Toroidal flux loss due to flux emergence explains why solar cycles rise differently but decay in a similar way

Akash Biswas, Bidya Binay Karak, and Robert Cameron

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

(Dated: October 14, 2022)

A striking feature of the solar cycle is that at the beginning, sunspots appear around mid-latitudes, and over time the latitudes of emergences migrate towards 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 study shows that the nonlinearity is effective if the flux emergence becomes efficient at the mean-field strength of the order of $10^5$ 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 time scale [2]. A stronger constraint is that all cycles behave in the same way towards the end of the cycle [3] hereafter CS16. This is illustrated in Figure 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 build-up of 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 study, we use an updated version of the code SURYA [13, 14], which solves the following equations:

\[ \frac{\partial A}{\partial t} + \frac{1}{s} (v_p \cdot \nabla)(sA) = \eta_p \left( \nabla^2 - \frac{1}{s^2} \right) A + \alpha B, \]  
\[ \frac{\partial B}{\partial t} + \frac{1}{r} \left[ \frac{\partial (rv_r B)}{\partial r} + \frac{\partial (v_\theta B)}{\partial \theta} \right] = \eta_r \left( \nabla^2 - \frac{1}{s^2} \right) B + s(B_p \cdot \nabla)\Omega + \frac{1}{r} \frac{\partial (r B)}{\partial r} \]  

where $r$ is radial distance from the center of the Sun, $\theta$ is the co-latitude, $B(r, \theta)$ is the toroidal field and $A(r, \theta)$ is the $\phi$ component of the magnetic vector potential of the poloidal magnetic field $B_p$, $s = r \sin \theta$, $v_p = v_r \hat{r} + v_\theta \hat{\theta}$ is the meridional circulation, $\Omega(r, \theta)$ is the local rotation rate, $\eta_p(r)$ and $\eta_r(r)$ 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 the Supple-
mental Material [18] except $\alpha$ and magnetic buoyancy, which play important role in the present work. In the regular model, the $\alpha$ has following form

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

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

In the Babcock–Leighton dynamo, generation of 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 $\alpha$ by replacing $\alpha_0$ by $\alpha_0 \left[ 1 + \sigma(t, \tau_{\text{corr}}) \right]$, where $\sigma(t, \tau_{\text{corr}})$ is a random number drawn from a uniform distribution in the range $[-1, 1]$ and $\tau_{\text{corr}}$ is the coherence time over which the value of $\alpha$ is constant. In all simulations, $\tau_{\text{corr}} = 1$ 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 [15] for details. With this slight modification in the flow, the equatorward migration of 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 modelled 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 Figure 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 perform the same analyses as done in CS16 for the observed data. In Figure 2(c-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 Figure 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 \( \alpha \) term. Suppose for some time \( \alpha \) 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 \( \alpha \) 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, sunspot starts 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 towards 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 towards 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 evident from the right part of the curves in Figure 3. The value of $B_c = 0.8\times10^4$ 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 Figure 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 Figure 3(b). In this Letter, we have been interested in investigating the role of flux emergence through buoyancy is found to be of the order 300 years.

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.
Supplemental Material

S1. Details of the Model Parameters

The profiles of all the ingredients of our dynamo model have been discussed in many previous publications [e.g., 10 14 15 17]. However, we discuss the profile of meridional circulation as it has been slightly modified here.

The meridional circulation is obtained from a stream function \( \psi \), such that \( \rho v_\theta = \nabla \times [\psi(r, \theta) \hat{\phi}] \), where \( \rho = C (\frac{R}{r} - 0.95)^{3/2} \), and

\[
\psi r \sin \theta = \psi_0 (r - R_\odot) \sin \left( \frac{\pi (r - R_\odot)}{(R_\odot - R_\odot)} \right) \left\{ 1 - e^{-\beta_2 \theta^2} \right\} \{ 1 - e^{-\beta_1 (\theta - \pi/2)} \} e^{(r - r_0)/\Gamma} \theta^2.
\]

Here the values of the parameters are: \( C = 0.407 \text{ kg m}^{-3}, \beta_1 = 1.5 \times 10^{-8} \text{ m}^{-1}, \beta_2 = 1.3 \times 10^{-8} \text{ m}^{-1}, r_0 = (R_\odot - R_\odot)/3.5, \Gamma = 3.1 \times 10^2 \text{ m}, R_p = 0.635 R_\odot, R_b = 0.55 R_\odot. \)

We take two profiles for the meridional flows which we call MC1 and MC2. In MC1, we take \( \epsilon = 3 \) and in MC2, we take \( \epsilon = 2 \). These profiles are shown in Figure 5. We note that MC1 is stronger in low latitudes and weaker in high latitudes. This MC1 profile has been used in our many previous publications 10 16 17.

S2. Results with MC2 meridional flow

In Figure 6, we show our results of the dynamo model in which the meridional circulation profile is changed to MC2 (see blue/dashed curves in Figure 5 for its profile). In this case, we notice that the agreement with observation is slightly poor, particularly, the butterfly wing begin at somewhat higher latitudes and the activity level of not all cycles strictly decline at the same rates.

The reason for this poor agreement with observation is that the equatorward drift of the meridional flow at low latitudes is weak which is critical in shaping the trajectory of the sunspot band over the cycle.

