Priest, E. R. 1981, Geophys. Astrophys. Fluid Dyn., 17, 297
Hu, Y. Q., Xia, L. D., Li, X., Wang, J. X., & Ai, G. X. 1997, Sol. Phys., 170, 283
Ireland, J., Wills-Davey, M., & Walsh, R. W. 1999, Sol. Phys., 190, 207
Ji, H., Prager, S. C., & Sarff, J. S. 1995, Phys. Rev. Lett., 74, 2945
Jordan, C. 2000, Plasma Physics and Controlled Fusion, 42, 415
Katsukawa, Y., & Tsuneta, S. 2001, ApJ, 557, 343
Kusano, K., Suzuki, Y., Kubo, H., Miyoshi, T., & Nishikawa, K. 1994, ApJ, 433, 361
Leka, K. D., Canfield, R. C., McClymont, A. N., & van Driel-Gesztelyi, L. 1996, ApJ, 462, 547
Lionello, R., Velli, M., Einaudi, G., & Mikic, Z. 1998, ApJ, 494, 840
Lothian, R. M., & Browning, P. K. 2000, Sol. Phys., 194, 205
Mandrini, C. H., Demoulin, P., & Klimchuk, J. A. 2000, ApJ, 530, 999
Melrose, D. B., Nicholls, J., & Broderick, N. G. 1994, J. Plasma Phys., 51, 163
Narain, U., & Ulmschneider, P. 1996, Space Sci. Rev., 75, 453
Ortolani, S., & Schnack, D. D. 1993, Magnetohydrodynamics of plasma relaxation (World Scientific)
Parker, E. N. 1988, ApJ, 330, 474
Parnell, C. E., & Jupp, P. E. 2000, ApJ, 529, 554
Portier-Fozzani, F., Aschwanden, M., Demoulin, P., Neupert, W., & Delaboudinere, J. P. 2001, Sol. Phys., 203, 289
Priest, E. R., & Forbes, T. G. 2000, Magnetic reconnection: MHD theory and applications (Cambridge University Press)
Reiman, A. H. 1980, Phys. Fluids, 23, 230
Sutcliffe, M. 2002, private communication
Taylor, J. B. 1974, Phys. Rev. Lett., 33, 1139
Taylor, J. B. 1986, Rev. Mod. Phys., 58, 741
Taylor, J. B. 2000, Phys. Plasmas, 7, 1623
Van der Linden, R. A. M. 1991, Ph.D. Thesis (Katholieke Universiteit Leuven)
Van der Linden, R. A. M., & Hood, A. W. 1999, A&A, 346, 303
Vekstein, G. E., Priest, E. R., & Steele, C. D. C. 1993, ApJ, 417, 781
Velli, M., Lionello, R., & Einaudi, G. 1997, Sol. Phys., 172, 257
Voslamber, D., & Callebaut, D. K. 1962, Phys. Rev. Lett., 128, 2016
Wolfson, R., Vekstein, G. E., & Priest, E. R. 1994, ApJ, 428, 345
Woltjer, L. 1958, Proc. Natl. Acad. Sci., USA, 44, 489
Zirker, J. B. 1993, Sol. Phys., 148, 43