

The Solar Dynamo and Its Phase Transitions during the Last Millennium

S. Duhau · C. de Jager

Received: 4 April 2007 / Accepted: 22 May 2008 / Published online: 14 June 2008
© Springer Science+Business Media B.V. 2008

Abstract We analyze the variation of the solar-dynamo magnetic-field components during the last millennium through a study of their proxy data. We introduce a phase diagram with as abscissa and ordinate the proxies of the values of the toroidal and poloidal magnetic field components. In this diagram the dynamo system appears to regularly cross a well-defined point, which we call the Transition Point. Such crossings occurred five times during the past millennium. Each of these crossings preceded a Grand Episode, either a Minimum or a Maximum one. In addition to these two types of quasiperiodic behavior, a third type consisting of weaker quasiregular oscillations (R) around the Transition Point's coordinates is identified. These periods appear to last one or two times the Gleissberg cycle length. Between the various types of episodes there are brief phase transitions. We identify two types of such phase transitions.

1. Introduction

The solar dynamo is a nonlinear system with chaotic elements (Weiss, 1987; Feynman and Gabriel, 1990; Ostryakov and Usoskin, 1990; Kremliovsky, 1995; Usoskin and Mursula, 2003; Duhau, 2003; Weiss and Tobias, 2004; de Jager, 2005). The problem that we investigate in this paper is whether it is possible to describe the observed variation of the dynamo system with time on the basis of its nonlinear character, in spite of its chaotic aspects.

The solar dynamo is seated in the strong magnetic fields that are generated in the tachocline, at the bottom of the convection region. Magnetic loops, detached from it or formed elsewhere in the convection zone, may rise and appear at the surface. The dynamo system is a nonlinear interplay between the poloidal and toroidal field components of the

S. Duhau (✉)

Departamento de Física, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires,
1428 Buenos Aires, Argentina
e-mail: duhau@df.uba.ar

C. de Jager

Royal Netherlands Institute for Sea Research, P.O. Box 59, 1790 AB Texel, The Netherlands
e-mail: cdej@kpnplanet.nl

dynamo (see, *e.g.*, Knobloch, Tobias, and Weiss, 1998; Durney, 2000; Dikpati, de Toma, and Gilman, 2004, 2006).

In the present paper, the evolution of these fields is inferred from the sunspot-cycle maximum values (1610–2005) and from the minimum values in the geomagnetic index *aa* around the times of the sunspot cycle minima (1844–2005). These two series of data are proxies for the long-term variations of the dynamo's toroidal and poloidal magnetic components, respectively. These series are extended over the past millennium on the basis of an analysis by Nagovitsyn (1997, 2006, 2007). A helpful way of analyzing the evolution of the solar dynamo system is a two-dimensional phase diagram that shows the relation between (proxies of) the poloidal and toroidal components of the dynamo field. By studying the time series of these proxies while subjecting them to a formalism based on the Morlet wavelet representation, we split them in two parts, on decadal and long-term time scales. Their respective temporal evolutions appear to be different. We describe the nonlinear temporal dependence of the relationship between the poloidal and the toroidal solar field components for each of these two components. That allows us to describe long-term solar-activity evolution, thus contrasting the results based on other nonlinear methods that state that the horizon of predictability is only a few years (*e.g.*, Ostryakov and Usoskin, 1990; Sello, 2001, and references therein). Our results agree with those of Ogurtsov (2004), also based on a nonlinear method.

2. The Toroidal and Poloidal Magnetic Field Components of the Dynamo and Their Proxies

Direct measurements of the solar magnetic field components are available for the past three and a half cycles only (Wang, Lean, and Sheeley, 2000; Wang and Sheeley, 2002). Hence, in view of the strong and irregular variability of the Sun and the consequent need to cover a longer interval of time, we need proxy data to understand the evolution of the two components of the dynamo field.

2.1. Toroidal Field Component

The strength of the toroidal component of the solar magnetic field is proportional to the total sunspot area (*S*; Nagovitsyn, 2005). Therefore we may consider the maximum *S* values as measures for the amplitudes of the Schwabe cycle in the toroidal component of the solar dynamo magnetic field. A comparison of the annual mean values of *S* with those of the Zürich Sunspot Number *R* shows that they are equivalent after 1780 and during the Maunder Minimum. By scaling the *R* values to the *S* scale we derived the empirical relation $S [M_{\text{sh}}] = 15.7R$, where M_{sh} is the spot area expressed in millionths of the solar hemisphere. Only in the period 1704 to 1780 are some *R* values, thus scaled, slightly larger compared with *S*. This overall similarity allows us to use *R* values as proxy data for the toroidal field component. An extension of these proxies to earlier periods is based on the Schöve (1955) time series of sunspot maxima as reconstructed from sunspot data and auroral occurrences. The time series extends, with interruptions, from 648 BC to present but only after AD 290 is it continuous. Based on the nonlinear relationship between the Schöve time series and the simultaneous Zürich Sunspot Number time series, and using a wavelet analysis, Nagovitsyn (1997, 2006, 2007) extended the Zürich Sunspot Numbers to the past. In this paper we will, if needed for clarity, denote Nagovitsyn's reconstructed series of *R* values by *RN*; when we are dealing with the maximum *R* values of solar cycles we call it R_{max} .

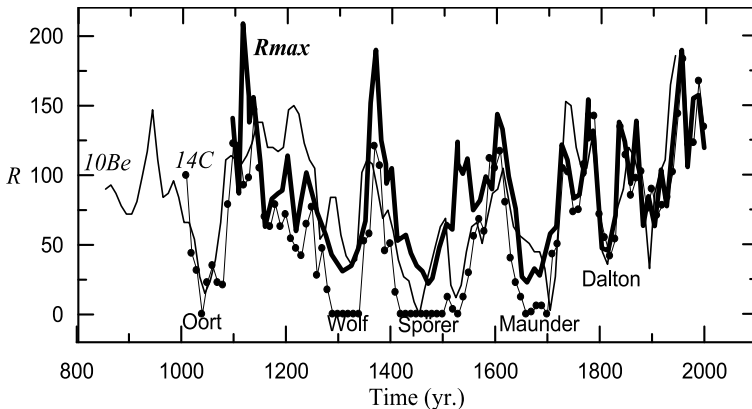


Figure 1 Sunspot number time series as inferred from ^{14}C by Stuiver and Quay (1980) (points) and from ^{10}Be by Usoskin *et al.* (2004) (thin line), respectively, and sunspot cycle maxima in the Nagovitsyn (1997, 2007) from <http://www.gao.spb.ru/database/esai/> annual sunspot number time series (thick line). Five episodes known as Grand Minima occurred during the last millennium. These are the Oort, Wolf, Spörer, Maunder, and Dalton Minima, respectively. There are also several Grand Maxima; the most pronounced are those peaking in 1110, 1375, and 1960.

