Mon. Not. R. astr. Soc. (1990) **243**, 241–251

Sunspot activity as originating in interference of the Sun's global MHD oscillations

M. H. Gokhale and J. Javaraiah
Indian Institute of Astrophysics, Bangalore 560 034, India

Accepted 1989 September 8. Received 1989 September 6; in original form 1988 November 20

SUMMARY

Spherical-harmonic-Fourier analysis of sunspot activity during 1902–54 shows that the sunspot occurrence probability $p(\theta, \phi, t)$ is given by superposition of, mainly, axisymmetric even degree SHF modes of $l \leq 22$, all having periods ~ 11 yr. The amplitudes and the phases of these modes are found to remain approximately constant over the five cycles, but the small variations in the phases seem to be systematic and coherent. The combination of the amplitudes and the phases determined is unique in reproducing the main qualities of the butterfly diagrams.

A 'nominal toroidal field', defined in terms of $p(\theta, \phi, t)$, yields axisymmetric modes of *odd* degrees, $l \leq 13$, and periods ≈ 22 yr. The amplitude spectrum of these modes is similar to that of the 'magnetic' modes obtained recently by Stenflo from the 'radial' fields observed during 1960–86. Reasons are given for believing that these 'magnetic' modes represent global slow MHD oscillations of the Sun.

If confirmed by further studies, the above conclusions can be used to study the 'global magnetic oscillations' during the pre-magnetogram decades.

1 INTRODUCTION

Maunder's well-known 'butterfly diagrams' show that the sunspot activity originates in 'waves' with periods ~ 11 yr (or multiples thereof), which propagate from middle latitudes in each solar hemisphere towards the solar equator (Becker 1955). Waves of one-way migrations, limited in latitudes, are also seen in the equatorward migrations of fast rotation zones (Snodgross 1987 and references therein) and poleward migrations of weak magnetic fields (Howard 1974), prominence belts and polar faculae (Makarov & Sivaraman 1983, 1986). It could be that all these waves are in fact global (i.e. pole-to-pole), but appear confined to certain latitudes in their manifestations in the respective observations. Supposing this is true, each wave contributing to the overall pattern would be equivalent to a set of at least approximately stationary global oscillations of the Sun with appropriate phase differences.

The following questions would then arise: (i) which set of oscillation modes contributes with appreciable amplitudes; (ii) what are the amplitudes and phases of these modes, (iii) how do the amplitudes and phases vary in time, (iv) what is the physical nature of these oscillations, and so on.

In an attempt to answer these questions we define sunspot occurrence probability and a rough measure of the underlying toroidal magnetic field empirically, as functions of heliographic coordinates and time, and subject these functions to spherical-harmonic-Fourier ('SHF') analysis using sunspot data obtained during 1902–54.

We find that the distribution of sunspot occurrence-probability can be considered as given by a superposition of mainly axisymmetric ($m = 0$) *even* degree 'SHF modes' with $l \leq 22$, and periods ~ 11 yr, whose amplitudes and phases remain fairly constant throughout the 53 years of the data analysed. The amplitudes and phases as well as their secular variations are as given in Section 3.

In Section 4 we illustrate that this combination of amplitudes and phases is the only combination that can reproduce the qualitative features of the butterfly diagrams.

In Section 5 we show that the SHF power of the 'nominal toroidal field' is concentrated mainly in axisymmetric *odd* degree modes of $l \leq 13$ and periods ~ 22 yr, with amplitude spectra similar to that derived by Stenflo (1988) from the photospheric magnetic fields observed during 1960–85.

In Section 6 we list our conclusions and discuss their significance. We also discuss the necessity and directions of further data analysis and theoretical work.

2 DATA AND METHOD OF ANALYSIS

2.1 The data

For this analysis we have used the data from Ledgers I and II of the Greenwich photoheliographs results. A magnetic tape of this data for the years 1874–1976 was kindly provided by H. Balthasar of Gottingen University. From this data we chose for this analysis the heliographic coordinates (λ, ϕ) and the time of observation (t) of all the sunspot groups

during the five sunspot cycles (Waldmeier Nos. 14–18) from 1902 to 1954. We have in all 50795 data points (λ, ϕ, t) wherein each day's observation of a spot group is treated as an independent data point. In this way the data corresponding to each spot group are weighted by the number of days on which it is observed.

2.2 Definition of sunspot occurrence probability

For each time interval (T_1, T_2) chosen for analysis (e.g. a sunspot cycle or a sequence of cycles) we define 'sunspot occurrence probability' $p(\mu, \phi, \tau)$ as

$$p(\mu, \phi, \tau) = \begin{cases} \dfrac{1}{N}\,\delta(\mu-\mu_i, \phi-\phi_i, \tau-\tau_i) \\ \text{at } (\mu_i, \phi_i, \tau_i),\ i=1, 2, \ldots N, \\ 0 \text{ elsewhere in } (\mu, \phi, \tau) \text{ space} \end{cases} \quad (1)$$

where $\tau=(t-T_1)/(T_2-T_1)$, $\mu=\cos\,\theta$, $\theta=90°-\lambda$, δ represents a delta function in (μ, ϕ, τ), $N=N(T_1, T_2,)$ is the number of data points during the interval (T_1, T_2) and t is the time of observation (in days, including the fraction of the day of the observation) from the *zero* hour of the *first day* of the interval (T_1, T_2).

Owing to the non-zero (though small) uncertainties in μ_i, ϕ_i and τ_i, the delta functions here must be considered as mathematical idealizations of properly normalized 'physical' delta functions of large finite values over small finite domains.

