

A new morphological theory on the formation of sunspots

Abstract

The formation of sunspots has been one of the most important problems in solar physics. Since Babcock (1961) presented his theory, it has been universally accepted and become the foundation of later studies. However, his model has at least three obvious problems, which have been left behind. This paper considers these long-standing problems which contradict with Babcock's model: (1) The presence of magnetic monopole-like single spots, (2) the delay of the appearance of f -spots after p -spots, (3) the flux unbalance between p - and f -spots. After considering the three problems, it is suggested: (a) the unipolar (positive and negative) magnetic regions are one of the basic structures generated by the internal solar dynamo, not remnants of old active regions; (b) single positive spots (unipolar magnetic spots) are born in a positive unipolar magnetic region (vice versa), and p -spots induces f -spots across the boundary of two unipolar magnetic regions, forming pairs of spots (positive p -spots on the positive unipolar region and negative f -spots in the negative unipolar regions. This consideration is based on three observations: (i) positive single spots tend to be present in positive unipolar magnetic regions (vice versa); (ii) pairs of spots tend to be present at the boundary unipolar magnetic regions, not elsewhere (iii) unipolar magnetic regions are of internal origin, perhaps generated by the internal dynamo. It seems that the three problems are explainable all together by the present consideration.

Keywords: single spots, pairs of spots, unipolar magnetic regions

Volume 2 Issue 3 - 2018

Syun Ichi Akasofu

International Arctic Research Center, University of Alaska Fairbanks, Alaska, USA

Correspondence: Syun Ichi Akasofu, International Arctic Research Center, University of Alaska Fairbanks, PO Box-757340 930, Koyukuk Dr. 415, Akasofu bldg Fairbanks, Alaska 99775-7340, USA, Tel 9074746012, Emails sakasofu@iarc.uaf, sakasofu@alaska.edu

Received: April 9, 2018 | **Published:** May 01, 2018

Introduction

The rising magnetic flux tube model by Babcock et al.¹ has long been almost universally accepted in considering the formation of sunspots (hereafter, this model is referred as the RF model), in which sunspots appear as a pair of positive and negative spots as the basic unit. In spite of a great observational and theoretical progress in studying sunspots during the last few decades, these studies have basically based on the RF model.

The formation of sunspots had been considered by many researchers in the past. After the discovery of magnetic fields by Harvey J et al.² the magnetic fields in pores and spots have become the main issue of sunspots. Early studies of the formation of sunspots are summarized.³⁻⁸ Recently, showed that their observations based on details of the emergence of an active region do not support the RF model. Thus, a few long-standing problems, which are contradictory with the RF model, have remained unsolved. They are:

- (1) The presence of magnetic monopole-like single spots.
- (2) The delay of appearance of f -spot after p -spot.
- (3) The unbalance of the magnetic flux between a p -spot and f -spot.

The purpose of this paper is to suggest that these three problems may be explainable on the basis of three observations.⁹⁻¹¹

- (i) Positive single spots tend to be present in positive unipolar magnetic
- (ii) Region (Vice Versa).
- (iii) Pairs of spots tend to appear at the boundary of two unipolar magnetic polar regions, not in the middle of unipolar magnetic

regions.

- (iv) Unipolar magnetic regions are not always remnants of old active regions, but are the basic feature of solar magnetism.

These three observations may be illustrated in Figure 1. An example of solar magnetic records is shown in Figure 1a: the region marked "1" are positive (red) and negative (blue) unipolar magnetic regions, respectively.

The regions marked by "2" have pores, single spots and groups of single spots (positive and negative). A distinct group of pairs of positive and negative spots are marked by "3". Based on a large number of magnetic records, Figure 1b summarizes schematically the close relationship among single spots, pairs of spots and unipolar regions, including also the well-known other features of pairs of spots, such as the reversal of the polarity of spots in the northern and southern hemispheres and during the even and odd sunspot cycles, and also the p -spots are located a little equatorward of f -spots.⁹⁻¹¹ It is particularly emphasized that the unipolar magnetic regions have a crucial importance in explaining the three long-standing problems (1), (2) and (3).