Other proxies for past solar activity are sunspot numbers that have been inferred from ^{14}C and ^{10}Be cosmogenic isotopes. The physics of the production of these isotopes is described by many authors. We just mention Beer (2000) and a review by Scherer *et al.* (2006). The transformations of deposit rates into time series of sunspot numbers was done from ^{10}Be data by Usoskin *et al.* (2004) and from ^{14}C by Stuiver and Quay (1980); *cf.* Figure 1). The three time series – those for the ^{14}C , the ^{10}Be , and the R_{max} data – appear to coincide as regards the periods of occurrence of *Grand Minima*. We also note periods of high activity; we will call these *Grand Maxima*.

A matter that deserves attention is that of the reality of the heights of the medieval Grand Maxima. The analysis by Usoskin, Mursula, and Kovaltsov (2003) shows lesser heights of the medieval maxima. In addition, Weber, Crowley, and Van der Schier (2004), Dergachev *et al.* (2004), and Solanki *et al.* (2004), from a reconstruction of the sunspot numbers during the largest part of the Holocene, on the basis of dendrochronologically dated radiocarbon concentrations, have concluded that the level of solar activity during the second half of the 20th century was exceptionally high and that the previous period of equally high activity occurred many thousand years ago. That conclusion appears to conflict with the data of our Figure 1. Also, Muscheler *et al.* (2005) and Steinhilber, Abreu, and Beer (2008) remarked that this time interval is most probably considerably shorter.

We have analyzed the assumptions that have led to the proxy data of Figure 1. The auroal data depend strongly on the relative geographic position of the auroral oval with respect to the observational location. It varies with time owing to changes in the geomagnetic–geographic relative position (Roederer, 1974). The time dependence of the corrections of cosmogenic data for the effect of changes in geomagnetic intensity is a fairly smooth function over a millennium scale. Hence, a simple correction procedure would be adequate if the data were of global coverage, but all of the time series of cosmogenic abundances have been taken at fixed locations. This implies that the westward drift of the geomagnetic field in the course of centuries may have caused changes in the local magnetic field intensity by the shifting of longitudinal inhomogeneities in the field structure over times of the order of a century. These changes are in some locations as large as those from the millennial trend in

the dipolar intensity in the same period of time, as may be seen for example in the archeomagnetic data from western Europe as presented by Gallet, Genevey, and Fluteau (2005). Hence, the significant differences found in the reconstructed sunspot time series taken in different locations, such as those shown in Figure 1, may be because the geomagnetic field variations were different in each of these locations. Robust conclusions for periods longer than a few hundred years cannot be made (Snowball and Muscheler, 2007; Ogurtsov, 2007). So we can neither discard nor confirm the possible existence of short periods of seemingly extreme RN activity in the 12th and 14th centuries.

2.2. The Poloidal Field Component

The minimum value of the geomagnetic activity (aa_{\min}) is a direct measure of the amplitude of the polar field during the Schwabe cycles (*cf.* Russell, 1975; Russell and Mulligan, 1995; Hathaway, Wilson, and Reichmann, 1999; Duhau and Chen, 2002; Duhau, 2003). Therefore, they will be used to study the variation of the polar components of the magnetic field. The longest time series of geomagnetic activity (1868–present) is the series of aa geomagnetic index values deduced by Mayaud (1975). It was extended back to 1844 by Nevanlinna and Kataja (1993). Nagovitsyn (2006) has reconstructed the aa time series since 1090 from his RN time series, on the basis of a wavelet analysis of the nonlinear relationship between R and aa from 1868 to present. We will henceforth call it the Naa time series. However, as with the RN time series as a proxy of the solar toroidal field strength, the use of geomagnetic indexes as a proxy for the solar poloidal field component might have been confused if geomagnetic activity had been modulated by geomagnetic field intensity and morphology. We examine this problem.

The aa index (Mayaud, 1975) was found by averaging a northern index measured at Greenwich and a southern one at Melbourne. The main field changes did not modify the oval position at Greenwich and changed it by only 1° at Melbourne (Feymann and Ruzmaikin, 1999). Therefore, the minima in the Mayaud time series may be safely used as a proxy for the poloidal component of the dynamo field. Also, during the interval 1705 to present the RN data coincide with the directly measured R values. Nagovitsyn based his extension to the past on the nonlinear relationship between the aa index and the R time series in the time interval 1868–1998. During that period the aa index was not markedly affected by geomagnetic field changes. Hence, the aa_{\min} time series in the Nagovitsyn data may well be used as a measure for the amplitude of the poloidal component of the solar magnetic field. So we use the Naa data for the interval 1090–1897, and the Mayaud index from 1868 to present (*cf.* Figure 2). Similarly as with Figure 1, the poloidal component shows the Grand Maxima and Minima, although there are remarkable differences in their amplitudes.

Summarizing this section, we have described the proxies of the time evolution of the two components of the dynamo field. For the R_{\max} data we checked their reliability by comparing several sources for the proxies. Such checking is not possible for the aa_{\min} time series, for which only Nagovitsyn's (2006) synthetic time series is available before 1844. Both the poloidal and toroidal components of the dynamo show striking periods of extreme amplitudes: the Grand Maxima and Minima.

