Mon. Not. R. Astron. Soc. **273**, 491–498 (1995) Printed 27 July 1995

The abrupt development of penumbrae in sunspots

A.M. Rucklidge,[1] H.U. Schmidt[2] and N.O. Weiss[1]

[1] *Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge CB3 9EW*
[2] *Max Planck Institut für Astrophysik, Karl Schwarzschildstrasse 1, 85740 Garching-bei-München, Germany*

ABSTRACT

A sunspot is distinguished from a pore by having a filamentary penumbra, corresponding to convective motions that carry energy into the spot from the surrounding field-free plasma. A simplified model of energy transport in sunspots is developed in order to model the transition from pores to spots as the magnetic flux is varied. The observed overlap between the radii of large pores and small spots implies that the filamentary convective mode sets in suddenly and rapidly, as in the idealized case where pore solutions lose stability at a bifurcation.

Key words: Sun: magnetic fields – sunspots – convection – MHD

1 INTRODUCTION

Most of the magnetic flux that emerges from the surface of the Sun is confined to isolated flux tubes, with fields that are locally intense and then drop suddenly to zero. These features form a family, parametrized by magnetic flux, that ranges from elements too small to be resolved, through pores with radii of 1–2 Mm, to sunspots with radii of up to 25 Mm. A sunspot (and hence, presumably, a starspot) has a dark central umbra surrounded by a filamentary penumbra, while a pore is just an isolated umbra. Despite the apparent regularity of their appearance, sunspots raise many difficulties for theory (Thomas & Weiss 1992). Recently, however, it has become clear that the penumbra is the key to understanding their detailed structure.

In this paper we first consider the question: what distinguishes a sunspot, possessing a penumbra, from a pore that lacks one? Then we go on to develop a simple model that allows us to explore the transition from pores to sunspots as the enclosed magnetic flux is varied. The approach to these problems is through considering energy transport. Sunspot umbrae are dark because convection is inhibited by the vertical magnetic field. Observations show that radiative transport is dominant at the photosphere, though umbral dots provide evidence of convective activity below the radiative blanket (Muller 1992; Sobotka, Bonet & Vázquez 1993). Theoretical models confirm that solutions diverge unless convective transport takes over just below the surface (Jahn 1992), and there has been some progress in describing oscillatory convection in the umbra (Weiss et al. 1990; Proctor 1992). The filamentary penumbra is interpreted as a consequence of convective interchanges that are responsible for carrying energy across the boundary of a spot (the magnetopause) from the external plasma (Schmidt 1991; Jahn 1992; Jahn & Schmidt 1994).

We should therefore assess the overall energy budget in magnetic features. Small flux tubes can be heated laterally by radiation (Steiner, Knölker & Schüssler 1994), but the energy emitted from pores has to be carried vertically by convective processes and is diluted owing to the increase of area as the field fans out with height. Sunspots, however, require a substantial lateral influx of energy, as indicated in Fig. 1. If the total energy flux from a spot of radius R is $\mathcal{F} = \pi R^2 F_{\rm sp}$, then the averaged energy flux per unit area $F_{\rm sp} = 0.66 F_{\rm ph}$, where $F_{\rm ph}$ is the normal photospheric energy flux. Of the total energy flux $\mathcal{F}$, only 5 per cent emerges through the umbra, with radius $0.4R$, while the remainder is emitted from the penumbra, with a local energy flux $F_{\rm p} = 0.75 F_{\rm ph}$. We estimate that only about 20 per cent of the total energy flux is carried upwards along the flux tube, while the remaining 80 per cent enters the penumbra from

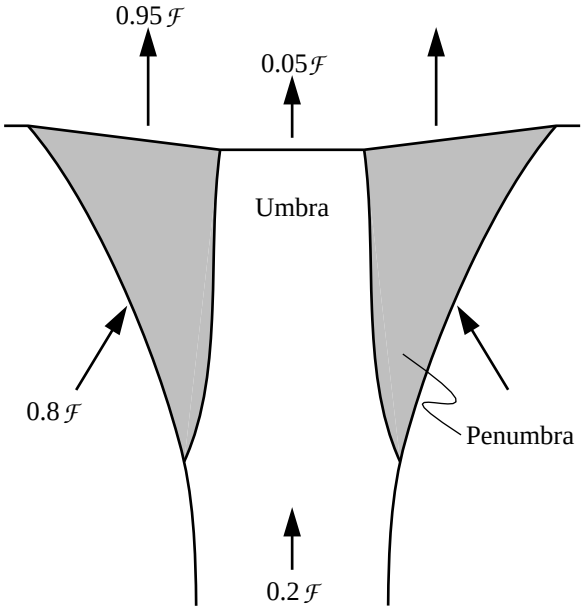

Figure 1. Cross-section of an idealized axisymmetric sunspot, showing the flux of energy from below and the lateral flux into the penumbra, together with energy fluxes from the umbra and penumbra.