Unipolar magnetic regions

It has long been considered that some of the unipolar magnetic regions are remnants of old active regions, which are elongated by the non-uniform rotation of the sun.¹² However, (α) the distribution of unipolar regions and their size are quite different from locally elongated old active regions; (β) unipolar regions are present throughout the sunspot cycle, although their intensity waxes and wanes, following the sunspot cycle^{10,11} at the beginning and end of each cycle, they are still present, but their intensity is very low; (γ)

at the beginning of the cycle, unipolar regions appear before sunspots are born and (δ) some of the unipolar regions are often extend to the polar unipolar regions during certain periods of the sunspot cycle; this feature may be related to the migration of the magnetic fields toward the polar region, the reversal of the polarity of the solar dipole,¹³ and thus to the internal dynamo. It is unfortunate that the fundamental importance of unipolar regions has not been recognized in the past.

One of the internal dynamo models by Nakagawa et al.¹⁴ specifically considers the origin of positive and negative unipolar regions as shown in Figure 1c: he noted: “The unipolar regions in

Figure 1 are identified with the variations of the magnetic field near the surface, so that the flux coil which appeared near the surface is identified as one polarity, while the submerged region is the opposite polarity.” Therefore, unipolar magnetic regions are different from remnants of old active regions.^{15,16} It is suggested that *unipolar magnetic regions are one of the basic features in solar magnetism related to the internal dynamo and that positive unipolar regions are the birth place of positive single spots (vice versa)*. Thus, they should be distinguished from old remnants of active regions. This interpretation of unipolar magnetic regions is crucial in explaining the three contradictory problems.

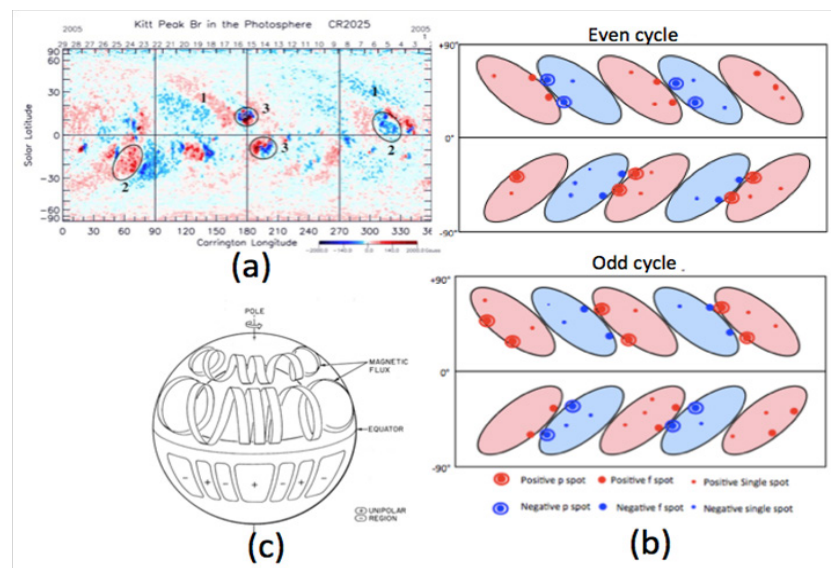


Figure 1 The relationship between unipolar regions and sunspots. (a) An example of solar magnetic records: the regions marked by “1” are positive (red) and negative (blue) unipolar regions, respectively. The regions marked by “2” have pores, single spots and groups of single spots (positive and negative). A distinct group of pairs of positive and negative spots are marked by “3”. (In this study, a color presentation of the magnetic fields is crucial, because for examples, it is not easy to recognize unipolar regions (which tend to have a low intensity)). (b) Assembly of the available information, regarding single spots, sunspot pairs and their relationship among positive and negative magnetic unipolar regions, for both even and odd cycles.¹¹ (c) The solar dynamo model,¹⁴ showing magnetic fluxes in the northern hemisphere; its photospheric fields are shown as unipolar magnetic field regions.

Single spots

Although single magnetic monopole-like spots (or often called ‘isolated’ or ‘independent’) spots are illustrated in many books or review papers,^{3,8,17–19} they have not been considered in recent years. Figure 2a shows an example of single spot. Since sunspots must appear as a pair of positive and negative spots in the RF model, the presence of single, isolated (or independent) spots has long been a puzzle, in addition to the fact that magnetic mono-poles cannot exist.