3. Decomposition of the Proxy Data

Solar activity shows a number of components: The Schwabe, Hale, Gleissberg, and Suess (de Vries) cycles, with average periods of $\approx 11, 22, 88,$ and 205 years, respectively, are the most

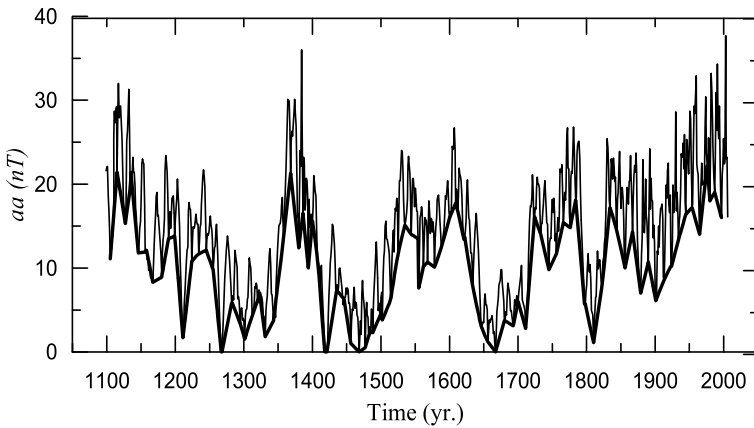


Figure 2 Geomagnetic index aa as defined by Mayaud (1975) (thin line). Data from 1868–2006 are derived from <http://www.ngdc.gov>, those prior to 1868 are from RN , and the time series is reconstructed by Nagovitsyn (2006) (from <http://www.gao.spb.ru/database/esai/>). The thick line shows the minimum aa values. It is assumed to be representative of the strength of the polar magnetic field.

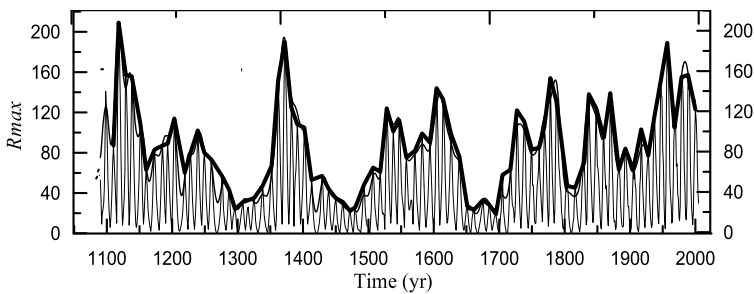


Figure 3 The Nagovitsyn (2007) time series together with the maximum sunspot-number envelope computed by a previously used method (Duhau, 2003, light line) and by the method introduced here (see text).

prominent ones (de Jager, 2005). Duhau and Chen (2002), who decomposed the proxies of the two magnetic components on the basis of Morlet wave base functions, found three classes of their components with periodicities that are harmonics or subharmonics of the basic components: the so-called Schwabe, decadal, and long-period (secular) components. The division between the decadal and long-term components is located between ≈ 65 and ≈ 85 years.

3.1. Filtering the Time Series

To study the modulation of the sunspot cycle amplitudes, wavelet components with Fourier periods less than 15.1 years were subtracted from the yearly means time. The envelope of the maxima was obtained by multiplying the result by two (Duhau and Chen, 2002; Duhau, 2003; thin line in Figure 3). Since we base our research directly on the R_{\max} and aa_{\min} time series the Schwabe signal *per se* does not play a role in our analysis.

3.2. Decomposing The Time Series

Duhau and Chen (2002) studied the sunspot cycle maxima envelope since 1844 on the two aforementioned time scales and found that it can be decomposed into a decadal oscillation that contains the well known odd–even effect and a long-term variation that can be decomposed into its value at 1923 and a fairly harmonic function of which the average period is in the Gleissberg band of values. It was called the Gleissberg cycle (Duhau, 2003). A similar but more precise procedure is followed here. Splitting the time series into the long-term variations and the decadal oscillation proceeds in three steps:

- (1) *Determination of the years and amplitudes* of the extremes in R_{\max} and aa_{\min} for each 11-year sunspot cycle.
- (2) *Interpolation*. The sunspot-cycle length varies between wide limits, and so the time series of the maximum or minimum years is not evenly distributed. Since the wavelet analysis requires a constant sampling time, the need for linear interpolation arose. To that end we inserted annual data points, obtained by linearly interpolating between successive maxima or minima. The thick line in Figure 3 represents this envelope of maxima. By this procedure we reduced the point-to-point differences from 5% to 0.05% in the representation of R_{\max} and aa_{\min} .
- (3) *Wavelet decomposition*. With regard to the technique for representing time series related to solar dynamo variables, we cannot use the Fourier base functions because they apply to stationary processes only. As the dynamo system is not stationary, a base function of compact support is needed. Therefore our analysis is based on the wavelet base functions (Torrence and Compo, 1998), a procedure that was applied earlier to similar problems (see, e.g., Duhau and Chen, 2002; Duhau, 2003; Nagovitsyn, 2006). It was also used for reconstructing sunspot numbers from the Schöve time series (Nagovitsyn, 2007).

As compared to previous investigations, we introduced a larger number of wavelet components and a larger value for the nondimensional characteristic frequency than was done earlier (Duhau and Chen, 2002; Duhau, 2003). This procedure leads to a broader separation between peaks in the spectrum, thereby improving the separation between frequency bands. Specifically, in the actual wavelet decomposition for the last millennium we have used the 127 wavelet components that result from including 12 subyears. To perform the computation of the Morlet wavelet components, a software package was used after fixing the characteristic nondimensional frequency at 12 (*cf.* Torrence and Compo, 1998; <http://paos.colorado.edu/research/wavelets/>).

For actual practice, the decadal oscillation is defined as the superposition of the wavelet components whose Fourier periods – 3.42 to 71.95 years – are in the higher frequency band of the spectrum. The long-term variation is defined as the linear trend added to the long-term oscillation, this last being the superposition of all the wavelet components in the lower frequency band – Fourier periods 76.22 to 2584.2 years. This last includes periods in the Gleissberg and Suess bands and beyond. Because of edge effects (Torrence and Compo, 1998) the decomposition is accurate only over a shorter time interval than that of the original time series.

