Toroidal Flux Loss due to Flux Emergence Explains why Solar Cycles Rise Differently but Decay in a Similar Way

Akash Biswas[●],[1] Bidya Binay Karak[●],[1,*] and Robert Cameron[●][2]

[1]*Department of Physics, Indian Institute of Technology (Banaras Hindu University), Varanasi 221005, India*
[2]*Max Planck Institute for Solar System Research, Justus-Von-Liebig-Weg 3, D-37077, Göttingen, Germany*

A striking feature of the solar cycle is that at the beginning, sunspots appear around midlatitudes, and over time the latitudes of emergences migrate toward the equator. The maximum level of activity (e.g., sunspot number) varies from cycle to cycle. For strong cycles, the activity begins early and at higher latitudes with wider sunspot distributions than for weak cycles. The activity and the width of sunspot belts increase rapidly and begin to decline when the belts are still at high latitudes. Surprisingly, it has been reported that in the late stages of the cycle the level of activity (sunspot number) as well as the widths and centers of the butterfly wings all have the same statistical properties independent of how strong the cycle was during its rise and maximum phases. We have modeled these features using a Babcock-Leighton type dynamo model and show that the flux loss through magnetic buoyancy is an essential nonlinearity in the solar dynamo. Our Letter shows that the nonlinearity is effective if the flux emergence becomes efficient at the mean-field strength of the order of 10^4 G in the lower part of the convection zone.

Solar activity, as measured, for example, by the number of sunspots on the solar surface, takes place in 11-year cycles. The strength of each cycle (e.g., the maximum number of sunspots) varies from cycle to cycle. Strong and weak cycles have systematic differences: compared to weak cycles, strong cycles begin at higher latitudes, rise more rapidly, reach their maxima earlier and consequently have a longer decline phase [1]. This is known as the Waldmeier effect which has also been confirmed in the reconstructed solar activity data over millennial timescale [2]. A stronger constraint is that all cycles behave in the same way toward the end of the cycle ([3], hereafter CS16). This is illustrated in Fig. 1, and we refer the reader to the work by Hathaway [4] and CS16 for the observational analysis. The fact that the cycles all have the same properties (amplitude, spatial distribution) in the late phases of the cycle despite having different amplitudes in the early phase of the cycle indicates a nonlinearity acting as the cycle progresses. The nonlinearity is particularly important because during the late phase of the cycle sunspots emerge closer to the equator and hence contribute more to the buildup of the Sun's polar fields [5]. [The magnetic flux of the leading polarity of the bipolar magnetic regions (BMRs) that emerge closer to the equator gets easily carried across the equator by small-scale convective motions. It is the cross-equatorial cancellation of this flux that changes the net flux in each hemisphere and the polar fields at the end of each cycle [6]]. The nonlinearity hence leads to the change in polar field from one cycle to the next being only weakly dependent of the cycle strength [7].

In recent years, Babcock-Leighton dynamo models have explained many features of the irregular solar cycle [9]. For example, Karak and Choudhuri [10] explained the Waldmeier effect using a Babcock-Leighton type flux transport dynamo model. More recently, Mandal *et al.* [11] reproduced the basic correlations described above using a Babcock-Leighton dynamo model. However, why they were successful in this regard was not discussed. Here, we shall employ a Babcock-Leighton type dynamo model to explain the features reported by Waldmeier [12] and CS16. We shall further show that in order to reproduce these features, some constraints on the nonlinearity in terms of the equipartition field strength and the variation of meridional flow can be inferred.

For our Letter, we use an updated version of the code SURYA [13,14], which solves the following equations:

$$\frac{\partial A}{\partial t} + \frac{1}{s}(\boldsymbol{v}_p \cdot \boldsymbol{\nabla})(sA) = \eta_p\left(\nabla^2 - \frac{1}{s^2}\right)A + \alpha B, \quad (1)$$

$$\frac{\partial B}{\partial t} + \frac{1}{r}\left[\frac{\partial(rv_rB)}{\partial r} + \frac{\partial(v_\theta B)}{\partial \theta}\right]$$
$$= \eta_t\left(\nabla^2 - \frac{1}{s^2}\right)B + s(\boldsymbol{B}_p \cdot \boldsymbol{\nabla})\Omega + \frac{1}{r}\frac{d\eta_t}{dr}\frac{\partial(rB)}{\partial r}, \quad (2)$$