The 'first' days of the cycles 14–18 were chosen as the sunspot minima 1902.0, 1912.5, 1923.0, 1933.5 and 1944.0 obtained from three point averages of six-monthly Zurich sunspot numbers.

2.3 Formulae for SHF components

The harmonic components $H_{cc}\ (l, m, n,|\, T_1, T_2)$, $H_{cs}\ (l, m, n,|\, T_1, T_2)$, $H_{sc}(l, m, n,|\, T_1, T_2)$ and $H_{ss}(l, m, n,|\, T_1, T_2)$ in the expansion

$$p(\mu, \phi, \tau) = \sum_{\alpha, l, m, n} H_\alpha(l, m, n|\, T_1, T_2) \times P_l^m(\mu)^{\cos}_{\sin}(m\phi)^{\cos}_{\sin}(2\pi n\tau) \quad (2)$$

during the interval (T_1, T_2) are given by

$$H_\alpha(l, m, n|\, T_1, T_2) = C(l, m, n) \int_0^1 d\tau \int_0^{2\pi} d\phi \int_{-1}^{+1} d\mu\, p(\theta, \phi, \tau) \times P_l^m(\mu)^{\cos}_{\sin}(m\phi)^{\cos}_{\sin}(2\pi nt) \quad (3)$$

$$= \frac{C(l, m, n)}{N} \sum_i P_l^m(\mu_i)^{\cos}_{\sin}(m\phi_i)^{\cos}_{\sin}(2\pi n\tau_i), \quad (4)$$

where α is a symbol representing the subscript 'cc', 'cs', 'sc' or 'ss', depending upon the combination of the cosines or sines of $(m\phi)$ and $(2\pi n\tau)$ in the respective term, and

$$C(l, m, n) = \Lambda \frac{(l-m)!(2l+1)}{(l+m)!\quad \pi}$$

with

$$\Lambda = \begin{cases} 1 \text{ for } m\neq 0,\ n\neq 0, \\ \tfrac{1}{2} \text{ if } m=0 \text{ or } n=0, \\ \tfrac{1}{4} \text{ if } m=0 \text{ and } n=0. \end{cases}$$

2.4 Determination of amplitudes and phases

2.4.1 *Referred to $t=T_1$ as zero epoch*

The amplitudes $A_c(l, m, n)$ and $A_s(l, m, n)$ of the modes $P_l^m(\cos\theta)\cos(m\theta)e^{2\pi in\tau}$ and $P_l^m(\cos\phi)\sin(m\phi)e^{2\pi in\tau}$ during (T_1, T_2) are given by

$$A_c(l, m, n)=[H_{cc}^2(l, m, n)+H_{cs}^2(l, m, n)]^{1/2},$$
$$A_s(l, m, n)=[H_{sc}^2(l, m, n)+H_{ss}^2(l, m, n)]^{1/2}. \quad (5)$$

We also define the rms amplitude of $P_l^m(\mu)$ term as

$$A(l, m, n)=[A_c^2(l, m, n)+A_s^2(l, m, n)]^{1/2}.$$

For $m=0$: $A(l, m, n)=A_c(l, m, n)$.

By phases $\varphi_c\ (l, m, n)$ and $\varphi_s(l, m, n)$ of the above modes during (T_1, T_2) 'referred to $t=T_1$ as the zero epoch' we mean the values of φ in their time dependence expressed as

$$\sin[2\pi\nu(t-T_1)+\varphi],$$

where

$$\nu=n/(T_1-T_2).$$

These phases are given by

$$\varphi_c(l, m, n)=\tan^{-1}[H_{cc}(l, m, n)/H_{cs}(l, m, n)]+0 \text{ or } \pi \text{ and}$$
$$\varphi_s(l, m, n)=\tan^{-1}[H_{sc}(l, m, n)/H_{ss}(l, m, n)]+0 \text{ or } \pi, \quad (6)$$

where 0 *or* π is chosen to ensure the correct signs for the sines and consines.

For axisymmetric $(m=0)$ modes φ_s is undefined and the symbol φ_c will be replaced by φ.

2.4.2 *Amplitudes and phases referred to other zero epochs*

It can be shown that the above formulae also give the amplitudes and phases referred to *any epoch* T_0 *other than* T_1 *as the zero epoch* if in equation (4) one takes

$\tau=(t-T_0)/(T_2-T_1)$ instead of $(t-T_1)/(T_2-T_1)$.

In such a shift of zero epoch, the amplitudes remain invariant and the phases shift by $2\pi\nu(T_1-T_0)/(T_2-T_1)$.

2.5 Estimates of uncertainties and errors

The uncertainties δA_c, δA_s, $\delta\varphi_c$ and $\delta\varphi_s$ in A_c, A_s, φ_c and φ_s arise from errors and uncertainties $\delta\lambda$, $\delta\phi$ and δt in λ, ϕ and t and from the sunspot activity *missed* by observations.

The relative errors $\delta\lambda/\lambda$, $\delta\phi/\phi$ and $\delta t/t$ are $<10^{-2}$. The corresponding relative uncertainties $\delta A_c/A_c$, $\delta A_s/A_s$, etc. are $\sim(1/\sqrt{N})\,(\delta\lambda/\lambda+\delta\phi/\phi+\delta t/t)$. These are $<10^{-4}$ for A_c, A_s, etc. computed from data lengths of a sunspot cycle or longer.