S3. Computations of Cycle Parameters

We have explicitly studied the two aspects of Waldmeier effect, namely WE1 and WE2. WE1 refers to the anti-correlation between the rise time and the amplitude of the cycle, while WE2 refers to the positive correlation between the rise rate and the amplitude [10]. To check these features, we need to compute the rise rate, rise time and amplitude of the cycle. For this, we first smooth the three-month binned data using a Savitzky–Golay filter. Due to the overlap between the consecutive cycles, finding out the minimum is difficult. To avoid this problem we follow the same procedure as in Karak and Choudhuri [10]. We take the rise time as the time taken by a cycle to grow its activity from 20% to 80% of its peak value. The rise rate is computed by dividing the difference between values of 80% and 20% strength by the rise time.

S4. Results with time delay in flux emergence

So far in our model, we have ignored the delay induced due the rise of the magnetic flux from the BCZ to the surface. The magnetic buoyancy depends on the strength of the magnetic field at the BCZ and as a result the buoyancy induced delay in the source term of the poloidal field depends on magnetic field [23 25]. Here, we assume a simple inverse square relationship of the delay with the magnetic field strength such that the maximum delay is 6 months. The formula to incorporate this process in the code is following: \( D = 180 \times (B_c/B)^2 \), where \( D \) is the number of days taken by the flux tube to rise from the BCZ to the surface (i.e. the delay in flux eruption), \( B_c \) is the critical field strength to satisfy eruption condition and \( B \) is the value of toroidal field strength at the grid point where the eruption condition \( (B > B_c) \) has been met. In our simulations we however do not observe a significant change in the magnetic cycle due to the magnetic field dependent delay (results shown in Figure 7). The reason for this is that in this model, once the magnetic field \( B \) surpasses \( B_c \) the flux loss \( (B \rightarrow B/2) \) condition does not allow the magnetic field to increase too much higher than \( B_c \) and thus the buoyancy delay varies only a little. However, it is interesting to note that on an average the mediocre and the weaker cycles (blue and green ones, respectively) shift slightly towards high latitude regions.

We note that the simulation presented here used the same initial condition that was used for the one presented in the Letter. The slight differences in the solutions (compare Figure 7 with Figure 5) are due to the time delay introduced in the flux emergence.

We thank both the anonymous referees for providing insightful comments which helped us to improve the manuscript to a large extent. A.B. and B.B.K. acknowledge financial support provided by ISRO/RESPOND (project No. ISRO/RES/2/430/19-20). RHC acknowledges the support of ERC Synergy Grant WHOLE SUN 810218.
FIG. 5. The meridional circulation profiles used in the model. (Left) Shows the latitudinal variation of the surface meridional circulation and (right) shows its streamlines in the CZ. Red/solid and blue/dashed curves correspond to MC1 and MC2 profiles, respectively.

FIG. 6. Results obtained from a dynamo run with MC2 profile for the meridional flow (as shown by the blue/dashed line in Figure 5).

FIG. 7. Same as Figure 3 of the Letter but here a delay has been included in the flux eruption process.

[4] D. H. Hathaway, Sol Phys 273, 221 (2011) [arXiv:1108.1722 [astro-ph.SR]].
[5] K. Petrovay, M. Nagy, and A. R. Yeates, Journal of Space Weather and Space Climate 10, 50 (2020) [arXiv:2009.02299 [astro-ph.SR]].
[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) [arXiv:2007.07069 [astro-ph.SR]].
[8] D. H. Hathaway, R. M. Wilson, and E. J. Reichmann, Sol Phys 151, 177 (1994).
[9] P. Charbonneau, Living Reviews in Solar Physics 17, 4 (2020).
[10] B. B. Karak and A. R. Choudhuri, MNRAS **410**, 1503 (2011), arXiv:1008.0824 [astro-ph.SR].

[11] S. Mandal, B. B. Karak, and D. Banerjee, Astrophys. J. **851**, 70 (2017), arXiv:1711.00222 [astro-ph.SR].

[12] M. Waldmeier, Ergebnisse und Probleme der Sonnenforschung (Leipzig: Geest & Portig) (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), astro-ph/0405027.

[15] B. B. Karak, Astrophys. J. **724**, 1021 (2010), arXiv:1009.2479 [astro-ph.SR].

[16] B. B. Karak, S. Mandal, and D. Banerjee, Astrophys. J. **866**, 17 (2018), arXiv:1808.03922 [astro-ph.SR].

[17] B. B. Karak and R. Cameron, Astrophys. J. **832**, 94 (2016), arXiv:1605.06224 [astro-ph.SR].

[18] See Supplemental Material for Model details and some supporting results, (2022).

[19] D. H. Hathaway, Living Reviews in Solar Physics **12**, 4 (2015), arXiv:1502.07020 [astro-ph.SR].

[20] A. V. Mordvinov, B. B. Karak, D. Banerjee, E. M. Golubeva, A. I. Khlystova, A. V. Zhukova, and P. Kumar, MNRAS **510**, 1331 (2022), arXiv:2111.15585 [astro-ph.SR].

[21] R. H. Cameron and M. Schüssler, Astrophys. J. **636**, A7 (2020), arXiv:2002.05436 [astro-ph.SR].

[22] R. Howe, Living Reviews in Solar Physics **6**, 1 (2009), arXiv:0902.2406 [astro-ph.SR].

[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), arXiv:1005.2283 [astro-ph.SR].

[25] Y. Fournier, R. Arlt, and D. Elstner, Astrophys. J. **620**, A135 (2018), arXiv:1808.08135 [astro-ph.SR].