where r is radial distance from the center of the Sun, θ is the colatitude, $B(r,\theta)$ is the toroidal field, and $A(r,\theta)$ is the ϕ component of the magnetic vector potential of the poloidal magnetic field $\boldsymbol{B}_p$, $s = r\sin\theta$, $\boldsymbol{v}_p = v_r\hat{\boldsymbol{r}} + v_\theta\hat{\boldsymbol{\theta}}$ is the meridional circulation, $\Omega(r,\theta)$ is the local rotation rate,
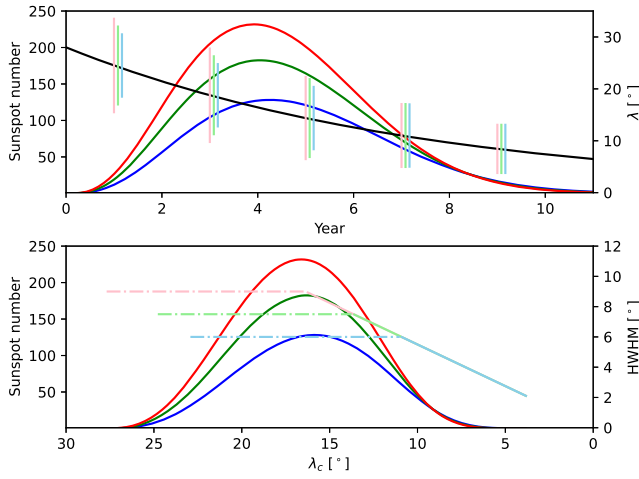
FIG. 1. Illustration of the behavior of cycles of different strengths. Upper panel: the three colored curves show the sunspot number for cycles of weak (blue), moderate (green), and high (red) levels of activity. The curves are based on the empirical curve fitting of Hathaway *et al.* [8]. The sunspots appear in butterfly wings, and the central latitude of the butterfly wings drift toward the equator in a way which is independent of cycle strength (black curve) [4]. Finally, the width of the butterfly wings (light vertical lines) depends on the cycle strength early in the cycle but is independent of the cycle strength late in the cycle [3]. The properties of the butterfly diagram in the late phase of all the cycles are the same, i.e., all cycles die in the same way although they have different properties during the rise phase of the cycle [3]. Lower panel: because all cycles drift toward the equator in the same way, the central latitude of the butterfly wings can be used as the independent variable rather than time from the start of the cycle. The solid light lines indicate the width of the butterfly wings as a function of latitude. The dashed light lines are the upper limit for the width during the early phases of the cycle: the width increases as the cycle progresses until it reaches this level. Thereafter, the cycle begins to decline.

$\eta_p(r)$ and $\eta_t(r)$ are the turbulent diffusivities for the poloidal and toroidal fields. The coefficient $\alpha(r, \theta)$ captures the generation of the poloidal field near the solar surface through the decay and dispersal of the tilted BMRs (Babcock-Leighton process) in our axisymmetric model.

As the profiles of all parameters are given in many publications (e.g., [10,14–17]), we discuss them in Supplemental Material [18] except α and magnetic buoyancy, which play important roles in the present Letter.

In the regular model, the α has the following form:

$$\alpha(r, \theta) = \frac{\alpha_0}{4} \cos \theta \left[1 + \mathrm{erf}\left(\frac{r - 0.95 \, R_\odot}{0.03 \, R_\odot} \right) \right]$$
$$\times \left[1 - \mathrm{erf}\left(\frac{r - R_\odot}{0.03 \, R_\odot} \right) \right], \tag{3}$$

where $\alpha_0 = 30 \, \mathrm{m \, s^{-1}}$.

In the Babcock-Leighton dynamo, generation of the poloidal field involves some randomness, primarily due to scatter around Joy's law and the random flux emergences. To capture these effects in our model, we include stochastic noise in the above α by replacing α_0 by $\alpha_0[1 + \sigma(t, \tau_{\mathrm{corr}})]$, where $\sigma(t, \tau_{\mathrm{corr}})$ is a random number drawn from a uniform distribution in the range $[-1, 1]$ and τ_{corr} is the coherence time over which the value of α is constant. In all simulations, $\tau_{\mathrm{corr}} = 1$ month (typical lifetime of sunspots). With these parameters, the variations in the solar cycle remain consistent with what is seen in the last 300 years.