If each sunspot group is viewed as providing a single data point weighted by the number of days on which it is observed then the spreads in t_i, together with those in λ_i and ϕ_i due to the proper motions of sunspots, will raise the above upper

Table 1. Values of SHF amplitudes $A(l, m, \nu)$ for $l=0$–9, $m=0$–2 and $\nu=0$–5 (Units $=\frac{1}{22}$ yr^{-1}), obtained from analysis of $p(\theta, \phi, t)$ during the whole sequence 1902–54.

l	m	$\nu=0$	$\nu=1$	$\nu=2$	$\nu=3$	$\nu=4$	$\nu=5$
0	0	0.159E+0	0.282E−1	0.652E−1	0.107E−1	0.798E−2	0.288E−2
1	0	0.291E−3	0.127E−2	0.331E−2	0.206E−2	0.313E−2	0.284E−2
1	1	0.191E−1	0.345E−2	0.158E−1	0.112E−1	0.127E−1	0.749E−2
2	0	0.311E+0	0.563E−1	0.130E+0	0.203E−1	0.124E−1	0.766E−2
2	1	0.407E−2	0.210E−2	0.609E−2	0.384E−2	0.397E−2	0.632E−2
2	2	0.320E−2	0.447E−2	0.292E−2	0.286E−2	0.265E−2	0.332E−2
3	0	0.291E−2	0.375E−2	0.954E−2	0.677E−2	0.845E−2	0.808E−2
3	1	0.702E−2	0.190E−2	0.533E−2	0.403E−2	0.479E−2	0.177E−2
3	2	0.122E−2	0.205E−2	0.128E−2	0.154E−2	0.109E−2	0.101E−2
4	0	0.203E+0	0.419E−1	0.144E+0	0.181E−1	0.936E−2	0.106E−1
4	1	0.422E−2	0.281E−2	0.553E−2	0.385E−2	0.403E−2	0.624E−2
4	2	0.106E−2	0.114E−2	0.860E−3	0.870E−3	0.573E−3	0.785E−3
5	0	0.919E−2	0.495E−2	0.136E−1	0.125E−1	0.107E−1	0.105E−1
5	1	0.366E−2	0.170E−2	0.116E−2	0.412E−3	0.198E−2	0.246E−2
5	2	0.695E−3	0.128E−2	0.697E−3	0.832E−3	0.725E−3	0.389E−3
6	0	0.163E−1	0.241E−1	0.123E+0	0.257E−1	0.255E−1	0.106E−1
6	1	0.342E−2	0.329E−2	0.335E−2	0.284E−2	0.355E−2	0.482E−2
6	2	0.541E−3	0.462E−3	0.568E−3	0.429E−3	0.183E−3	0.296E−3
7	0	0.159E−1	0.359E−2	0.139E−1	0.187E−1	0.975E−2	0.849E−2
7	1	0.328E−2	0.164E−2	0.151E−2	0.193E−2	0.266E−2	0.254E−2
7	2	0.381E−3	0.789E−3	0.349E−3	0.428E−3	0.532E−3	0.261E−3
8	0	0.212E+0	0.464E−1	0.162E+0	0.350E−1	0.367E−1	0.954E−2
8	1	0.273E−2	0.348E−2	0.194E−2	0.182E−2	0.301E−2	0.361E−2
8	2	0.276E−3	0.173E−3	0.387E−3	0.167E−3	0.136E−3	0.128E−3
9	0	0.186E−1	0.364E−2	0.141E−1	0.231E−1	0.835E−2	0.530E−2
9	1	0.267E−2	0.147E−2	0.185E−2	0.144E−2	0.316E−2	0.286E−2
9	2	0.200E−3	0.471E−3	0.163E−3	0.218E−3	0.356E−3	0.261E−3

limit on $\delta A_c/A_c$ etc., but to a value still $<10^{-3}l$. Further, the differential rotation will cause a spread in ϕ_i whose magnitude in degrees will be $\sim(2.7)\times(\sin^2\lambda)\times$(life span in days). The upper limit on the average value of this spread will be $<10°$, i.e. less than 10 times the maximum error in the daily values of ϕ. Hence the upper limit on $\delta A_c/A_c$, etc. for non-axisymmetric ($m\neq 0$) will be $<10^{-2}m$. Thus, $\delta A_c/A_c$ etc., will be $<10^{-3}l$ for the axisymmetric modes, and $<10^{-2}m$ for the non-axisymmetric modes.

Since the definition of $p(\lambda, \phi, t)$ is normalized with respect to $N(T_2-T_1)$, the errors due to the 'missing' of activity on the unobservable side of the Sun will be small for $(T_2-T_1)\gg 27$ d. Since the sunspot activity is confined to $\sin^2\lambda\lesssim\frac{1}{4}$, the errors due to its 'undersampling' at different latitudes caused by the differential rotation $\Omega=\Omega_0-\Omega_1\sin^2\lambda$, with $\Omega_1/\Omega_0\approx\frac{1}{5}$, will be <5 per cent.

3 RESULTS

For $p(\theta, \phi, t)$ we have calculated the values of $A(l, m, \nu)$ for $l=0$–9, $m=0$–l, $\nu=0$–9, and of $A(l, m, \nu)$ as well as $\varphi(l, m, \nu)$, with $l=0$–29, $m=0$, $\nu=0$–9, taking units of ν as $\frac{1}{11}$ yr^{-1} for the data during single cycles, $\frac{1}{22}$ or $\frac{1}{44}$ yr^{-1} for sequences of three cycles and $\frac{1}{22}$ or $\frac{1}{44}$ yr^{-1} for the sequence of all the five cycles.

3.1 Frequencies

In Table 1 we illustrate the values of $A(l, m, \nu)$ for $l=0$–9, $m\leq 2$ and $\nu=0$–5 in units of $\frac{1}{22}$ yr^{-1} derived from $p(\theta, \phi, t)$ during the whole sequence (1902–54).