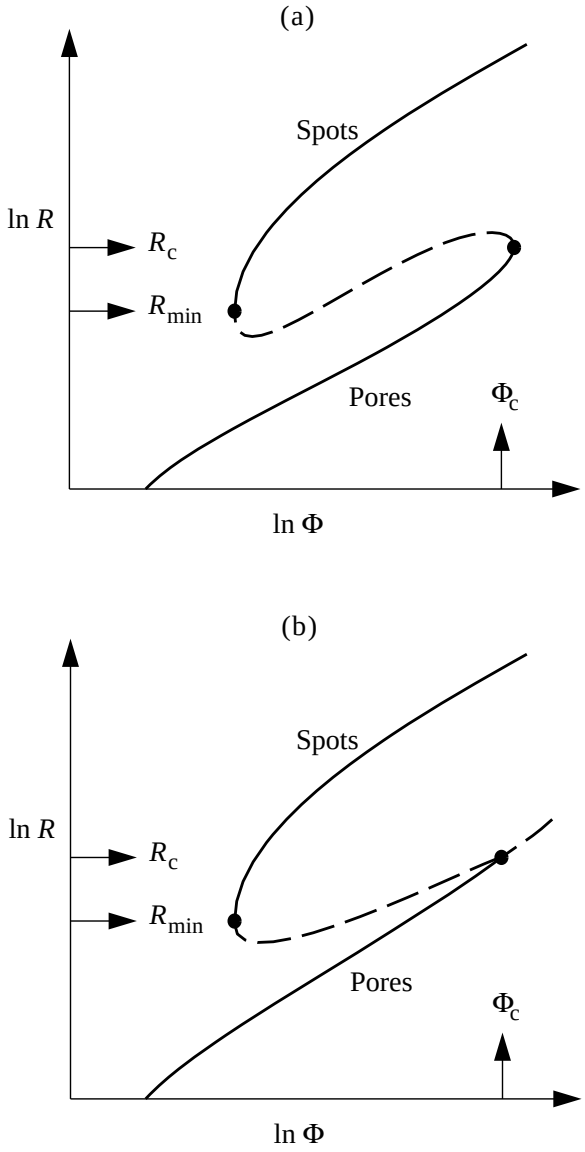

Figure 2. Hysteresis in the transition from pores to sunspots: diagrams showing the radius R as a function of the magnetic flux Φ. (a) Pores and sunspots form a single family with two turning points. (b) The branch of sunspots emerges from the branch of pores at a subcritical bifurcation. Solid and broken lines indicate stable and unstable solutions, respectively. Stable pores exist for $\Phi < \Phi_c$ with $R < R_c$, while stable spots exist for $R > R_{min}$, where $R_{min} \approx 0.5 R_c$.

outside.

This picture is based upon the detailed models of sunspots in thermodynamic and magnetohydrostatic equilibrium that have been computed by Jahn & Schmidt (1994; see also Jahn 1992). Further observational support comes from the curious variation of umbral and penumbral intensity with the solar cycle. At a wavelength of 1.67 μm the umbral intensity increases by about 20 per cent from solar minimum to solar maximum (Albregtsen & Maltby 1978, 1981; Maltby 1992). This can only be explained if the umbra is almost completely insulated from the external plasma. The penumbral intensity varies in phase with the umbra but the relative amplitude is much less, suggesting that only 10 per cent of the penumbral energy flux is transported up the flux tube, while the rest is brought in from outside (Thomas & Weiss 1992; Jahn & Schmidt 1994).

Magnetic flux in an emerging active region first appears in pores, which may coalesce to form sunspots (Zwaan 1992). The important physical parameter is the magnetic flux Φ, though it is more convenient to measure the radius R. Most pores with $R > 2$ Mm develop penumbrae within a few hours but there are examples of pores with radii of up to 3.5 Mm that survive (Bray & Loughhead 1964). Conversely, the smallest spots, with rudimentary penumbrae, have radii of only 1.8 Mm. Thus a plot of R against Φ shows two families of solutions with a region of overlap as indicated by full lines in Fig. 2(a). We explain this through the sudden appearance of filamentary convection which provides a mechanism for lateral heat transport. Simple pore models show that the field fans out as the magnetic flux increases, so that θ, the inclination to the vertical of the photospheric magnetic field at the edge, increases with increasing Φ (Simon & Weiss 1970; Spruit 1976; Simon, Weiss & Nye 1983). Simon & Weiss (1970) conjectured that a penumbra is formed when the field is nearly horizontal; we propose that a new mode of filamentary convection sets in when θ exceeds some critical value θ_c. Then pores and spots can form a single family parametrized by θ and the two stable branches are linked by an unstable branch, as shown in Fig. 2(a). This pattern of hysteresis, with two stable branches terminating in saddle-node bifurcations, is an example of a cusp catastrophe.

If all lateral energy transport across the magnetopause into the flux tube were suppressed by enclosing it in a thermal sheath, we could construct a family of pores for all values of Φ, as sketched in Fig. 2(b). The inclination θ increases monotonically along this branch, which lies close to the branch of pores in Fig. 2(a) for $\theta < \theta_c$. We then presume that filamentary convection appears as an instability when $\theta = \theta_c$, so that there is a subcritical bifurcation when $\Phi = \Phi_c$, as indicated in the Figure. We expect that this would be a Hopf bifurcation, giving rise to oscillatory motion, and that the width of the penumbra would increase along the unstable segment of the branch of spot solutions, so that stable spots first appear with finite penumbrae at the turning point, where $\Phi = \Phi_{min} < \Phi_c$. This idealized version of the bifurcation structure allows us to investigate behaviour near the bifurcation at Φ_c and to study the transition from pores to sunspots in much greater detail.

In the next section we develop a simple model with idealized geometry, incorporating conservation of magnetic flux and a balance between the energy emitted at the surface and the energy entering laterally and from below. This model is then applied both to pores and to well-developed sunspots. In section 3 we consider bifurcations from the pore solutions. Linearization about the critical point, where $\theta = \theta_c$, allows us to find conditions for subcriticality, and quadratic effects can also be included. This treatment is extended in the final section to provide a more realistic representation of pores and sunspots in the ΦR-plane; this description is consistent both with theoretical models (Jahn & Schmidt 1994) and with the observations. In conclusion, we comment on the significance of our results and emphasize the need for detailed modelling of penumbral convection.

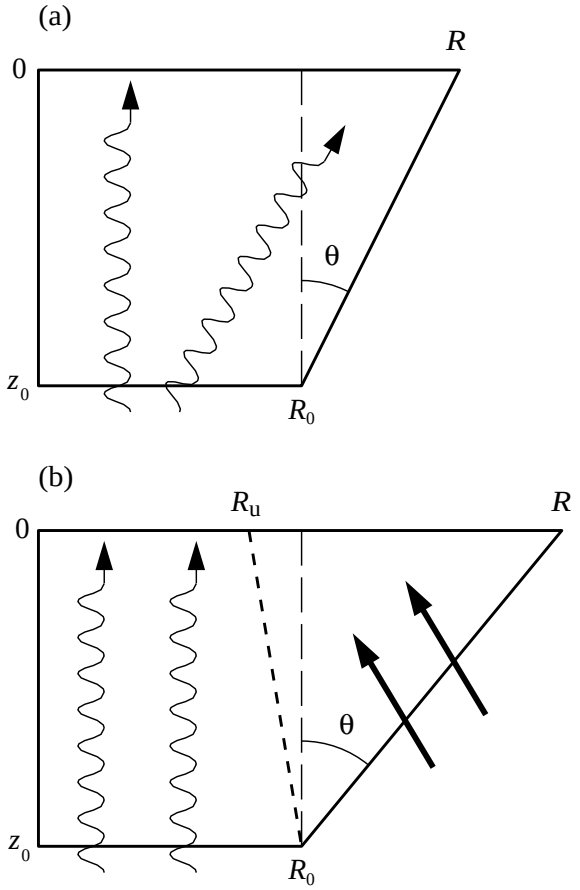

Figure 3. Cross-sections of simplified axisymmetric models of (a) pores and (b) sunspots. The magnetic field has a uniform inclination θ at the outer boundary. Upward and lateral energy fluxes are indicated.