We take the profile for meridional circulation to be of the same functional form as in [14] with one change in the parameters: we take $\epsilon = 3$ instead of $\epsilon = 2$ as given there. This change makes the meridional circulation to increase slightly at the lower latitudes and decrease at higher latitudes compared to the previous profile [14]; see Ref. [18] for details. With this slight modification in the flow, the equatorward migration of the sunspot belt is in better agreement with the observations.

Finally, we describe the only nonlinearity that has been included in the model. The nonlinearity concerns the time at which flux is assumed to emerge. Every 0.6 days, the code checks the amplitude of the toroidal field at each grid point above the base of the convection zone (BCZ). If the value is above a to-be-described critical value (B_c) then it is assumed that half of the flux emerges—its value at that grid point is halved and this half is added to the toroidal field near the surface at the same latitude. This is how the sunspot eruptions are modeled and this mechanism ensures that with each eruption a part of the toroidal flux from the CZ is lost. When the threshold is met, the toroidal flux is removed from the lower CZ so that it is no longer available to emerge at lower latitudes. This flux loss associated with a threshold is the only nonlinearity we have included. Such a nonlinearity tends to make all cycles behave in the same way during the late phases of the cycles because in the late phase a lot of the toroidal flux associated with a strong cycle has already been removed. It acts to saturate the dynamo because emergences at high latitudes tend to be ineffective in terms of changing the polar fields [5].

We now discuss the results of the dynamo simulation. The latitudinal distribution of the surface radial field for eight cycles from a long run of 40 cycles is shown in Fig. 2(a). We observe that this plot resembles the basic features of the solar cycle reasonably well [9,19]. We find that when $B_c = 0.8 \times 10^4$ G, the strength of the radial field on the surface is in agreement with the observed range [20] and the subsurface flux loss through magnetic buoyancy becomes consistent with the observed flux loss through BMR emergences [21]. We use this value of B_c for the rest of this Letter.

We compute the theoretical sunspot number by tracking the latitudes of eruptions of the toroidal flux (where the bottom field surpasses B_c) and with this number, we
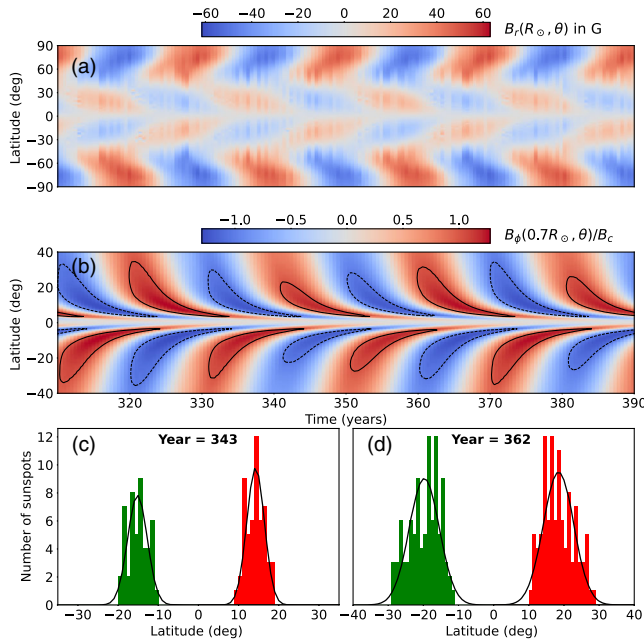
FIG. 2. A typical portion from our simulation. Latitudinal distributions of (a) the surface radial field and (b) the toroidal field at BCZ (the contour represents $B = B_c$). (c) and (d) show the annual latitudinal distributions of sunspots for the years 343 and 362, respectively, and the fitted Gaussian profiles (black curves).
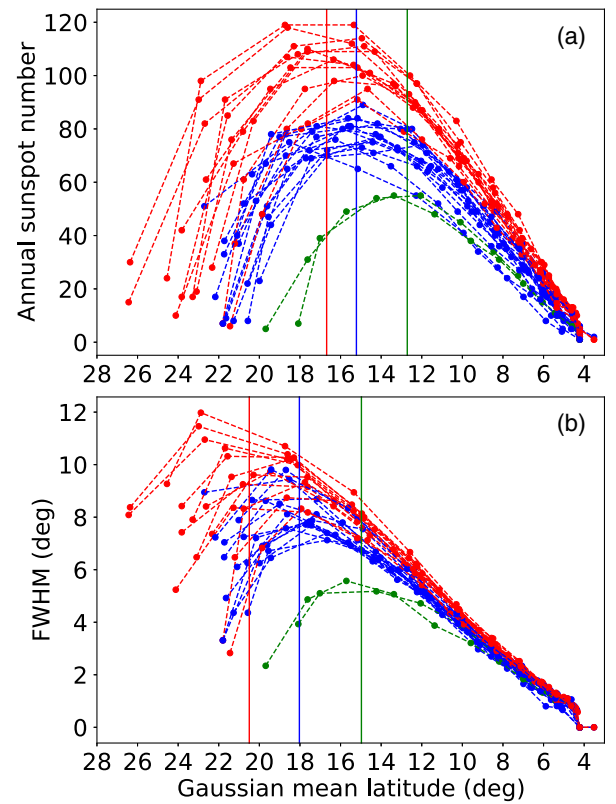