It is clear that with the exception of the mode ($l=1$, $m=1$), the SHF amplitudes of the non-axisymmetric ($m>0$) modes are quite small compared to those of the axisymmetric ($m=0$) modes of the same degree (l).

Among the axisymmetric modes, the even degree modes have much higher amplitudes than the odd degree modes. All axisymmetric modes of even degree have maximum amplitudes at $\nu=\frac{1}{11}$ yr^{-1}.

For the axisymmetric ($m=0$) even degree modes with $l\leq 22$, the dependence of $A(l, m, \nu)$ on ν, in units of $\frac{1}{44}$ yr^{-1}, is shown in Fig. 1.

Since the power is proportional to the square of the amplitude, it is clear that most of the SHF power is concentrated in *axisymmetric modes of even degrees and frequencies* $\approx$ 1—11 yr^{-1} (see also Section 3.2). Since the resolution in ν is $\approx\frac{1}{53}$ yr^{-1}, the periods of all these modes can be considered as *approximately* 11 yr.

3.2 The amplitude spectrum with respect to l

3.2.1 *As given by the real data*

In Fig. 2(a) we show $A_l=A(l, m, \nu)$ as a function of l for $0\leq l\leq 29$, $m=0$ and $\nu=\frac{1}{11}$ yr^{-1} derived from the data during the individual sunspot cycles 14–18 and from the combined sequence of these five cycles.

Each of these time intervals yields *distinct* smooth forms of $A(l)$ for the *even* and the *odd* values of l. For $l<24$ the amplitudes of the odd degree modes are much smaller, though not negligible, compared to those of the even degree modes.

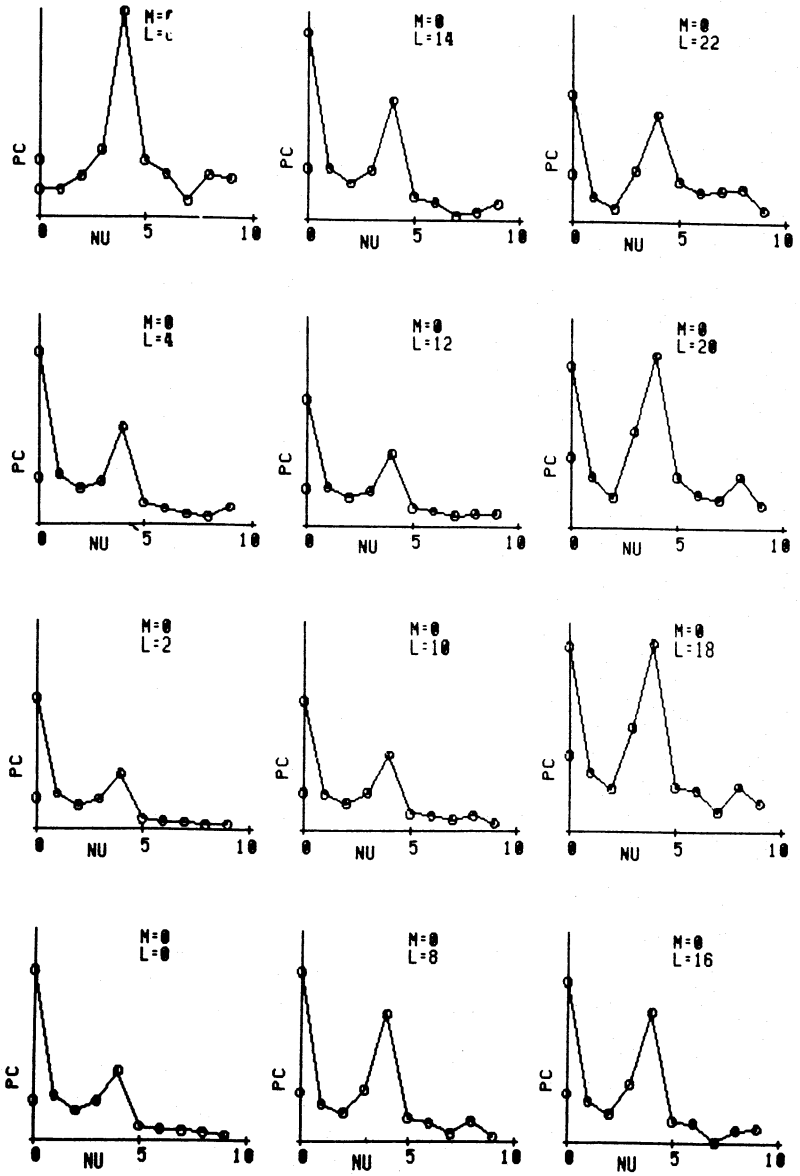

Figure 1. SHF amplitudes $A(l, m, \nu)$, of $p(\theta, \phi, t)$ as functions of ν (in units of $\frac{1}{44}$ yr^{-1}) for axisymmetric ($m=0$) modes of even degree up to $l=22$ during the whole sequence 1902–54. (Symbols: PC = A, L = l, M = m, NU = ν and the dot on the ordinate represents the average of A.)