To judge the accuracy of the proxy data in the two components, we compare the R_{\max} data with the two time series found from cosmogenic radionuclide data (Figures 4(a) and (b)). Since the present investigation aims at comparing the data on the decadal time scale with those in the range of the Gleissberg and the Suess periods, we have excluded periods longer than 440 years. By examining the contribution of the disregarded components to the total signal in R_{\max} based on the sunspot number time series since 1705, we found that they contribute at most 5% to the total signal.

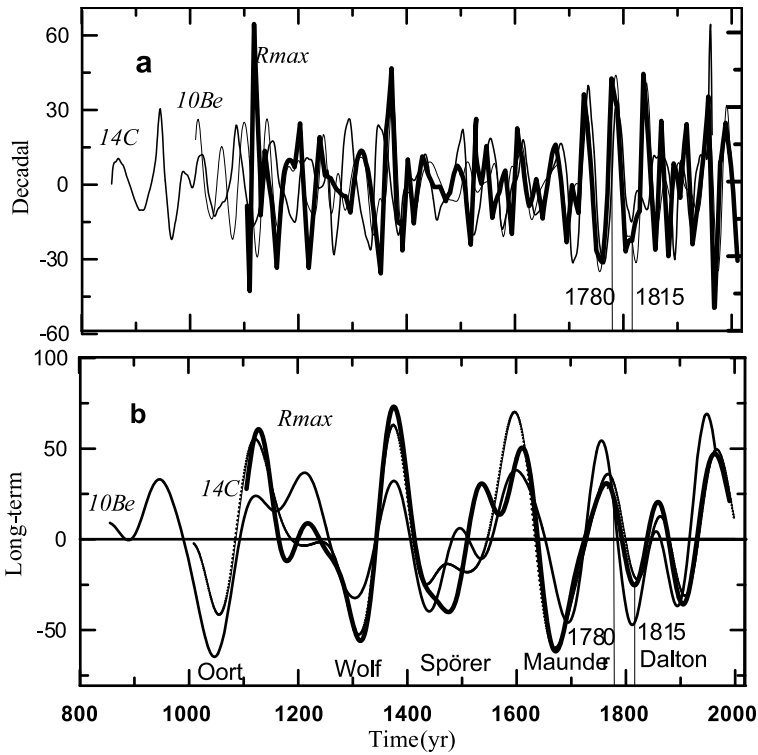


Figure 4 The (a) decadal and (b) long-term oscillations in the sunspot-number proxy. These decompositions are shown in the R_{max} time series and in those of the ^{10}Be and ^{14}C data as presented in Figure 1. The vertical lines labeled 1780 and 1815 are discussed in the text.

The decadal signal in the three kinds of proxies is reasonably similar insofar as the amplitudes of the oscillations are concerned. These amplitudes are quite weak during most of the Grand Minima, particularly during the Spörer and Maunder Minima. An exception is the Dalton Minimum. The amplitudes are stronger during Grand Maxima. However, for the three proxies they differ in the times of maxima, minima, and zero crossing. We note that the time delay between solar activity and the ^{14}C signal plays a part in the decadal signal.

However, there is reasonable agreement between the long-term oscillations in ^{14}C and R_{max} . The exception is the short interval 1420–1570. But even then they still coincide in the positions of highs and lows. The overall agreement is evidence for the reality of the Grand Maxima in the 12th and 14th centuries as observed in the R_{max} time series (Figure 1). With regard to the Dalton Minimum this fairly brief episode does not show itself significantly in the long-term oscillations. This is clear from the long-term oscillation in the historical R_{max} time series (heavy line in Figure 4(b)). Instead, this “short” Grand Minimum is due to the synchronicity of maximum negative amplitude in the Gleissberg cycle and the decadal oscillations, this last being very strong. Compare, for example, the vertical lines labeled 1815 in Figures 4(a) and (b).

During the Dalton Minimum the R_{max} values were 50 spot number units smaller than the average level of 94 (see Figure 1). It may be significant that during the previous positive phase (the vertical lines labeled 1780), when this strong oscillation was synchronous with

positive values of the long-term oscillation, an analogous situation occurred, leading to a departure of the same magnitude (49 spot number units) above the average.

Our analysis confirms that the decadal signal is weak, sometimes even absent during Grand Episodes, the Maunder Minimum being a well-studied case. In contrast, the Grand Episodes are well visible in the long-term decomposition with the Dalton Minimum being a remarkable exception. This latter case may be an interesting topic for later investigation.

4. A Well-Observed Period: 1620 to 2000

For illustrative purposes and for providing the basis for the topics that are dealt with in the next section, we describe here the well-observed period from the Maunder Minimum through the recent Grand Maximum.

Securing instrumental data of the geomagnetic index aa time series started in 1844. The relationship between aa_{\min} and R_{\max} for the period 1844–2000 was analyzed by Duhau and Chen (2002), who found that in 1923 a sudden change occurred in the decadal oscillation and in the Gleissberg cycle, the latter being defined as the long-term variation from which its value in 1923 was subtracted. Note that the same result would not have been found if instead of subtracting this constant value from the long-term variation we had subtracted the linear trend. However, this trend does approach a constant value that is very close to the 1923 value when the length of the time series is increased. It differs by less than 5% from the value of the long-term variation at 1923 when that would have been computed from the 1800 year time series (1955). The decadal oscillations and the long-term variation during the period 1630–2006 were obtained as described in Section 3, and by subtracting their values in 1923 from the long-term variation, the Gleissberg cycle was computed. The results are shown in Figure 5. Characteristic elements are the Maunder Minimum of which the lowest R_{\max} values occurred in 1650–1700, the recent Grand Maximum, peaking around 1960, and during a ≈ 200 -year period of quasiregular oscillations between these two. We note that the durations of the Grand Minimum and Maximum are of the order of the Gleissberg cycle as computed.