FIG. 3. (a) Model (pseudo) sunspot number for each year (based on the number of flux emergences) as a function of the central latitude of the Gaussian distribution. Different curves correspond to different cycles. (b) Same as (a) but for the distribution width (FWHM). Cycles begin at the left side of the plots and migrate to the right as they evolve. Vertical lines guide the average of the Gaussian-mean latitudes for strong (red), moderate (blue), and weak (green) cycles at their peaks.

perform the same analyses as done in CS16 for the observed data. In Figs. 2(c) and 2(d), we show that the latitudinal distribution of annual (model) sunspot emergence can be approximated with a Gaussian distribution. The solar cycle variations of the total annual sunspot number and the width of sunspot distribution with the central latitudes of distribution are shown in Fig. 3. These behaviors largely agree with the observations (CS16).

We now explain these results based on the dynamo theory. In our model, cycles vary in strength due to the fluctuations introduced in the Babcock-Leighton α term. Suppose for some time α is large in a cycle, then more poloidal field will be produced in that cycle. The poloidal field is sheared by the differential rotation to produce the next cycle's toroidal field (we note that the shear is strongest at about 55° [22]). In our model, this toroidal field emerges once a critical threshold is reached, and gives rise to the (pseudo) sunspots on the solar surface. Hence, a high value of α in a cycle produces a strong poloidal field. This generates a strong toroidal field for the next cycle and the eruption condition ($B > B_c$) is satisfied at earlier times when the field is at high latitudes. This explains why, for a strong cycle, sunspots start appearing at high latitudes and the width of the latitude belt is large. On the other hand, when a cycle is weak, the model takes a long time for the toroidal field to satisfy the spot-eruption condition by that time the meridional circulation drags the toroidal field toward the low latitudes. Hence, in a weak cycle, the

sunspot belt begins at lower latitudes. The band of the sunspot latitudes is also narrow when the toroidal field is weak (as the spot eruption condition is satisfied only in a narrow latitude band).

Now we consider the effect of the loss of toroidal flux due to flux emergence [21]. Again we consider a strong cycle for which emergence begins early in the cycle at high latitudes. Each emergence reduces the subsurface toroidal flux [21] so that a strong cycle which has many early flux emergences rapidly loses toroidal flux until the subsurface mean field strength is just above B_c. Thereafter the cycle begins to decline and the toroidal flux is advected equatorward by the meridional flow while maintaining a strength just above B_c. The decrease in the meridional flow as we approach the equator will tend to increase the subsurface field strength, but this will be compensated for by further flux loss due to flux emergence.