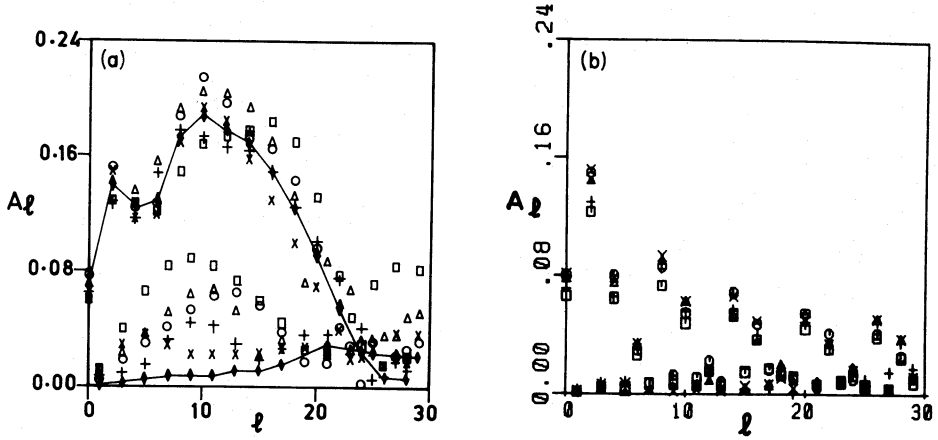

Figure 2. (a) SHF amplitude $A_l = A(l, m, \nu)$, of $p(\theta, \phi, t)$, as a function of l for $0 \le l \le 29$, $m=0$ and $\nu = \frac{1}{11}$ yr^{-1}, during individual cycles 14–18 represented by the symbols □, ○, △, + and ×, respectively. The filled diamonds represent values during the whole sequences, and are connected separately for the even and the odd values of l. (b) Same as above, obtained from simulated data sets for cycles 14–18, with real values of t_i but random values of λ_i between ± 30°.

The form of $A_l(l)$ for the even degree modes is approximately constant from one time interval to another and most of the power is concentrated in $l<24$. Thus the amplitude spectrum with respect to the even degree modes is more or less independent of the choice of the interval T_1, T_2, at least when $(T_2-T_1)\geq 11$ yr.

3.2.2 *Comparison with the spectrum given by simulated random data*

In Fig. 2(b) we give the amplitude spectrum obtained from simulated data sets with real values of t_i but random values of λ_i between $\pm 30°$. Comparison of Fig. 2(a) and (b) shows that the amplitude spectrum in part (a) would not be obtained if sunspots occurred at random latitudes within $\pm 30°$ *even if* they occurred at the *observed epochs.*

3.3 The phases of the SHF modes and their cycle-to-cycle variation

In Table 2(a) we give the values of $\varphi_l=\varphi(l,0,11\ \mathrm{yr}^{-1})$ derived from the individual cycles referred to their initial minima as the zero epochs, and from the whole sequence referred to 0.0 UT on 1902 January 1 as its zero epoch.

The cycle-to-cycle changes in φ_l have values $\sim 35°$, and the same signs, for *all* even degree modes up to $l\approx 24$ (see Table 2b). The sign of these changes alternates in time showing that the phase variations are not caused by the difference between the values of the actual periods of the modes and the value chosen in the computations. The changes could be reduced to negligible values by shifting the zero-epochs for odd numbered cycles backward in time (or those of the even numbered cycles forward in time) by about 12.5 months. However, the alternate lengthening and shortening of the oscillation period by such shifts would itself imply a *real and simultaneous* variation in the phases of all the *even* degree modes up to $l\approx 24$.

3.4 The constancy of the relative phases

As a consequence of the near equality of the changes in φ_l for the even degree modes, the *relative* phases $(\varphi_l-\varphi_0)$ from cycle to cycle are *much more constant* than the *absolute* phases, and have values nearly the same as those derived from the whole sequence (see Table 3a).

Constancy of the relative phases can result even from 'simulated' data sets with *real* values of t_i but *random* values of λ_i between $\pm 30°$ (see Table 3b). However, these relative phases are either 0° or 180° since all 'l' modes have phases φ_0 or $\varphi_0+180°$ where φ_0 is the phase of the variation of the *total amount of data.* Moreover the corresponding amplitude spectrum is noisy, as seen in Fig. 2(b).

Table 2. (a) Phases φ_l (in degrees modulo 360) during each of the sunspot cycle nos 14–18 and during the whole sequence 1902–54, with their initial epochs as zero epochs.

Cycle nos	$l=0$	$l=2$	$l=4$	$l=6$	$l=8$	$l=10$	$l=12$	$l=14$	$l=16$	$l=18$	$l=20$	$l=22$
14	275	88	245	23	164	315	107	259	56	221	32	208
15	244	56	207	326	108	264	58	205	356	156	322	132
16	275	83	229	1	145	295	85	266	37	198	4	179
17	255	62	202	332	121	271	59	214	18	187	0	178
18	292	103	252	8	149	304	102	260	59	222	29	194
1902–54	277	88	236	1	144	286	90	244	42	206	16	191

Cycle nos	$l=1$	$l=3$	$l=5$	$l=7$	$l=9$	$l=11$	$l=13$	$l=15$	$l=17$	$l=19$	$l=21$
14	353	170	345	157	325	129	288	85	239	19	119
15	357	165	324	119	270	77	245	61	248	100	8
16	84	275	115	318	155	347	179	30	327	162	343
17	158	317	83	230	40	217	32	194	355	175	0
18	211	33	214	28	181	322	106	245	34	200	24
1902–54	217	86	327	156	319	90	254	98	324	174	14

Table 2. (b) Cycle-to-cycle variation of φ_l

Cycle nos $l=$	0	2	4	6	8	10	12	14	16	18	20	22	Mean	rms
14→15	−31	−31	−37	−56	−57	−49	−49	−54	−61	−66	−59	−77	−49	±12
15→16	30	27	22	35	38	30	27	34	41	43	41	48	33	±6
16→17	−19	−21	−28	−29	−24	−23	−25	−24	−18	−12	−4	−2	−22	±5
17→18	37	41	50	36	27	33	43	46	41	36	29	16	39	±6
													36	±10
Mean magnitude	29	30	34	39	36	34	26	40	40	39	33	36:	34.7	±4.8
rms	±6	±7	±10	±10	±13	±10	±16	±11	±15	±19	±20	±34		

Table 3. (a) 'Relative phases' $\varphi_l - \varphi_0$ during each cycle, their mean values and rms deviations over the five cycles, and the values during the whole sequence 1902–54.