Table 1. Pore and spot models.

	R (Mm)	z_0 (Mm)	θ	Φ (TWb)
Pores				
	1	1.0	16°	0.6
	2	1.0	30°	2.5
	3.4	1.0	45°	7.3
Spots				
	10	1.8	70°	36
	20	3.6	70°	144

2 PORE AND SPOT MODELS

To model energy transport in a pore we drastically simplify its geometry, as shown in Fig. 3(a). The region with an inclined field is represented by the frustum of a cone, with a radius R at the photosphere ($z = 0$) and a radius $R_0 < R$ at a depth of z_0. Thus the inclination θ of the field at the outer boundary satisfies

$$\tan\theta = \frac{R - R_0}{z_0}. \tag{1}$$

We assume that the vertical component of the field at the photosphere is effectively uniform with a value B_u and that there is a uniform field B_0 at the base of the pore; similarly, the energy fluxes, normalized with respect to the normal photospheric flux F_ph, at $z = 0$ and $z = z_0$ are given by F_u and F_0 respectively. Conservation of magnetic flux requires that

$$\frac{1}{\pi}\Phi = R^2 B_\mathrm{u} = R_0^2 B_0. \tag{2}$$

Since no energy flows into the flux tube from outside,

$$R^2 F_\mathrm{u} = R_0^2 F_0. \tag{3}$$

It follows that

$$B_\mathrm{u}/B_0 = F_\mathrm{u}/F_0 = R_0^2/R^2 = \beta^2, \tag{4}$$

say, where $\beta < 1$. Hence

$$\tan\theta = \frac{R}{z_0}(1-\beta) = \frac{1}{z_0}\left(\frac{\Phi}{\pi B_\mathrm{u}}\right)^{\frac{1}{2}}(1-\beta). \tag{5}$$

Thus the inclination θ increases monotonically as the magnetic flux Φ (and hence the radius R) is increased, provided that the logarithmic gradient $m \equiv \mathrm{d}\ln z_0/\,\mathrm{d}\ln R < 1$.

We take $B_\mathrm{u} = 2000\,\mathrm{G}$, as suggested by observations (Sütterlin, Thim & Schröter 1994) and, for simplicity, we regard pores as isolated umbrae with $F_\mathrm{u} = 0.2$. We further assume that the magnetic field strength and the energy flux are doubled at the base of the pore, so that $B_0 = 4000\,\mathrm{G}$, $F_0 = 0.4$ and $\beta = 1/\sqrt{2}$. The inclination θ of the field at the outer boundary of a pore is ill-determined (Sütterlin et al. 1994); we shall adopt a value of 45° for the inclination in the largest pores. If, for the moment, we assume that z_0 is independent of R and set $z_0 = 1\,\mathrm{Mm}$ we obtain the values in Table 1. This choice is consistent with the observed upper limit to the radii of pores.

Next, we extend the model to describe a spot with umbral radius R_u and penumbral radius R. The geometry is shown in Fig. 3(b). Then equations (2) and (3) are replaced by

$$\frac{1}{\pi}\Phi = R_\mathrm{u}^2 B_\mathrm{u} + (R^2 - R_\mathrm{u}^2)B_\mathrm{p} = R_0^2 B_0 \tag{6}$$

and

$$\begin{aligned} R^2 F_\mathrm{sp} &= R_\mathrm{u}^2 F_\mathrm{u} + (R^2 - R_\mathrm{u}^2)F_\mathrm{p} \\ &= R_0^2 F_0 + (R^2 - R_0^2) f(\theta), \end{aligned} \tag{7}$$

where B_p, F_p are the mean penumbral field strength and normalized energy flux respectively. Here we have assumed that penumbral convection carries energy across the magnetopause at a rate that is proportional to the total projected area (cf. Jahn & Schmidt 1994) and depends on the inclination of the boundary through a function $f(\theta)$, $0 \le f < 1$. The function $f(\theta)$, which represents the efficiency of lateral heat transport into the inclined penumbra, is the key to our treatment of this problem.

Let us consider well-developed sunspots, the gross properties of which are similar and independent of their radii. Then we may take $R_\mathrm{u} = 0.4R$, $B_\mathrm{p} = 1000\,\mathrm{G}$, $F_\mathrm{p} = 0.75$ and $\theta = 70°$ (Thomas & Weiss 1992), while other quantities retain the values given above. There are two constraints on the value of f. First, we require that $R_\mathrm{u} > 0$: then, from (6) and (7),

$$R^2/R_0^2 > B_0/B_\mathrm{p}, \qquad R^2 F_\mathrm{p} > R_0^2 F_0 + (R^2 - R_0^2) f, \tag{8}$$

whence it follows that

$$f < [(B_0/B_\mathrm{p})F_\mathrm{p} - F_0]/[B_0/B_\mathrm{p} - 1] = 0.87. \tag{9}$$

Secondly, we expect $R_\mathrm{u} < R_0$, to give the geometry of Fig. 1. Hence

$$(R_\mathrm{u}/R)^2 < (R_0/R)^2 = (f - F_\mathrm{sp})/(f - F_0) < 1. \tag{10}$$

Since $F_0 < F_\mathrm{sp}$ it follows that $f > F_\mathrm{sp}$. So $0.66 < f < 0.87$; we choose $f = 0.75$, which is consistent with the models of Jahn & Schmidt (1994). Then $R_0/R = 0.51$ and $z_0/R = 0.18$. Table 1 shows some representative values. In these models 84 per cent of the energy emitted through the spot enters the penumbra from outside.

To complete our simple model we need to prescribe z_0 as a function of R. Although the logarithmic gradient $m = 1$ for large sunspots it is clear that m must be less for pores, for which $(z_0/R) \approx 0.29 \cot\theta$, from (5). So we expect m to rise from some small value when $R = 0$ to unity when R is very large, with a value around 0.5 to 0.7 for the largest pores.