A crucial year is 1923. The amplitude and length of the oscillations changed significantly after 1923 (*cf.* Figure 5). Another aspect of that year was that the decadal oscillations in the

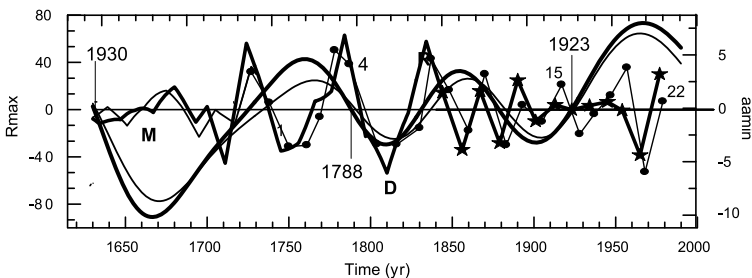


Figure 5 The decadal oscillation and the Gleissberg cycle in aa_{\min} (heavy lines) and R_{\max} (light lines), respectively. The points and stars at the peaks of the decadal oscillation indicate the times of sunspot maxima and geomagnetic minima since 1705 and 1844, respectively. The Gleissberg cycles were computed by subtracting from the long-term variations in aa_{\min} and R_{\max} their values in 1923: 10.28 nT and 93.44 respectively. The three thin vertical lines indicate the dates during which the Gleissberg cycles (see text) in both components were simultaneously zero. The two even numbers indicate the leading even cycle of the two pairs of sunspot cycles that violated the odd–even rule since 1705. Sunspot cycle maximum 15 occurred at the beginning of the polar cycle peaking at 1923.

aa_{\min} values were close to zero during a period that lasted from prior to the maximum of cycle 15 until after that of 16 (Figure 5). Hence, the amplitude of the 11-year polar cycle was at that time given by the constant level in Figure 6(a) (*i.e.*, $aa_{\min} = 10.28$). At that time an abrupt change in the amplitude of the Gleissberg cycle occurred. Also, the decadal oscillations in the annual means of R_{\max} were zeroing simultaneously in 1923. The values of the decadal variation at 1923 differ from zero by less than 1%. After that year both R_{\max} and aa_{\min} increased in length and amplitude (see the lower curves in Figure 5). These sudden changes in the oscillatory behavior that occurred simultaneously in the two coordinates may be interpreted as a manifestation of the nonlinear nature of the solar dynamo. We also note that before sunspot cycle maximum 15 the maxima of the decadal cycle occurred at each even sunspot cycle, and the minima at the odd ones. Therefore the period of the decadal cycle was about 22 years, which is the regular Hale cycle length. After sunspot cycle 16 the length of the decadal cycle increased sharply, occurring just during the polar cycle that was active between cycles 15 and 16 and peaking during the 1923 geomagnetic minimum. Therefore we may conclude that the observed change of behavior around 1923 was very rapid and that the period of change did not last longer than the polar Schwabe cycle that peaked in 1923.

The temporal evolution of the magnetic-field components was similarly analyzed for the shorter interval 1844–2000 (*cf.* Figure 2 in Duhau, 2003), but the resulting values for aa_{\min} and R_{\max} in 1923 differ from the earlier ones, owing to the different methodology applied here for finding the maxima (minima) in the time-series modulation and for splitting it into the decadal oscillation and the long-term variation.

The Gleissberg cycles in both components are simultaneously close to the origin of the coordinates in 1630, 1788, and 1923, with small numerical differences that are not discernible in Figure 5. However, there are noticeable differences in the decadal oscillations in 1788 on one hand and in 1630 and 1923 on the other hand, whereas the decadal oscillations were very small or passing through the origin in the last two cases; this was not so in the first. On the other hand, while in the 1930 and 1923 cases a strong change in the quasiperiodic behavior occurred simultaneously in the decadal oscillation as well as in the Gleissberg cycle, the same did not occur in the 1788 case. This fact indicates that the condition of a Grand Episode, either a Grand Minimum or a Grand Maximum, to occur is that long-term variations in aa_{\min} and R_{\max} assume simultaneously values that are very close to those in 1923 while simultaneously the decadal oscillation is very small or near zero. Therefore we will call the pair of values (aa_{\min} , $R_{\max} \approx 10.26, 93.44$) a “Transition Point” (TP). The precise values of these data are determined in the next section.

Also, in Figure 5 we see that at the end of the Maunder Minimum two complete Gleissberg cycles of moderate amplitude and period started at different dates: 1721 and 1730 for the polar and the toroidal components, respectively. Therefore a change of behavior occurred that is quite different from that just discussed. This type of event will be further studied in the next section.

To conclude this section, the well-observed Grand Minimum and Grand Maximum of the 17th and 20th centuries are pronounced in the long-term oscillations, with a period of more regular oscillations in between. The year 1923 marks a change, when the regular oscillations were succeeded by the Grand Maximum of the 20th century. It may be significant that around that year the decadal oscillations in aa_{\min} and R_{\max} were close to zero.

Table 1 Lengths of the various episodes in years.

Type of episode	M	H	R
-2	119	≈ 80	59
-1	176	56	105
0	91	≈ 80	204

5. The Solar Dynamo during the Past Millennium

Guided by the observations presented in the previous section, we analyze the whole past millennium. A helpful way to study the time changes of the toroidal and poloidal components of the dynamo magnetic field is by plotting their proxies in an R_{\max} versus aa_{\min} diagram (Duhau and Chen, 2002; Figure 3 in Duhau, 2003; cf. our Figure 6). This diagram projects proxies of the two magnetic field components of the dynamo system along the abscissa and ordinate and it has therefore the properties of a phase diagram. It may also be viewed as a representation of two quasiharmonic motions with a phase difference.

5.1. Three Kinds of Episodes

The dominant features in this diagram are the two kinds of episodes: the large upward loops (H) during which the dynamo passed through a Grand Maximum and the downward loops (M) corresponding to Grand Minimum. Here, M-2, M-1, and M indicate the Wolf, Spörer, and Maunder Grand Minima and H-2, H-1, and H designate the Grand Maxima in the 12th, 14th, and 20th centuries, respectively. Apart from these episodes there is a third kind of episode, consisting of loops with smaller amplitudes. They appear to occur when the track happens to just miss the TP during its return to this point after a Grand Episode. The best observed case is the oscillation between 1721 and 1923 (cf. Sections 3 and 4 and Figures 5 and 6(d)), but similar episodes started in 1171 and 1525 (cf. Figures 6(a) and (c)). A characteristic of such episodes is the fact that the path performs a closing system of quasi-ellipsoidal loops to finally hit the TP in the various diagrams. We will designate these episodes as R (Regular) type. Those starting in 1175, 1525, and 1721 will be identified by R-2 through R, respectively.