Cycle nos	$\varphi_2-\varphi_0$	$\varphi_4-\varphi_0$	$\varphi_6-\varphi_0$	$\varphi_8-\varphi_0$	$\varphi_{10}-\varphi_0$	$\varphi_{12}-\varphi_0$	$\varphi_{14}-\varphi_0$	$\varphi_{16}-\varphi_0$	$\varphi_{18}-\varphi_0$	$\varphi_{20}-\varphi_0$	$\varphi_{22}-\varphi_0$	$\varphi_{24}-\varphi_0$	$\varphi_{26}-\varphi_0$	$\varphi_{28}-\varphi_0$
14	172	−31	108	−111	40	191	−17	141	−54	117	−67	124	9	−170
15	172	−37	82	−136	20	173	−39	112	−88	78	−112	62	31	171
16	168	−46	86	−130	19	170	−37	122	−77	83	−96	108	−32	154
17	167	−53	77	−139	16	164	−41	123	−68	89	−77	103	−90	−45
18	171	−40	76	−143	12	170	−32	127	−70	96	−98	50	179	−35
Mean	170±2	−41±8	86±12	−132±11	21±10	174±9	−33±9	125±9	−71±11	93±14	−90±16	89±28	19±90	15±130
1902–54	170	−42	84	−134	19	172	−34	125	−71	92	−86	92	−44	171

Table 3. (b) Relative phases $\varphi_l - \varphi_0$ during each cycle and during the whole sequence, obtained from the 'simulated' data sets with real values of t_i and random values of λ_i between ±30°.

Cycle nos	$\varphi_2-\varphi_0$	$\varphi_4-\varphi_0$	$\varphi_6-\varphi_0$	$\varphi_8-\varphi_0$	$\varphi_{10}-\varphi_0$	$\varphi_{12}-\varphi_0$	$\varphi_{14}-\varphi_0$	$\varphi_{16}-\varphi_0$	$\varphi_{18}-\varphi_0$	$\varphi_{20}-\varphi_0$	$\varphi_{22}-\varphi_0$	$\varphi_{24}-\varphi_0$	$\varphi_{26}-\varphi_0$	$\varphi_{28}-\varphi_0$
14	179	−4	16	−174	6	−9	−178	4	1	−176	10	−18	175	−3
15	−180	1	−2	−180	1	−1	−177	10	−91	163	−7	−25	165	−16
16	−180	2	−6	−180	6	−22	176	3	−81	172	9	−53	156	−19
17	180	−3	12	−175	7	−4	−175	15	−39	170	−2	−32	164	−19
18	−180	0	1	−178	6	−26	172	−10	7	177	−13	22	−171	12
1902–54	180	−1	2	−179	2	−6	179	2	−14	174	−9	16	179	11

3.5 Year-to-year variation of the phases

We determine the year-to-year variation of φ_l's from ten-year intervals, each shifted by one year with respect to the previous one, all referring to 0.0 UT on 1902 January 1 as the zero epoch. This way each interval samples almost a full cycle of each mode. Values of φ_l for even degree modes up to $l=22$, determined from the intervals 1902–11 to 1945–54, are shown in Fig. 3.

We note that: (i) the approximate constancy of φ_l and the much *better* constancy of the *relative* phases is valid on a year-to-year basis also; (ii) phases of $l=0$, 2 and 4 are nearly the same as those of $l=10$, 12 and 14, respectively, and (iii) phases of many modes show a small but rapid advance during a few intervals centred near sunspot minima, followed by a slow retreat during the following several intervals.

3.6 Interpretation of the results in Sections 3.1–3.5

From Sections 3.1–3.5 it is clear that the sunspot occurrence probability during 1902–54 can be expressed as

$$p(\theta, \phi, t) = p(\theta, t) = \sum_{l}^{22} A_l P_l(\cos\theta) \sin(2\pi\nu t + \varphi_l), \qquad (7)$$

where (i) $\nu \approx \frac{1}{11}$ yr^{-1} for all l, (ii) A_ls of even modes are dominant and approximately constant throughout the 53 years, (iii) the phases of the even modes undergo small but coherent variations and (iv) the contribution from the odd modes is small but non-negligible.

4 UNIQUENESS OF AMPLITUDES AND PHASES IN QUALITATIVELY REPRODUCING THE BUTTERFLY DIAGRAMS

4.1 Latitude–time distribution from the derived amplitudes and phases

Yearly means of the monthly values of expression (7) during 1902–54, with amplitudes A_l given in Fig. 2 and phases φ_l given in Table 2, yield the latitude–time distribution of $p(\theta, \phi, t)$ shown in Fig. 4(a). This distribution agrees qualitatively with the butterfly diagrams and the migrations of background field features in the following respects: (i) high values of p are confined to latitudes ≤30°, (ii) these migrate towards the equator, (iii) the highest values occur at ±10°–15°, (iv) the 'very small probability points' in higher latitudes may appear to migrate *either* towards the lower latitudes, *or* towards the poles, *depending upon one's point of view.*

There are some important differences between Fig. 4(a) and the real butterfly diagrams. For example (i) the values of p in the 'overlap regions' of successive butterflies are too small, (ii) appreciable (though not high) values are found in high latitudes in some years, and (iii) the sum of the values over all latitudes does not show the asymmetry of the sunspot cycle, etc. However, such detailed comparison is not called for until we determine how the observational parameters in the various butterfly diagrams are related to the function $p(\theta, \phi, t)$ defined here.