On the other hand, in a weak cycle, there will be very few eruptions early in the cycle, and the field strength will simply build up as the flux is advected toward the stagnation point at the equator. At some point, however,

the mean field will be comparable to B_c and the cycle will enter its decline phase. During this phase, the situation is entirely the same as for a strong cycle: the increase in the field strength as the flux builds up near the stagnation point is balanced by the loss of flux due to emergence, and the mean field strength is kept around B_c. A nonlinear process which causes flux emergence rates to be enhanced when the mean field exceeds B_c can explain why all cycles decline in the same way as is evident from the right part of the curves in Fig. 3. The value of $B_c = 0.8 \times 10^4$ G was chosen so that the model matches the observed range of the radial magnetic field on the solar surface and the amount of flux loss. Incidentally, this value is close to the equipartition field strength at BCZ ($\sim 10^4$ G). Although we did not choose the value of B_c keeping this in mind, it is interesting to note that just by constraining our model parameters through observations, we get a value of B_c, which is close to the equipartition field strength. A nonlinearity is expected because at this field strength the magnetic field has an energy density comparable to the kinetic energy density of the turbulent motions.

One quantitative discrepancy between our model presented in Fig. 3(b) with respect to the observations (CS16) is that in the latter, we observe that as soon as the distance between the center of the Gaussian and the equator is roughly equal to the FWHM, all cycles begin to decay. The width of the butterfly wings is substantially smaller in our model; see Fig. 3(b). In this Letter, we have been interested in investigating the role of flux loss in combination with a threshold for the emergence of active regions. For this purpose we have used a simple sharp threshold which involves two parameters B_c and the fraction of flux which emerges when an emergence takes place. A better match with the observed widths might be possible if the flux emergence recipe had more degrees of freedom, however, the basic physical idea is captured with this simple threshold. An additional possibility is that the butterfly wing widths are broadened by convective buffeting of the rising flux tubes before they reach the surface.

We have also checked the Waldmeier effect, i.e., WE1 and WE2 [18]. We find linear (Pearson) correlation coefficients -0.30 and 0.51 for these two cases (Fig. 4). Thus, the classical Waldmeier effect is also reproduced in this simulation. Earlier, Karak and Choudhuri [10] showed that fluctuations in meridional circulation are needed to reproduce WE1. Our result differs from this because WE1 is a weak anticorrelation and is sensitive to the way in which the data (and simulations) are treated. We find that a weak anticorrelation in WE1 exists even without including fluctuations in the meridional circulation. When fluctuation in meridional circulation is introduced, it enhances this anticorrelation. However, it leads to the model disagreeing with the fact that the widths of the sunspot latitude bands of all cycles are the same function of the central latitudes of the sunspot bands (CS16; Fig. 3). Hence, observational
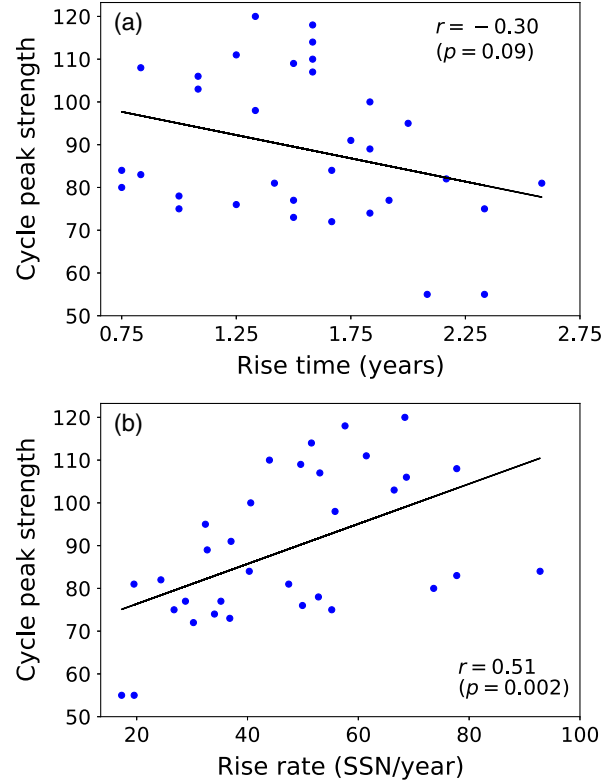

FIG. 4. The scatter plots between the peak of sunspot cycle (amplitude) and (a) the rise time and (b) the rise rate, i.e., WE1 and WE2, respectively.

features of sunspot cycle [12] as analyzed in CS16 suggest that there was no large variation in the deep meridional circulation in the past 300 years.