3 THE TRANSITION FROM PORES TO SPOTS

We hypothesize that a new form of convective transport sets in when the inclination θ exceeds a critical value θ_c, carrying energy across the magnetopause into the flux tube and giving rise to a filamentary penumbra. Thus the efficiency factor $f(\theta)$ in (7) varies as indicated schematically in Fig. 4(a), rising rapidly but continuously from a small value for $\theta < \theta_\mathrm{c}$ and tending to a maximum value $f_0 = 0.75$ as θ tends to 70°. We represent this behaviour for sunspots by an idealized function

$$f(\theta) = \begin{cases} 0 & (\theta < \theta_\mathrm{c}), \\ g(\theta - \theta_\mathrm{c}) & (\theta_\mathrm{c} < \theta < \pi/2), \end{cases} \tag{11}$$

such that $g(0) = 0$, $g'(\psi) > 0$, $g(\psi) \to f_0$ as $\psi \to (\pi/2 - \theta_\mathrm{c})$, which has the form shown in Fig. 4(b). For pore models we simply set $f(\theta) = 0$ for all $0 < \theta < \pi/2$; then there is a bifurcation at $\theta = \theta_\mathrm{c}$ and the pores must be unstable for $\theta > \theta_\mathrm{c}$.

We shall investigate behaviour in the immediate neighbourhood of this bifurcation in order to see whether it is possible to reproduce the bifurcation pattern in Fig. 2(b). At the bifurcation point we have $\theta = \theta_\mathrm{c}$, $R = R_\mathrm{c}$, $R_0 = R_{0\mathrm{c}}$ and $z_0 = z_{0\mathrm{c}}$, where $\tan\theta_\mathrm{c} = (R_\mathrm{c}/z_{0\mathrm{c}})(1 - \beta)$, from (5), and $\Phi_\mathrm{c} = \pi R_\mathrm{c}^2 B_\mathrm{u}$. Near the bifurcation we hope to find solutions with $\theta > \theta_\mathrm{c}$ but with $R < R_\mathrm{c}$ and $\Phi < \Phi_\mathrm{c}$; this requires that $R_0 < R_{0\mathrm{c}}$ and $R_\mathrm{u} < R < R_\mathrm{c}$, as indicated in Fig. 5. So we set

$$\begin{aligned} R &= (1 + r)R_\mathrm{c}, \qquad R_\mathrm{u} = (1 + r_\mathrm{u})R_\mathrm{c}, \\ R_0 &= (1 + r_0)R_{0\mathrm{c}} = (1 + r_0)\beta R_\mathrm{c}, \end{aligned} \tag{12}$$

expecting to find $r_0 < 0$ and $r_\mathrm{u} < r < 0$. Furthermore, we assume that, locally, $z \propto R^m$ for some m $(0 < m < 1)$. Then from (1) and (12) we have, to leading order in $\psi \equiv \theta - \theta_\mathrm{c}$,

$$\tan\theta = \tan\theta_\mathrm{c} + \psi \sec^2\theta_\mathrm{c} = \frac{R_\mathrm{c}[1 + r - \beta(1 + r_0)]}{z_{0\mathrm{c}}(1 + mr)}, \tag{13}$$

whence, from (5), it follows that

$$[1 - m(1 - \beta)]r = \beta r_0 + 2\psi(1 - \beta)\,\mathrm{cosec}\,2\theta_\mathrm{c}. \tag{14}$$

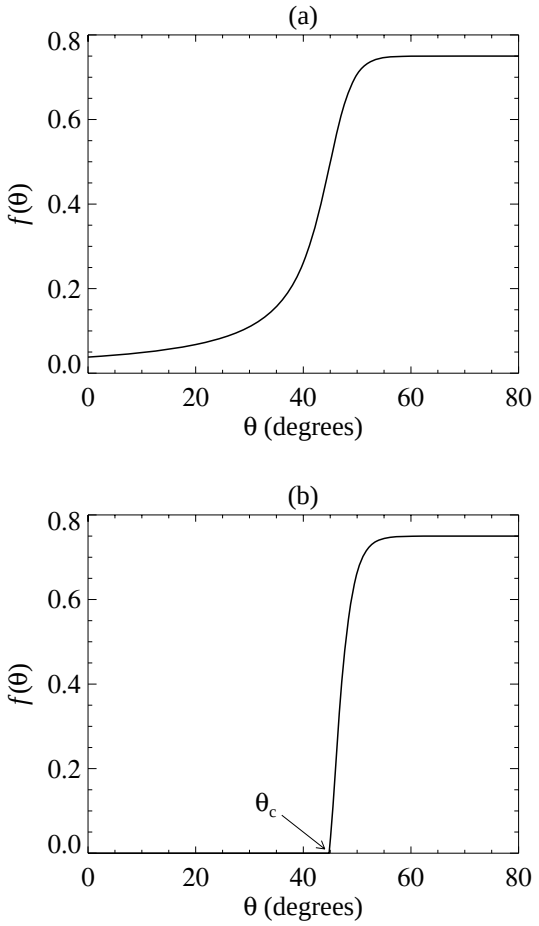

Figure 4. The efficiency $f(\theta)$ of energy transport across the magnetopause. (a) The efficiency varies smoothly with the inclination θ, increasing rapidly with the onset of filamentary convection. (b) Idealized behaviour, with $f(\theta)$ rising sharply after a bifurcation at $\theta = \theta_\mathrm{c}$.

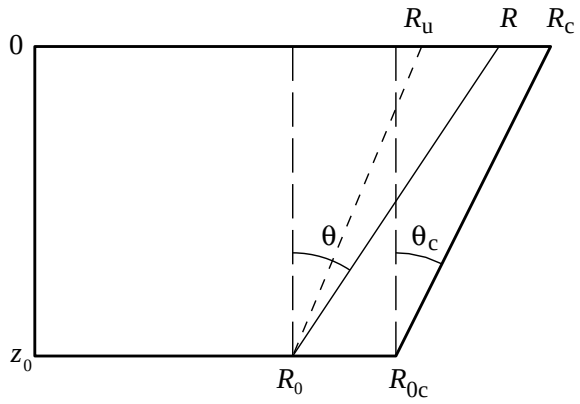

Figure 5. Geometry of sunspots near the critical pore radius R_c. The spots have $\theta > \theta_\mathrm{c}$ but $R_\mathrm{u} < R < R_\mathrm{c}$ and $R_0 < R_{0\mathrm{c}}$.

For pore models we have $r = r_0 = 2\psi\,\mathrm{cosec}\,2\theta_\mathrm{c}/(1 - m)$. For spots we linearize the equations for conservation of magnetic flux and energy flux. From (6) we then obtain, to leading order,

$$(1 - \gamma)r_\mathrm{u} = r_0 - \gamma r, \tag{15}$$

where $\gamma \equiv B_{\rm p}/B_{\rm u} < 1$. Near the bifurcation we set

$$g(\psi) = 2\alpha\psi + \zeta\psi^2 + \mathcal{O}(\psi^3), \qquad \alpha > 0, \tag{16}$$

so that the linearized form of (7) yields

$$\alpha(1-\beta^2)\psi = F_{\rm p} r - (F_{\rm p} - F_{\rm u}) r_{\rm u} - F_{\rm u} r_0; \tag{17}$$

α is now the key parameter that determines the growth of lateral heat transport at the bifurcation, and ζ parametrizes non-linear effects.