4.2 Uniqueness of amplitude spectrum and relative phases

There is nothing surprising in reproducing qualitatively the butterfly diagrams using the amplitudes and phases determined from the very data that yield the butterfly diagrams.

However, butterfly diagrams *cannot be* reproduced if: (i) the amplitude spectrum varies substantially from cycle to cycle even if the phases remain the same as in Table 2(a) (see Fig. 4b), *or* (ii) if the phases vary substantially from cycle to cycle, even if the amplitudes are chosen as in Fig. 2(a) (see Fig. 4c), *or* (iii) if the amplitude spectrum as well as the phases remain constant from cycle to cycle, but either the amplitude spectrum is not as in Fig. 2(a) (see Fig. 4d), or the phases are not as in Table 2a (see Fig. 4e).

Thus there is a certain amount of uniqueness of l, m, ν, A_l and φ_l in a qualitative reproduction of the butterfly diagrams if sunspot activity is considered as originating in superposition of global oscillations of the Sun.

5 PHYSICAL NATURE OF THE BASIC OSCILLATIONS

5.1 The possibility that the sunspot activity may originate in superposition of a set of Sun's global oscillations

The results in Section 3 suggest the possibility that sunspot activity might originate in superposition of some global oscillations of the Sun.

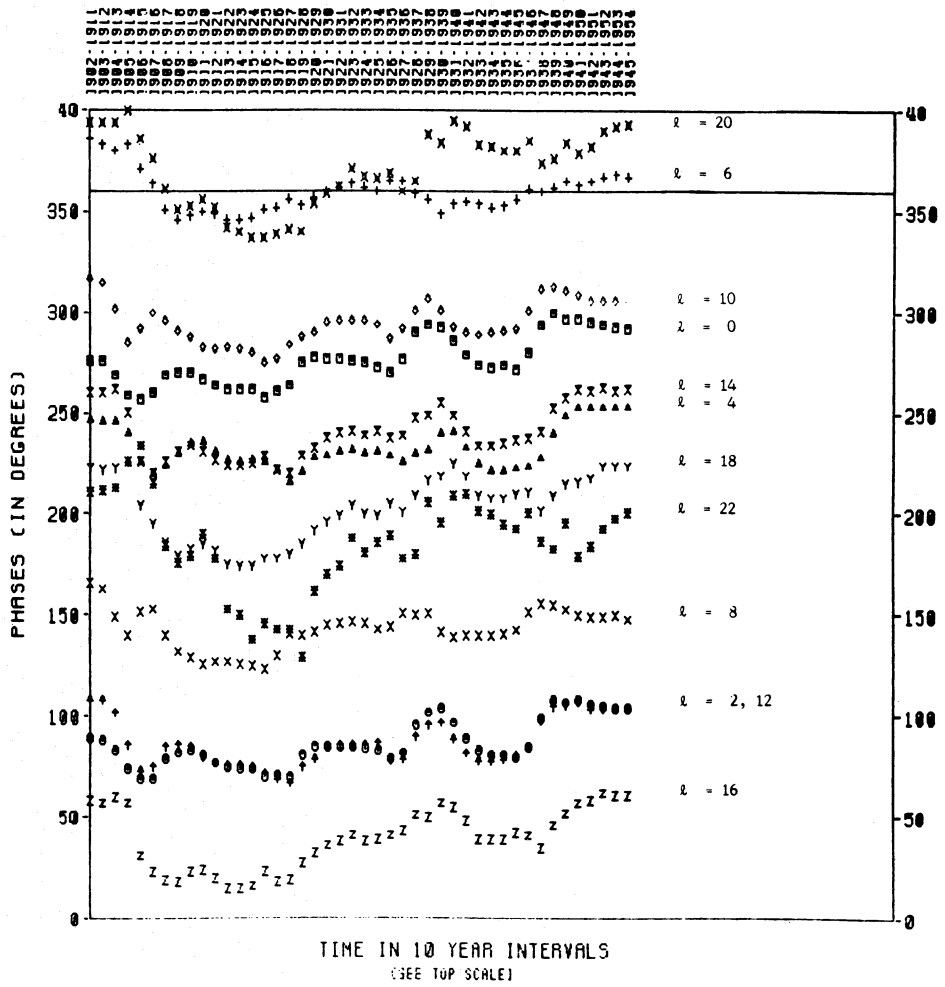

Figure 3. Phases φ_l of even degree modes up to $l=22$ during successive 10-year intervals between 1902 and 1954, all referred to 0.0 UT on 1902 January 1 as the 'zero epoch'.

The set of 'SHF modes' given by the data analysis may not be the same as that representing the 'modes of physical oscillations', if the data variable is *non-linearly* related to the variable describing the physical oscillations. The phase locking of different modes in Fig. 3 suggests that $p(\theta, \phi, t)$ may be a non-linear function of some basic variable(s) describing the global oscillations.

The approximate constancy of the amplitudes and phases of the SHF modes suggests that the physical modes probably have small amplitudes. Their long periods (e.g. 11 yr or multiples thereof) would then imply that the 'basic' modes must be '*slow-MHD*' modes.