The length of the three kinds of episodes cannot be determined very precisely, apart from those cases in which the relevant episode starts and ends in the TP, such as the M-2 and H-1 episodes. In computing an approximate value for the other cases we have chosen those data in which the path crosses the TP abscissa (cf. Figure 6). The values are given in Table 1. Most of them are in the range of the Gleissberg period. There are two cases where they rather equal twice the Gleissberg period. None of them are of the length of the Suess (de Vries) cycle, because the length of that cycle, as determined by several authors, is single peaked and it varies only between narrow limits. The Gleissberg cycle is doubly peaked, variable, and broad banded (see the review by de Jager, 2005, p. 233).

5.2. The Transition Point in the Phase Diagram

In the previous section we introduced the term “Transition Point” for the pair of values of the long-term variation that occur at the onset of a Grand Episode. Its numerical values are close to those of the long-term variation in 1923. We have found two more such events in which a “Grand Maximum” (in 1923) or a “Grand Minimum” (in 1630) started. Similarly, other Grand Episodes, either a Maximum or a Minimum, started in three cases during the

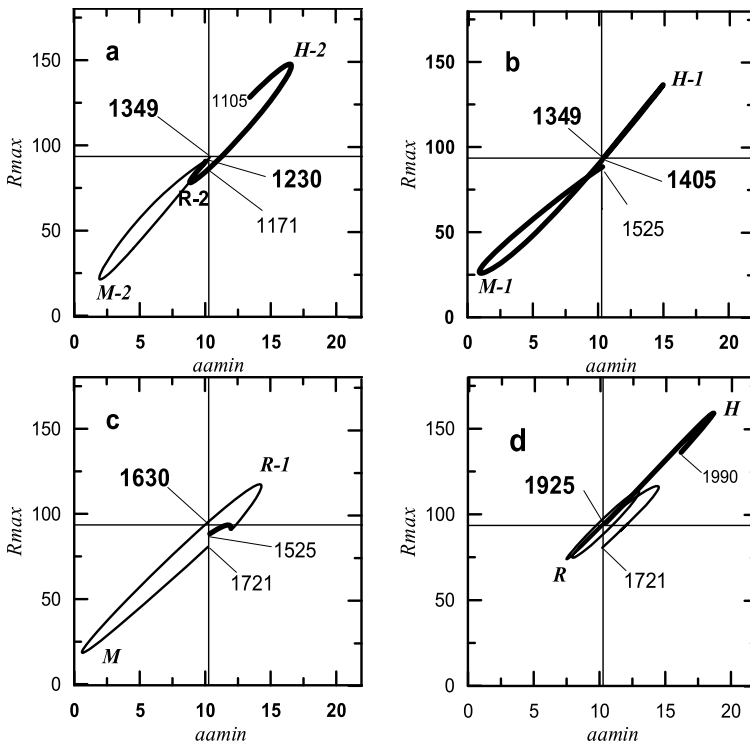


Figure 6 Four phase (R_{\max} , aa_{\min}) diagrams covering the past millennium. They show the long-term variation in sunspot cycle maxima (R_{\max} , characteristic for the toroidal component of the solar dynamo) versus those in aa_{\min} (poloidal component) for the four intervals (a) 1105 – 1349, (b) 1349 – 1525, (c) 1525 – 1730, and (d) 1721 – 1990. The horizontal and vertical lines correspond to the abscissa and the ordinate of the transition point (TP) (10.34, 93.38) (see Table 2). The years printed in bold are those in which the path collapsed into the TP; others correspond to the nearest approach of the path to the TP abscissa. The sense of motion along the paths is shown by thin lines (counterclockwise) and thick lines (clockwise).

last millennium. This happened when the coordinates of this point were equal to the 1923 values to within less than 1% in the phase diagram of the long-term variation.

These cases along with their coordinates in the phase diagram are listed in Table 2. The bottom rows give their average values, next to their one- σ mean errors, and the last column presents the character of the episodes preceding and following the transition through the TP. Of these the 1923 event is best studied. It is not only the point into which the path of the R episodes eventually returns but it also marks a phase transition (Duhau and Chen, 2002) during which the path changes suddenly in relative amplitude, while around that year the decadal oscillations in both coordinates practically reduce to zero. These observations lead us to improving the precision of the definition of the TP, as the point in the phase diagram, defined by two conditions:

- i) The long-term variations in aa_{\min} and R_{\max} take those of the TP coordinates to within a statistical error smaller than 1% (Table 2).
- ii) The amplitude of the decadal variation in aa_{\min} is close to zero (to within less than 1% of the TP’s value), whereas that in R_{\max} passes through zero. Hence the odd – even rule is fulfilled for these transitions (cf. the lower curves in Figure 5).

Table 2 Observed coordinates of the transition point, average values, and the one- σ mean errors.

Year	aa_{\min}	R_{\max}	Episodes
1230	10.01	90.61	R-2 to M-2
1349	10.37	93.23	M-2 to H-1
1405	10.54	94.52	H-1 to M-1
1630	10.22	95.12	R-1 to M
1923	10.26	93.44	R to H
Average	10.34	93.38	
Mean error (σ)	0.08	0.69	

Condition *i*) appears to be valid for all five cases that occurred during the last millennium and condition *ii*) is only well documented for the 1923 case. However there is a well-documented counterexample (the case of 1788) for which condition *i*) was fulfilled while condition *ii*) was not. In that case a phase transition did not occur. This gives some support to the working hypothesis that the necessary and sufficient conditions for a transition to a Grand Episode to occur are the two conditions discussed here.