A high resolution image of unipolar magnetic region shows the well-known magnetic network structure, resulting from the convective flow; an example is shown in Figures 3a & 3b. It has been understood that the convective flow in each cell continuously sweeps magnetic fields to its boundary. This feature is schematically illustrated in Figure 3 by combining two convective cell structures (in a negative unipolar region) constructed.²⁰

It has also been considered that some irregularities in the convection flow are expected to concentrate the magnetic fields more than other places along the boundary (two pores indicated by

arrows in Figure 3b). This feature is schematically shown by a large (blue) pore or small spot in Figure 3c. It has also been considered that pores thus formed assemble *along the cell boundary* to form small spots and further larger spots. This process is morphologically called ‘*coalescence*’.⁹ In fact, one of the highest resolution observations shows that a large spot consists of an assembly of small spots. There have been many ideas on this process.²¹

Since magnetic monopoles cannot exist, we consider the formation of single spots as a *local phenomenon within unipolar magnetic regions*, in which positive single spots are born in a positive unipolar region (vice versa). The polarities of single spots are the same as the polarity of the unipolar regions, so that positive single spots are result of the coalescence in positive unipolar magnetic regions (vice versa). Thus, monopole-like single spots within unipolar regions satisfy the condition of $\text{div} \cdot B = 0$, since this condition is globally satisfied so long as unipolar magnetic regions are a result of internal dynamo. The fact that a positive unipolar region is the birth place of a positive single spot (vice versa), not elsewhere, has advanced toward the solution of this problem. Therefore, the problem (1) may be explained in this way.

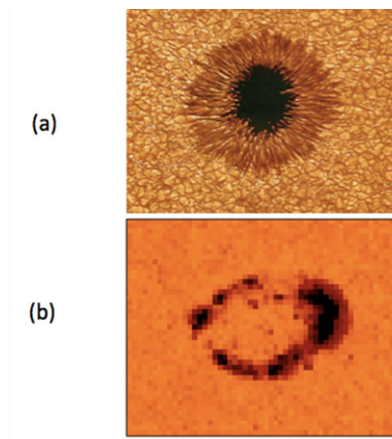


Figure 2 (a) An example of single spots (Courtesy of the Big Bear Observatory). (b) An example of a pair of positive and negative spots; a largest spot on the left side is a negative spot and largest spot on the right side is a positive spot (October 27, 2003; Courtesy of the Kitt Peak Solar Observatory).

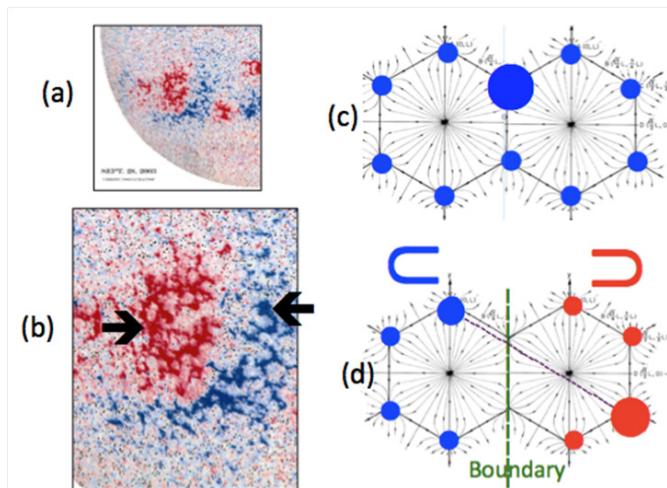


Figure 3 (a) & (b) show the magnetic network structure near the boundary of two unipolar magnetic regions; two pores are pointed out by arrows (September 28, 2003; Courtesy of the Kitts Peak Observatory). (c) Schematic illustration of the network in a negative unipolar region, 19 one large blue dot indicates a negative pore. (d) Schematic illustration of the magnetic network near the boundary of two unipolar magnetic regions, together with a pair of large negative (blue) and positive (positive) spots. Both might eventually form a pair of spots.

Pair of spots

Occurrence at the boundary

If magnetic buoyancy (or an upward flow) is the only reason for the rise of a magnetic flux tube in the RF model, a pair of positive and negative spots may appear together randomly at any location on the photosphere, even in the middle of unipolar regions. This is not the case, and thus is a serious problem for the RF model, as well as the three problems. We know now that a pair of sunspots tends to appear at or near the boundary of positive and negative unipolar regions, a positive spot in a positive unipolar regions (vice versa), not in the middle of unipolar magnetic regions.⁹⁻¹¹ Actually, this is collateral with the fact that a positive spot of the pair appears in a positive unipolar

region (vice versa). Based on this fact, there is a possibility that a positive p -spot in a positive unipolar magnetic region may induce a negative f -spot in a negative unipolar region across the boundary (vice versa). This possibility is explained in the following. In Figure 3b (and schematically illustrated in Figure 3d), there are many positive U-shaped “half” cells (lying horizontally) in a positive unipolar region and negative U-shaped “half” cells in a negative unipolar region at or near the boundary region of two unipolar regions. Thus, it is expected that positive pores develop *along* positive U-shaped cells at the boundary of a positive unipolar region (vice versa).