We now proceed to calculate the range of α that will give $r_0 < 0$ and $r < 0$. From (14), (15) and (17) we find that

$$\frac{r}{\psi} = \frac{2\,\mathrm{cosec}\,2\theta_{\rm c}}{(1-m)} - \frac{\alpha\beta(1+\beta)(1-\gamma)}{(1-m)(F_{\rm p}-\gamma F_{\rm u})}, \tag{18}$$

while

$$\frac{r_0}{\psi} = \frac{2\,\mathrm{cosec}\,2\theta_{\rm c}}{(1-m)} - \frac{\alpha[1-m(1-\beta)](1+\beta)(1-\gamma)}{(1-m)(F_{\rm p}-\gamma F_{\rm u})}. \tag{19}$$

So $r < 0$ for $\alpha > \alpha_{\rm c}$, where

$$\alpha_{\rm c} = \frac{2(F_{\rm p} - \gamma F_{\rm u})\,\mathrm{cosec}\,2\theta_{\rm c}}{\beta(1+\beta)(1-\gamma)}, \tag{20}$$

and is independent of m, while $r_0 < 0$ for $\alpha > \alpha_{0{\rm c}}$, where

$$\begin{aligned} \alpha_{0{\rm c}} &= \frac{2(F_{\rm p} - \gamma F_{\rm u})\,\mathrm{cosec}\,2\theta_{\rm c}}{[1-m(1-\beta)](1+\beta)(1-\gamma)} \\ &= \frac{\beta}{[1-m(1-\beta)]}\alpha_{\rm c} < \alpha_{\rm c}. \end{aligned} \tag{21}$$

Now the magnetic flux $\Phi = \Phi_{\rm c}(1+2r_0)$, so behaviour near the bifurcation point in Fig. 2(b) can be represented by considering r as a function of r_0. Fig. 6 shows how the initial slope of the sunspot branch changes as α is increased. For $\alpha = 0$ there is only the branch of pores with $r = r_0$. For $0 < \alpha < \alpha_{0{\rm c}}$ the bifurcation is supercritical, with r and r_0 both positive. For $\alpha_{0{\rm c}} < \alpha < \alpha_{\rm c}$ the bifurcation is subcritical, with $r_0 < 0$, but r remains positive. For $\alpha > \alpha_{\rm c}$ both r and r_0 are negative, as required by the observations. In the limit as $\alpha \to \infty$, $r \to \beta r_0/[1-m(1-\beta)]$. If we set $\theta_{\rm c} = \pi/4$ for convenience and adopt the parameter values in section 2, then $\alpha_{\rm c} = 2.15$, while $\alpha_{0{\rm c}} = 1.52/(1-0.293m)$; so, for example, $\alpha_{0{\rm c}} = 1.52\,(1.78)$ for $m = 0\,(0.5)$.

This simple linear treatment shows that the hysteresis implied by the observations is possible provided that lateral transport of energy sets in sufficiently rapidly at the bifurcation. Since $g < f_0 = 0.75$, the slope $g'(\psi)$ must decrease as θ increases along the sunspot branch. Thus we expect that addition of non-linear terms will allow r to reach a minimum, after which it will increase with increasing θ. This can be demonstrated by including quadratic terms in the expansion, so that $g(\theta)$ is a parabola. As an illustration, we show in Fig. 7 a solution with a turning point below the bifurcation, obtained for $\alpha = 3.25$, $\zeta = -30$ and $m = 0$. Thus our simple model allows us to construct a single family of sunspot solutions on a branch that emerges from the bifurcation, descends to smaller radii and fluxes as required by observations, and then turns round to yield stable sunspots.

4 DISCUSSION

The idealized sunspot model was constructed by conserving magnetic flux and the flux of energy, in a severely simplified geometry. We have shown that these ingredients suffice

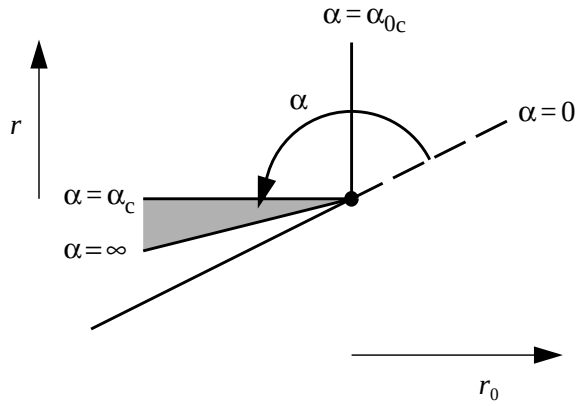

Figure 6. The pore–sunspot transition. Results from linear theory, showing the changes in r and r_0 as α is increased. Both r and r_0 are negative for $\alpha > \alpha_{\rm c}$.

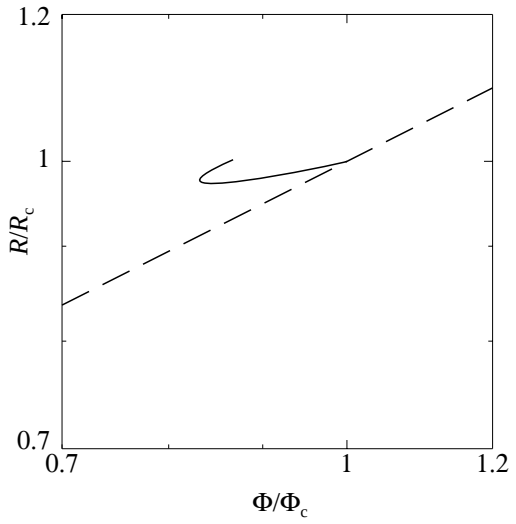

Figure 7. The pore–sunspot transition: second-order theory. Families of pores (broken line) and spots (solid line) in the ΦR-plane for the model with $\alpha = 3.25 > \alpha_{\rm c}$, $\zeta = -30$ and $m = 0$. The branch of sunspot solutions bifurcates subcritically but quadratic terms produce a turning point.