In this Letter we have focused on explaining why all cycles are statistically the same in their decay phase. The dynamo simulation can be modified to include other effects such as a time delay between the start of the rise of a flux tube and when it emerges at the surface. This delay time can be of the order of months [23]. Such a delay can affect the properties of the dynamo cycles [24,25]. We show [18] that the inclusion of such a delay in our model does not substantially affect the conclusions of this Letter, although it does have a weak effect on the latitudes at which the emergences take place.

In conclusion, we have shown that the main features of the latitudinal distribution of sunspots as reported in Waldmeier [12] and CS16 are reproduced in a Babcock-Leighton type flux transport dynamo model with stochastic fluctuations in the poloidal field source. We find that a constant equatorward flow near BCZ and a reduction of toroidal field due to flux emergence are essential to reproduce these results. Further, the critical strength of the mean magnetic field for the flux emergence through buoyancy is found to be of the order of 10^4 G. This is about the equipartition value where the magnetic energy density is equal to the kinetic energy density of the turbulent convective motions.

----

*karak.phy@iitbhu.ac.in

[1] M. Waldmeier, Astron. Mitt. Eidg. Sternwarte Zurich **14,** 105 (1935).

[2] I. G. Usoskin, S. K. Solanki, N. A. Krivova, B. Hofer, G. A. Kovaltsov, L. Wacker, N. Brehm, and B. Kromer, Astrophys. J. **649,** A141 (2021).

[3] R. H. Cameron and M. Schüssler, Astrophys. J. **591,** A46 (2016).

[4] D. H. Hathaway, Sol. Phys. **273,** 221 (2011).

[5] K. Petrovay, M. Nagy, and A. R. Yeates, J. Space Weather Space Clim. **10,** 50 (2020).

[6] C. J. Durrant, J. P. R. Turner, and P. R. Wilson, Sol. Phys. **222,** 345 (2004).

[7] J. Jiang, Astrophys. J. **900,** 19 (2020).

[8] D. H. Hathaway, R. M. Wilson, and E. J. Reichmann, Sol. Phys. **151,** 177 (1994).

[9] P. Charbonneau, Living Rev. Solar Phys. **17,** 4 (2020).

[10] B. B. Karak and A. R. Choudhuri, Mon. Not. R. Astron. Soc. **410,** 1503 (2011).

[11] S. Mandal, B. B. Karak, and D. Banerjee, Astrophys. J. **851,** 70 (2017).

[12] M. Waldmeier, *Ergebnisse und Probleme der Sonnenforschung* (Geest & Portig, Leipzig, 1955).

[13] D. Nandy and A. R. Choudhuri, Science **296,** 1671 (2002).

[14] P. Chatterjee, D. Nandy, and A. R. Choudhuri, Astrophys. J. **427,** 1019 (2004).

[15] B. B. Karak, Astrophys. J. **724,** 1021 (2010).

[16] B. B. Karak, S. Mandal, and D. Banerjee, Astrophys. J. **866,** 17 (2018).

[17] B. B. Karak and R. Cameron, Astrophys. J. **832,** 94 (2016).

[18] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.129.241102 for model details and some supporting results (2022).

[19] D. H. Hathaway, Living Rev. Solar Phys. **12,** 4 (2015).

[20] A. V. Mordvinov, B. B. Karak, D. Banerjee, E. M. Golubeva, A. I. Khlystova, A. V. Zhukova, and P. Kumar, Mon. Not. R. Astron. Soc. **510,** 1331 (2021).

[21] R. H. Cameron and M. Schüssler, Astrophys. J. **636,** A7 (2020).

[22] R. Howe, Living Rev. Solar Phys. **6,** 1 (2009).

[23] Y. Fan, G. H. Fisher, and A. N. McClymont, Astrophys. J. **436,** 907 (1994).

[24] L. Jouve, M. R. E. Proctor, and G. Lesur, Astrophys. J. **519,** A68 (2010).

[25] Y. Fournier, R. Arlt, and D. Elstner, Astrophys. J. **620,** A135 (2018).