Then, positive pores are coalesced *along the U-shaped boundary* to form a small positive single spots, and further then these small single spots form a larger single spot or spots by coalescence; for the coalescence process.²⁰ If one of a positive U-shaped half cell in a positive unipolar region and one of a negative U-shaped half cell in a negative unipolar region join in this way, they can form a pair of spots or pairs along the elliptical boundary. *McIntosh (1981) emphasized that the elliptical formation is a typical way by which a pair of spots are formed; in the past, this fact is not well appreciated.* Figure 2b is a good example of a pair of spots thus formed; in this particular case, small spots along the elliptical boundary coalesced to form small spots and finally a distinct pair of spots; this particular pair of spots developed almost in full in a period of less than 5 hours. A number of such examples are shown by McIntosh PS et al.⁹

Delay of f -spots

According to the RF model, a p -spot and f -spot should appear simultaneously. However, it has long been known that a pair of spots does not appear at the same time; as so named, p (primary) spots appear earlier than f (following) spots.^{5,6} There are only a few cases, in which f -spots appear first, so that such cases are rare.⁹

There is a possibility that a p -spot induces a f -spot across the boundary of two adjacent unipolar regions (of opposite polarity), if a positive U-shaped cell and a negative U-shaped cell are connected by magnetic field lines. Then, some delay might be expected as the communication time. This might be a possible explanation on the problem (2). The formation of a p -side of the U-shaped half cell could be communicated to f -side of U-shaped cell, because magnetic field lines are expected to be equi-potential in the solar atmosphere, so that any electromagnetic changes associated with the coalescence (e.g. a converging flow, $V \times B$) in a p -side U will be communicated to a f -side U by Alfvén waves or other processes, inducing a similar coalescence at the other end of the magnetic field lines, namely the formation of f -spots of the *opposite* polarity. In considering the converging flow, it is noted that Kotov et al.²² showed that there is 10^{12} A of electric currents around a single spot.

Unbalance of the magnetic flux

It is well known that there is often a significant unbalance of the magnetic flux between two spots of a pair.^{6,8} Figure 2b is a good such example. This unbalance is not easy to explain in terms of a simple RF model. However, if f -spots can be induced by p -spots as suggested in the above, it provides an opportunity to consider this problem of unbalance. When a p -spot is forming in a positive unipolar region, it does not have to induce an *identical* f -spot just across its neighboring unipolar region, because the field distribution in the two unipolar magnetic regions across the boundary or in the surrounding regions is different (Figure 3b). In this respect, it may be recalled that X ray

observations of the corona showed that some of the magnetic field lines in a pair of spots are connected to at least another nearby pair of spots even across the magnetic equator,²³ so that it is not necessary to consider the induction within just one pair of spots. This might provide a possible explanation on the problem (3).

Actually, instead of considering the delay of f -spot, it is more appropriate to consider *why one specific spot, namely p -spot, appears first before f -spot*. At this point, we cannot find definitive morphological observations in explaining why a p -spot appears first at the specific boundary of two unipolar regions, not the other side, except that the specific side of the unipolar region tends to be more active than the other side. Noted this particular feature of the boundary and called it the “Hale boundary”.

The sunspot cycle

Any idea of spot formation must explain the equatorward shift during the sunspot cycle. There are a number of dynamo theories on the butterfly diagram, explaining the equatorward shift of the formation of sunspot pairs during the sunspot cycle. Since we are considering only *morphologically* this problem in this paper, it may be noted that there is an interesting observation which shows that an east–west belt of ‘torsional oscillation’ (rotation or anti–parallel flow) on the photosphere shifts from poleward to equatorward during each sunspot cycle.²⁴ This observation may be related to the process of coalescence.²¹

In addition, the east–west alignment of sunspot pairs can be explained by the fact that the boundary of two unipolar magnetic regions tends to align along the meridian lines. Actually, it is somewhat deviated from the meridian line, so that it can explain why p -spots are located a little equatorward than f -spots; Figure 1b.

Conclusions

Although the RF model has long been well established and has many supporting observations,²⁵ there are a few long–standing problems (1, 2, 3), which are still very difficult to attempt to solve on the basis of a simple RF model, since they are basically contradictory to the RF model. Thus, a new morphological study on the formation of sunspots is made by taking into account of a number of past studies and observations: (i) positive single spots tend to be present in positive unipolar regions (and vice versa), (ii) a pair of spots is formed at the boundary of two unipolar regions and (c) unipolar regions have the internal origin. These observations have made it possible to consider the formation of sunspots, enabling to consider the three contradictory problems.