5.3. Transition Periods between the Various Episodes

Returning to the 1923 transition between the R-type and the H-episode, we remark that around that year a sudden change occurred in the parameters of the long-term component of the Gleissberg cycle while the decadal oscillation practically reduced to zero. Table 2 shows that such transitions took place five times during the past millennium. We designate these transitions as G (Grand)-type, because they are followed by Grand Episodes. Another type of transition happens when the path, while returning toward the TP, slightly misses it while an R episode starts thereafter. Table 3 lists three such transitions together with the differences between the point of closest approach (expressed in one- σ values of the transition point's coordinates) and the types of episodes concerned. We call these C-type transitions.

The difference between the two types of transitions is that a G-type transition (Table 2) is exclusively followed by a Grand Episode, whereas a C-type one is always followed by an R-type episode. In the phase diagram the necessary and sufficient condition for a G-type transition to occur is that the path is precisely drawn towards the transition point, with a deviation that is not larger than two- σ values. A larger difference on the way toward the transition point signals a C-type transition, preceding an R-episode (Table 3). The duration of these transitions is not yet certain. From Table 3 (last column) one may infer that it is of the order of a Schwabe cycle but precise information cannot be drawn from these three cases. We remark that the dynamo system is presently undergoing a transition from the past Grand Maximum (H) toward another episode. The ongoing observations may provide empirical data on the present transition.

5.4. The Transition State

A G transition occurs always after the years during which the long-term variation crossed the TP (Condition *i*)). There are indications that simultaneously condition *ii*) is also fulfilled: In that case the amplitude of the decadal variation in aa_{\min} is close to zero (to within less than 1% of the TP's value), whereas that in R_{\max} passes through zero. In these cases the odd–even rule is fulfilled. This means that the transition occurs in that polar cycle of which

Table 3 Duration of C-type episodes. The table gives the years of closest approach to the transition point, either for the abscissa or for the ordinate. The difference between the coordinates of closest approach and those of the transition point are expressed in one- σ values. The duration of the transition periods is found as the time between the year of closest approach and that of the start of the corresponding R-type episode. Three such cases are described.

Year	aa_{\min}	$\Delta aa_{\min}/\sigma$	R_{\max}	$\Delta R_{\max}/\sigma$	Episodes	Time delay
1165	11.28	1.25	93.75	0.54		
1171	10.27	0.16	88.26	7.10	H-2 to R-2	6
1525	10.33	0.13	88.32	7.19		
1540	11.75	1.76	93.34	0.07	M-1 to R-1	15
1721	10.21	0.29	80.56	4.06		
1730	11.98	2.05	93.35	0.04	M to R	9

the amplitude is equal to the TP's abscissa while the average amplitude of the preceding and following toroidal cycles is very close to the TP's ordinate. These properties of the polar and toroidal cycles around the transition date indicate that the corresponding dynamo state is a well-determined one. A striking feature in Figure 5 is the fact that in the coordinate system with its origin in the TP the path is asymmetrical: The loops in the first quadrant are smaller than those in the third, and the path passes from the first to the third quadrant mostly to the right of the TP coordinates (the fourth quadrant); this feature indicates that the solar dynamo has an intrinsic asymmetry.

Another way to look at these episodes and the transitions between them can be derived from the phase diagram. The amplitude and hysteresis of the quasi-ellipses may suddenly change when passing by or slightly missing the TP, and the final state of the amplitude and hysteresis of the newly formed loop do not depend of the actual value of the coordinates of the path at that instant. This fact shows how sensitive the future path in the phase space of the dynamo system is to the initial conditions. It again illustrates the typical characteristics of a nonlinear system with its "chaotic" transitions (Duhau, 2003). As the phase diagrams represent the long-term evolution of the amplitude of the 11-year cycle in the toroidal (ordinate) and poloidal (abscissa) components, the transition point is a limit circle.

In concluding this section, we give evidence that the solar dynamo system is characterized by one transition state. Its coordinates (expressed in the proxies for the two magnetic field components) are determined with good precision. The system frequently makes episodic excursions to deviating conditions. Three kinds of such episodes have been identified. Their durations are one or two times the Gleissberg cycle length. There are brief transition phases between such episodes. Of these we have identified two types. There are weak indications that these transition phases may last for a time of the order of the Schwabe cycle.

6. Conclusions

This investigation contains an observational analysis of the nonlinear character of the solar dynamo and its chaotic elements that characterize solar activity (Kremliovsky, 1995; Weiss and Tobias, 2004). This is done on the basis of observations of the proxies for the poloidal and toroidal magnetic field components over the past millennium. These proxies are the maximum sunspot number of the successive solar cycles (R_{\max}) and the minimum aa values (aa_{\min}), respectively. For the latter time series, the more recent data are more reliable than the

earlier ones. By means of a Morlet wavelet representation these two proxies are decomposed into decadal and long-term components.

The decadal component consists of the Hale cycle and its harmonics and subharmonics, whereas the long-term one is based on the Gleissberg and Suess (de Vries) cycles and their harmonics and subharmonics. Both the long-term poloidal and toroidal components of the dynamo show striking periods of extreme amplitudes: the Grand Maxima and Minima. The decadal signal is weak and sometimes even absent during Grand Minima, the Maunder Minimum being a well-studied case.

We illustrate the diagnostic value of a phase diagram (aa_{\min} , R_{\max}). Its study shows that the dynamo is characterized by one well-defined Transition Point, which it regularly crosses (five times during the past millennium). In addition there are three quasiperiodic states: the M-type, the H-type, and the R-type oscillating behavior. Their durations are one or two times the Gleissberg cycle length. A well-studied year is 1923 in which a transition occurred between two centuries of regular oscillations and the Grand Maximum of the 20th century. Around that year the decadal oscillations in aa_{\min} and R_{\max} were close to zero. That year preceded the Grand Maximum of the 20th century.

There are brief transition phases between such episodes. Of these we have identified two types. They depend on the distance of the path in the phase diagram to the TP at the moment of maximum approach, when returning to it: A Grand Episode occurs always (G transition) when the distance of the path to the TP is less than 1%; otherwise the transition leads to a regular episode (C transition) along which the path finally returns to the TP. Although in the last case there are weak indications that these transitory phases may last the Schwabe cycle length, in the first case the transition occurs exactly during the polar Schwabe cycle that has an amplitude equal to the TP abscissa whereas the average amplitude of the preceding and following toroidal cycles amplitudes are close to that of the TP ordinate. This indicates that the TP point corresponds to a well-defined dynamo state.