to explain hysteresis at the pore–spot transition. To proceed further we need to concoct complete prescriptions for $z_0(R)$ and $f(\theta)$. We begin with an analytical expression for the logarithmic gradient $m \equiv \mathrm{d}\ln z_0/\,\mathrm{d}\ln R$, involving a hyperbolic tangent, and thence obtain z_0, subject to the constraints that $m \to 1$ as $R \to \infty$ and $z_0(R_{\rm c}) = z_{0{\rm c}}$. Fig. 8 shows the functions that we shall use. The efficiency factor $f(\theta)$ is defined by setting $g(\psi) = A\tanh(2\alpha\psi/A)$ and choosing A so that $f = 0.75$ when $\theta = 70°$. We have already found that hysteresis occurs for $\alpha > 2.15$; Figs 9(a) and (b) show $f(\theta)$ for the two cases $\alpha = 5.0$ and $\alpha = 12.5$. The corresponding variations of the radius R with the magnetic flux Φ are displayed in Figs 10(a) and (b), which should be compared with Fig. 2. We see that the smallest spots in Fig. 10(a) have a radius $R_{\rm min} \approx 0.8R_{\rm c}$, while those in Fig. 10(b) have $R_{\rm min} \approx 0.6R_{\rm c}$, with $\Phi_{\rm min} \approx 0.25\Phi_{\rm c}$. To reproduce the range of overlap reported by Bray & Loughhead

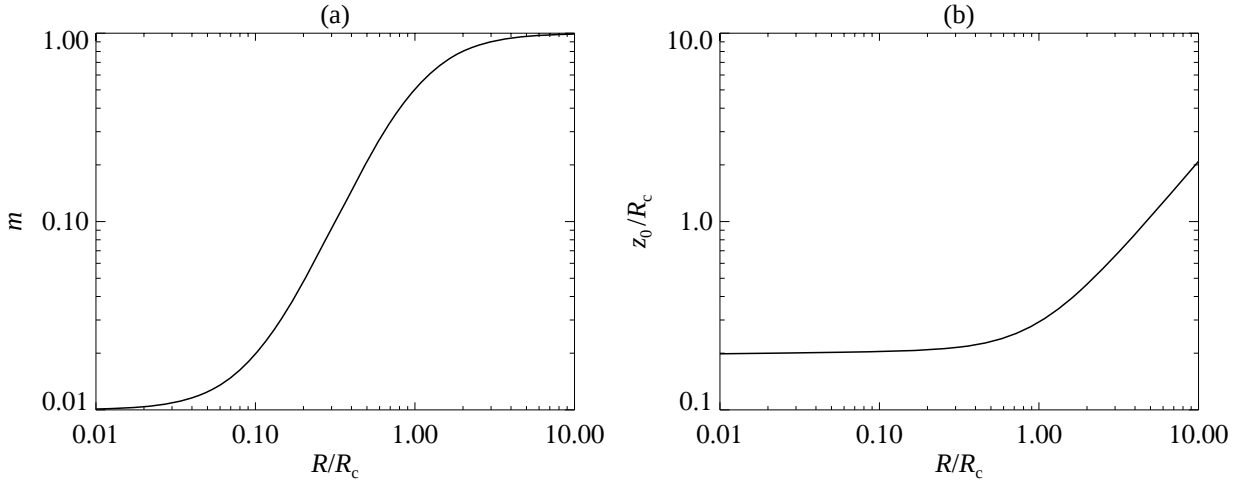

Figure 8. The relationship between depth and radius for the spot and pore models. (a) Assumed variation of the logarithmic derivative $m = \mathrm{d}\ln z_0/\mathrm{d}\ln R$ with R. (b) The corresponding dependence of z_0 on R.