It is concluded: (1) *single spots are the basic element in the formation of sunspots, not pairs of spots, and (2) unipolar magnetic regions are also one of the basic structures of internal origin of the solar dynamo process. The three contradictory problems seem to be explainable on these bases.*

In short, a new morphological study of unipolar magnetic regions and their relationship with sunspots are needed. The results presented here suggest that there may be new theories to consider the formation of sunspots than the RF model, which consider not only the three problems and others, such as why p -spots appear first before f -spots. It is hoped that this paper will become a ‘seed’ in considering the formation of sunspots anew.²⁶

Acknowledgements

The author would like to acknowledge Dr. Kees de Jager for his comments on drafts of this paper. He would like to acknowledge also the Kitt Peak Solar Observatory and the Big Bear Solar Observatory for their magnetic records used in this paper.

Conflict of interest

Author declares there is no conflict of interest.

References

1. Babcock HW. The topology of the sun’s magnetic field and the 22–year cycle. *The Astrophysical Journal*. 1969;133:572–587.
2. Harvey J. Hale’s Discovery of Sunspot Magnetic Fields. *The Astrophysical Journal*. 1999;525:60–60.
3. Abetti G. The Sun, Faber and Faber. London: 24 Russel Square; 1951.
4. Cowling TG. Solar Electrodynamics. In: Kuiper GP, editor. USA: The University of Chicago Press; 1953. p. 532–591.
5. Kiepenheuer KO. Solar activity. In: Kuiper GP, editor. 1953.
6. De Jager C. Structure and dynamics of the solar atmosphere. In: Flugge S, editor. Germany: Springer; 1959. p. 80–362.
7. Menzel DH. The Sun. USA: The Blackiston Comapany; 1959.
8. Bray RJ, Loughhead RE. Sunspots. USA: Dover Publications; 1964
9. McIntosh PS. The Physics of Sunspots. In: Crom LE, Thomas JH, editor. Mexico: Sacramento Peak Observatory; 1981. p. 7–54.
10. Akasofu SI. Single spots, unipolar magnetic regions, and pairs of spots. *Geophysical Research Letters*. 2014;41(11):3698–3700.
11. Akasofu SI. Single spots, unipolar magnetic regions, and pairs of spots: 2. The development of sunspot pairs and the Hale boundary. *Geophysical Research Letters*. 2015;42(8):2571–2576.
12. Leighton RB. *The Astrophysical Journal*. 1969;156:1–26.
13. Hathaway DH. The solar cycle. *Living Reviews in Solar Physics*. 2010;7:1–65.
14. Nakagawa Y. A numerical study of the solar cycle. *Solar Magnetic Fields*. 1971;43:725–736.
15. McIntosh PS, Wilson PR. A new model for flux emergence and the evolution of sunspots and the large–scale fields. *Solar Physics*. 1985;97(1):59–79.
16. Watson PR. The reversal of the solar polar magnetic fields, III, The large scale fields and the first major active regions of cycle 22. *Solar Physics*. 1992;138(1):11–21.
17. Noyes RW. The Sun, Our star. USA: Harvard University Press; 1982. 263 p.
18. Solanki SK. Sunspots: An overview. *The Astronomy and Astrophysics Review*. 2003;11(2–3):153–286.
19. Svalgaard L, Wilcox JW. The Hale solar sector boundary. *Solar Physics*. 1976;49(1):177–185.
20. Clark A, Johnson AC. Magnetic–field accumulation in supergranules. *Solar Physics*. 1967;2(4):433–440.
21. Parker EN. Vortex attraction and the formation of sunspots. *The Astrophysical Journal*. 1992;390(1):290–296.
22. Kotov VA. Solar Magnetic Fields. In: Howard RD, editor. Netherlands: Reidel Publishing Company; 1971. p. 213–219.

23. Sheeley NR. Energy released by the interaction of coronal magnetic fields. *Solar Physics*. 1976;47(1):173–180.
24. Howard R, Labonte BJ. *The Astrophysical Journal*. 1980;239:L33–L36.
25. Itonidis S, Zhao J, Kosovichev A. Detection of Emerging Sunspot Regions in the Solar Interior. *Science*. 2011;333(6045):993–996.
26. Jager CD, Akasofu SI, Duhau S, et al. The remarkable recent phase transition in the solar dynamo. *Space Science Reviews*. 2016;201(1–4):109–145.