Acknowledgements We are obliged to Ilya Usoskin and to an anonymous referee for many very helpful remarks. Their help is greatly appreciated.

References

- Beer, J.: 2000, *Space Sc. Rev.* **94**, 53.
 de Jager, C.: 2005, *Space Sci. Rev.* **120**, 197.
 Dergachev, V.A., Dmitriev, P.B., Raspopov, O.M., Van Geel, B.: 2004, *Russ. J. Earth Sci.* **6**, 323.
 Dikpati, M., de Toma, G., Gilman, P.A.: 2004, *Astrophys. J.* **601**, 1136.
 Dikpati, M., de Toma, G., Gilman, P.A.: 2006, *Geophys. Res. Lett.* **33**, L05102. doi:[10.1029/2005GL025221](https://doi.org/10.1029/2005GL025221).
 Duhau, S.: 2003, *Solar Phys.* **213**, 203.
 Duhau, S., Chen, C.: 2002, *Geophys. Res. Lett.* **29**, 1628. doi:[10.1029/2001GL013953](https://doi.org/10.1029/2001GL013953)
 Durney, B.: 2000, *Solar Phys.* **196**, 421.
 Feynman, J., Gabriel, S.B.: 1990, *Solar Phys.* **127**, 393.
 Feynman, J., Ruzmaikin, A.: 1999, *J. Geophys. Res. Lett.* **26**, 2057.
 Gallet, I., Genevey, A., Fluteau, F.: 2005, *Earth Planet. Sci. Lett.* **236**, 339.
 Hathaway, D.H., Wilson, R.M., Reichmann, E.J.: 1999, *J. Geophys. Res.* **104**, 22388.
 Knobloch, E., Tobias, S.M., Weiss, N.O.: 1998, *Mon. Not. Roy Astron. Soc.* **297**, 1123.
 Kremliovskiy, M.: 1995, *Solar Phys.* **159**, 371.
 Mayaud, P.N.: 1975, *J. Geophys. Res.* **80**, 111.
 Muscheler, R., Joos, F., Müller, S., Snowball, I.: 2005, *Nature* **436**, E3. doi:[10.1038/Nature04045](https://doi.org/10.1038/Nature04045).
 Nagovitsyn, Yu.A.: 1997, *Astron. Lett.* **23**, 742.
 Nagovitsyn, Yu.A.: 2005, *Astron. Lett.* **31**, 622.
 Nagovitsyn, Yu.A.: 2006, *Astron. Lett.* **32**, 344.
 Nagovitsyn, Yu.A.: 2007, *Astron. Lett.* **33**, 340.
 Nevanlinna, H., Kataja, E.: 1993, *Geophys. Res. Lett.* **20**, 2703.

- Ogurtsov, M.G.: 2004, In: Stepanov, A.V., Benevolenskaya, E.E., Kosovichev, A.G. (eds.) *Multi-Wavelength Investigations of Solar Activity, Proc. IAU Symposium 223*, Cambridge University Press, Cambridge, 135.
- Ogurtsov, M.G.: 2007, *Astron. Lett.* **33**, 419.
- Ostryakov, V.M., Usoskin, I.G.: 1990, *Solar Phys.* **127**, 405.
- Roederer, J.G.: 1974, *Geophys. Res. Lett.* **1**, 367.
- Russell, C.T.: 1975, *Solar Phys.* **42**, 259.
- Russell, C.T., Mulligan, T.: 1995, *Geophys. Res. Lett.* **22**, 328.
- Scherer, K., Fichtner, H., Borrman, T., Beer, J., Desorgher, L., Flükiger, E., Fhar, H.J., Ferreira, S.E.S., Langner, U.W., Potgieter, M.S., Heber, B., Masarik, J., Shaviv, N.J., Veizer, J.: 2006, *Space Sci. Rev.* **127**, 467.
- Schove, D.J.: 1955, *J. Geophys. Res.* **60**, 127.
- Sello, S.: 2001, *Astron. Astrophys.* **377**, 312.
- Snowball, I., Muscheler, R.: 2007, *Holocene* **6**, 851.
- Solanki, S.K., Usoskin, I.G., Kromer, B., Schüler, M., Beer, J.: 2004, *Nature* **341**, 1084.
- Steinhilber, F., Abreu, J.A., Beer, J.: 2008, *Astrophys. Space Sci. Trans.* **4**, 1.
- Stuiver, M., Quay, P.D.: 1980, *Science* **207**, 11.
- Torrence, C., Compo, G.P.: 1998, *Bull. Am. Meteorol. Soc.* **79**, 61.
- Usoskin, I.G., Mursula, K.: 2003, *Solar Phys.* **218**, 319.
- Usoskin, I.G., Mursula, K., Kovaltsov, G.A.: 2003, *Astron. Astrophys.* **403**, 743.
- Usoskin, I.G., Solanski, S.K., Schüssler, M., Mursula, K., Alanko, K.: 2004, *Astron. Astrophys.* **413**, 745. doi:[10.105/0004-6361:20031533](https://doi.org/10.105/0004-6361:20031533).
- Wang, Y.M., Lean, J., Sheeley, N.R.: 2000, *Geophys. Res. Lett.* **27**, 505.
- Wang, Y.M., Sheeley, N.R.: 2002, *J. Geophys. Res.* **107**(A10), 1302. doi:[10.1029/2001JA000500](https://doi.org/10.1029/2001JA000500).
- Weber, S.L., Crowley, T.J., Van der Schier, G.: 2004, *Clim. Dyn.* **22**, 539.
- Weiss, N.O.: 1987, In: Stephenson, F.R., Wolfendale, A.W. (eds.) *Secular Solar and Geomagnetic Variations in the Last 10 000 Years*, University of New Hampshire, Durham, 69.
- Weiss, N.O., Tobias, S.M.: 2004, *Space Sci. Rev.* **94**, 99.