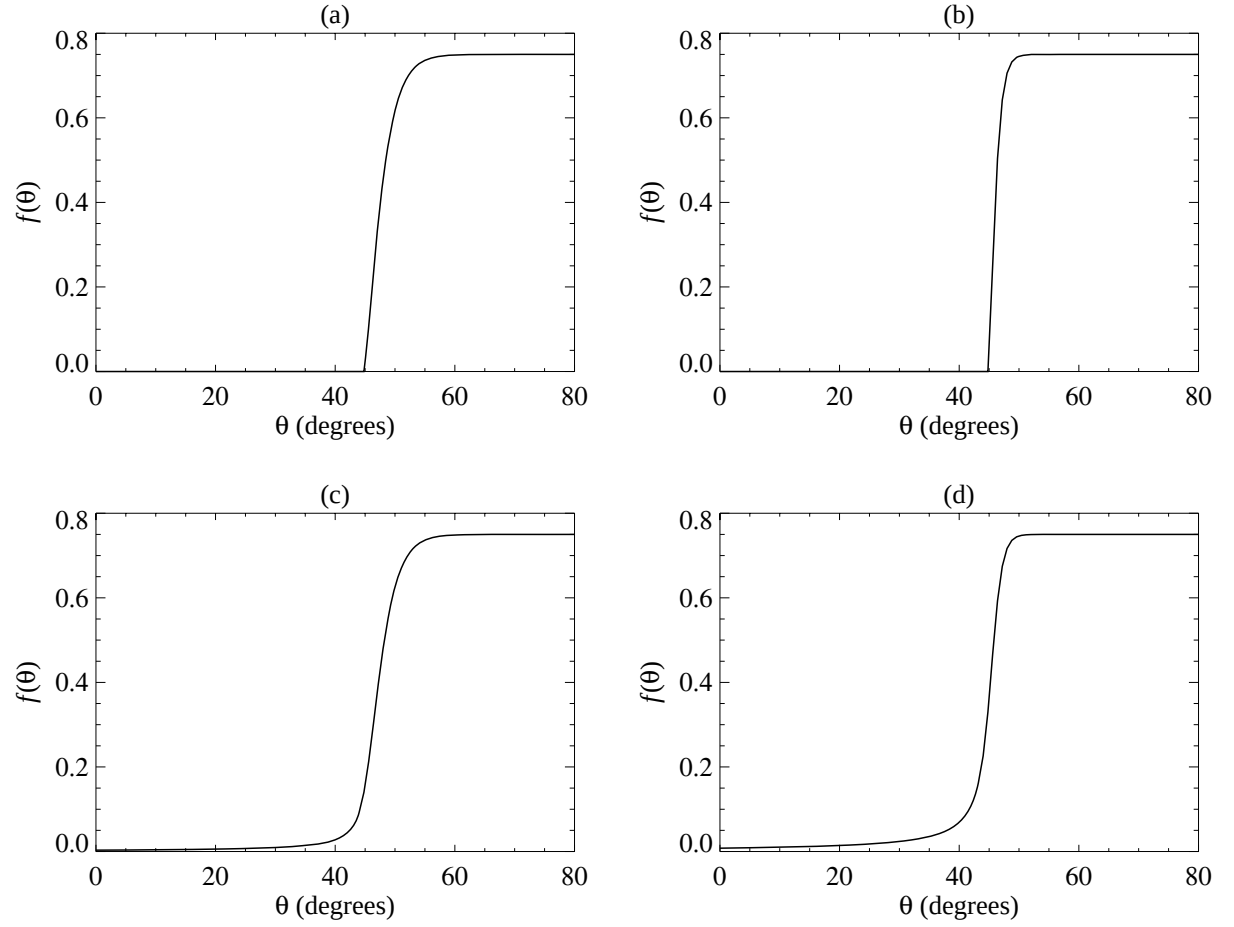

Figure 9. Different forms for the efficiency function $f(\theta)$. Discontinuous variation ($C = 0$) with (a) moderate slope ($\alpha = 5.0$), (b) steep slope ($\alpha = 12.5$). Smooth variation ($C = 10^{-3}$) with (c) $\alpha = 5.0$, (d) $\alpha = 12.5$.

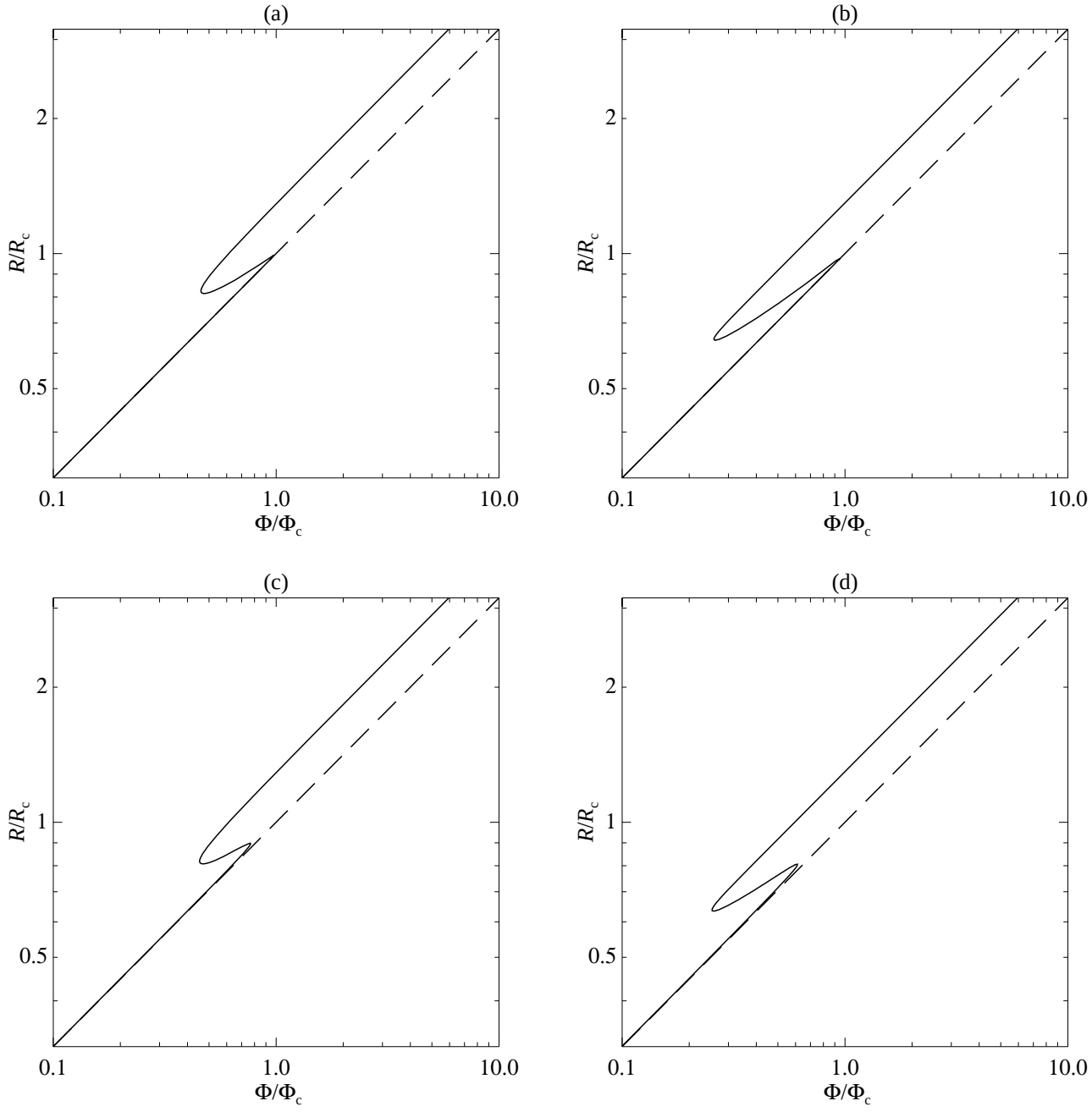

Figure 10. Four families of pore–spot solutions, with parameter values corresponding to the cases in Fig. 9. Solid lines denote sunspot solutions and broken lines indicate the pores that would be obtained with $f(\theta)$ set to zero. The top two panels show sharp transitions in $f(\theta)$ at $\theta = \theta_c$: (a) with moderate slope ($\alpha = 2.3\alpha_c$); (b) with $f(\theta)$ increasing steeply ($\alpha = 5.8\alpha_c$). The bottom two panels show continuous transitions with (c) $\alpha = 2.3\alpha_c$, (d) $\alpha = 5.8\alpha_c$. The observed overlap between pores and spots can only be matched if $f(\theta)$ rises suddenly and steeply, as in case (b), so that penumbrae appear abruptly.

(1964), with $R_{\min} \approx 0.5R_c$, the efficiency factor must rise even more steeply than in Fig. 9(b).

So far, we have assumed that there is no lateral transport for $0 < \theta < \theta_c$. A more realistic assumption is that $f(\theta)$ varies smoothly, as sketched in Fig. 4(a). We represent this behaviour by an expression of the form

$$f(\theta) = A\tanh(2\alpha\chi/A), \tag{22}$$

where $\chi = \frac{1}{2}[\psi + (\psi^2 + C)^{\frac{1}{2}}]$ and $C \ll 1$. For $C \leq 10^{-5}$ the variation of f with θ is indistinguishable from that in Figs 9(a) and (b). The curves in Figs 9(c) and (d) show $f(\theta)$ for the same values of α but with $C = 10^{-3}$. Now $\mathrm{d}f/\mathrm{d}\theta$ is continuous, the sharp bifurcation is smoothed out and replaced by a turning point, and there is a single family of solutions that includes both pores and spots. As a result, the extent of hysteresis in Figs 10(c) and (d) is significantly

reduced (though $R_{\min}$ itself is virtually unchanged). A further increase to $C = 10^{-2}$ completely eliminates the overlap in radius between pores and spots, as might be expected for a cusp catastrophe.

From Fig. 10 it is apparent that the observations can only be matched if C is very small, so that the change of slope in f is effectively discontinuous, and if α is substantially greater than the critical value α_c. This result is not affected by altering the values adopted for θ_c and β, or by changing the prescribed forms of $m(R)$ and $f(\theta)$. Since our simple model is so closely related to the detailed structures computed by Jahn & Schmidt (1994), we are confident that the qualitative results are robust. We conclude, therefore, that lateral transport of energy increases sharply when penumbral convection appears, and that the idealized bifurcation sketched in Fig. 2(b) is a very good approximation to the actual imperfect bifurcation.

These calculations are consistent with the observations and demonstrate that there is now a coherent physical picture of overall energy transport in sunspots. Moreover, they confirm the validity of the approach that emphasizes lateral heating (Schmidt 1991; Jahn 1992; Jahn & Schmidt 1994). In particular, the results in Figs 9 and 10 show that the penumbral mode of convection enters *abruptly* and *rapidly* at what is close to being an ideal bifurcation. Furthermore, they indicate that filamentary convection is a robust feature, which should not depend critically on details of a sunspot model.

The largest pores that are observed have diameters of about 10 arcsec, corresponding to a critical radius $R_c \approx 3.5\,\mathrm{Mm}$ and a magnetic flux $\Phi_c \approx 7\,\mathrm{TWb}$ (or $7 \times 10^{20}\,\mathrm{mx}$). On the other hand, the smallest sunspots have diameters of about 5 arcsec ($R_{\min} \approx 1.8\,\mathrm{Mm}$), and an estimated flux $\Phi_{\min} \approx 2\,\mathrm{TWb}$. Since the unstable portion of the sunspot branch in Figs 2 and 10 lies so close to the branch of stable pores, we expect the latter to have a narrow basin of attraction; in other words, any pore with a flux in the range $\Phi_{\min} < \Phi < \Phi_c$ will be unstable to finite amplitude perturbations that lead to the development of a penumbra. Since pores grow by amalgamation or by absorbing small flux elements, and are continuously buffeted by convection in surrounding granules, it is not surprising that they generally develop into sunspots if $\Phi > \Phi_{\min}$. In that case the penumbra forms rapidly. Of course, a real spot is not likely to be circular; indeed, filaments often appear first on the side that faces outward from the active region while pores are still being absorbed on the inner side, and an annular penumbra only develops later (Zwaan 1992). The reverse process cannot readily be observed, since small spots are usually torn apart rather than decaying gradually until $\Phi < \Phi_{\min}$.

For the future, we may expect that new high-resolution observations will clarify the relationship between pores and small spots, and the development of penumbrae. The main need, however, is for a proper theoretical description of the non-linear convective processes that generate a filamentary penumbra. As well as providing a mechanism for lateral heat transport, such a theory must explain the complicated interlocking comb-like structure of the observed magnetic field (Title et al. 1992, 1993; Schmidt et al. 1992). Our investigation is only a prelude to the full problem, which remains a major challenge.

ACKNOWLEDGMENTS

We thank Douglas Gough and Steven Tobias for helpful suggestions, and we are grateful to SERC and its successors, EPSRC and PPARC, for supporting this research.

REFERENCES

Albregtsen F., Maltby P., 1978, Nat., 274, 41
Albregtsen F., Maltby P., 1981, Sol. Phys., 71, 269
Bray J.R., Loughhead R.E., 1964, Sunspots. Chapman and Hall, London
Jahn K., 1992, in Thomas J.H., Weiss N.O., eds, Sunspots: Theory and Observations. Cambridge Univ. Press, Cambridge, p. 139
Jahn K., Schmidt H.U., 1994, A&A, 290, 295
Maltby P., 1992, in Thomas J.H., Weiss N.O., eds, Sunspots: Theory and Observations. Cambridge Univ. Press, Cambridge, p. 103
Muller R., 1992, in Thomas J.H., Weiss N.O., eds, Sunspots: Theory and Observations. Cambridge Univ. Press, Cambridge, p. 175
Proctor M.R.E., 1992, in Thomas J.H., Weiss N.O., eds, Sunspots: Theory and Observations. Cambridge Univ. Press, Cambridge, p. 221
Schmidt H.U., 1991, Geophys. Astrophys. Fluid Dyn., 62, 249
Schmidt W., Hofmann A., Balthasar H., Tarbell T.D., Frank Z.A., 1992, A&A, 264, L27
Simon G.W., Weiss N.O., 1970, Sol. Phys., 13, 85
Simon G.W., Weiss N.O., Nye A.H., 1983, Sol. Phys., 87, 65
Sobotka M., Bonet J.A., Vázquez M., 1993, ApJ, 415, 832
Spruit H.C., 1976, Sol. Phys., 50, 269
Steiner O., Knölker M., Schüssler M., 1994, in Rutten R.J., Schrijver C., eds, Solar Surface Magnetism. Kluwer, Dordrecht, p. 441
Sütterlin P., Thim F., Schröter E.H., 1994, in Schüssler M., Schmidt W., eds, Solar Magnetic Fields. Cambridge Univ. Press, Cambridge, p. 213
Thomas J.H., Weiss N.O., 1992, in Thomas J.H., Weiss N.O., eds, Sunspots: Theory and Observations. Cambridge Univ. Press, Cambridge, p. 3
Title A.M., Frank Z.A., Shine R.A., Tarbell T.D., Topka K.P., Scharmer G., Schmidt W., 1992, in Thomas J.H., Weiss N.O., eds, Sunspots: Theory and Observations. Cambridge Univ. Press, Cambridge, p. 195
Title A.M., Frank Z.A., Shine R.A., Tarbell T.D., Topka K.P., Scharmer G., Schmidt W., 1993, ApJ, 403, 780
Weiss N.O., Brownjohn D.P., Hurlburt N.E., Proctor M.R.E., 1990, MNRAS, 245, 434
Zwaan C., 1992, in Thomas J.H., Weiss N.O., eds, Sunspots: Theory and Observations. Cambridge Univ. Press, Cambridge, p. 75

This paper has been produced using the Blackwell Scientific Publications TEX macros.