5.2 A test of the above hypothesis: SHF analysis of the 'Nominal Toroidal Field'

A strong toroidal magnetic field is essential for production of sunspot activity in general. The toroidal field is not directly measurable, but using Hale's laws of magnetic polarities, we define a 'nominal toroidal field' 'B_ϕ' (θ, t) in the following manner. (We thank Professor M. Stix for this suggestion.)

$$'B_\phi'(\theta, t) = \begin{cases} \pm p(\theta, t) \text{ in the northern solar hemisphere} \\ \mp p(\theta, t) \text{ in the southern solar hemisphere,} \end{cases}$$

the upper signs being taken during the 'even' sunspot cycles and the lower ones during the 'odd' cycles.

The SHF analysis of 'B_ϕ'(θ, t), over two subsequences 1902–33 and 1923–54, and over the whole sequence 1902–54, yields the following results.

(i) Most of the SHF power is concentrated in axisymmetric modes of *odd* degrees and *periods of 22 yr* (Table 4). [This is merely a consequence of the results in Section 3 and the definition of 'B_ϕ'(θ, t).]

(ii) For $\nu = \frac{1}{22}$ yr^{-1} modes, the amplitude spectrum of 'B_ϕ'(θ, t) with respect to l is nearly the same during the two subsequences (1902–33, 1923–54) as during the whole sequence (1902–54) [see Fig. 6].

The definition of 'B_ϕ'(θ, t) implies

$$p(\lambda, t) = |'B_\phi'(\lambda, t)|,$$

which is essentially a '*rectification*' in latitude 'λ' as well as in time 't'.

This has led to the transfer of SHF power (i) *from* each 'ν' *to* '0' and '2ν' in the spectra w.r.t. ν for fixed values of l and (ii) *from* each 'l' *to* '0' and '$2l$' in the spectra w.r.t. l for fixed values of ν. This accounts for the shapes of spectra in Figs 1 and 2 in terms of those in Figs 5 and 6.

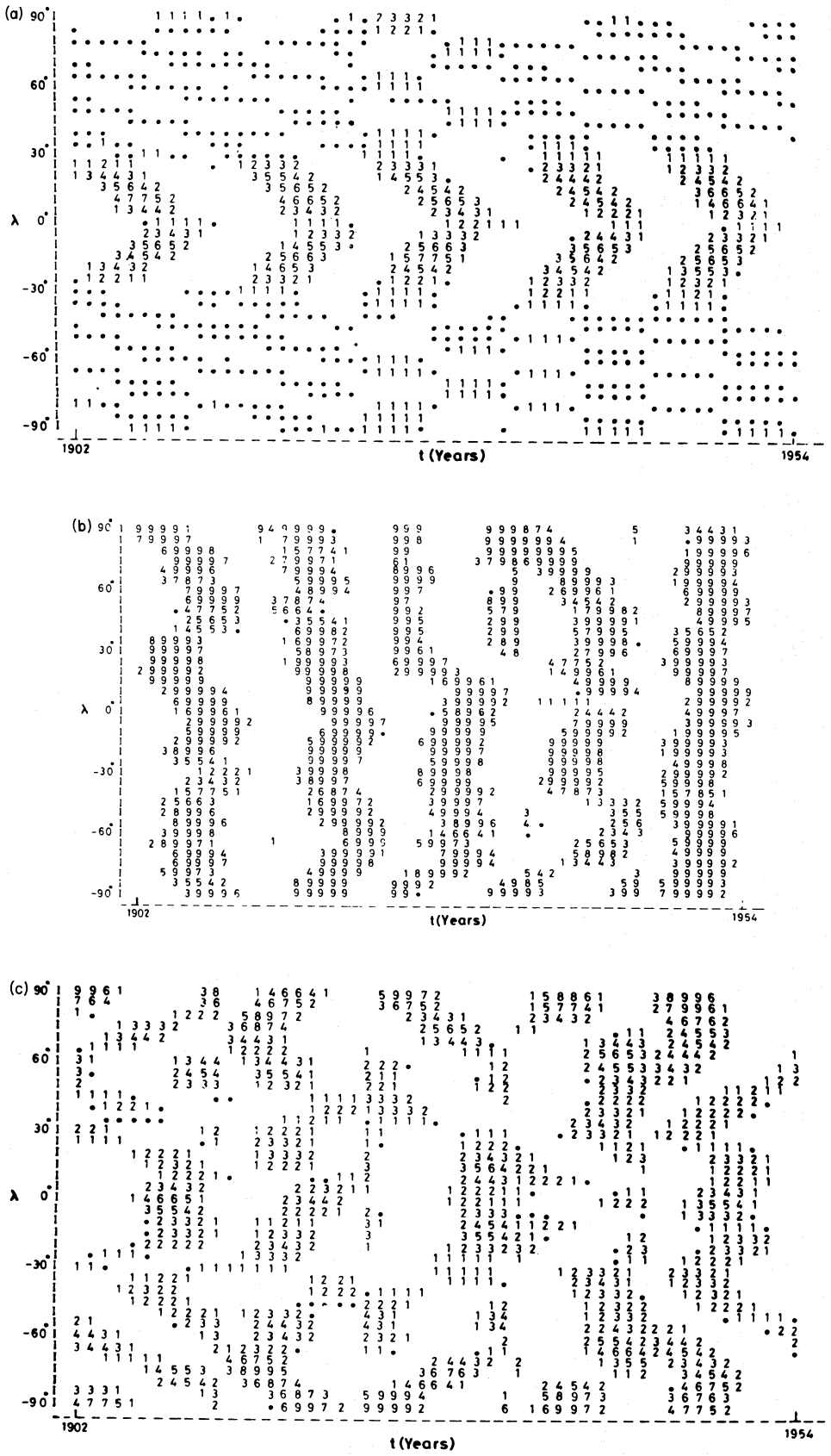

Figure 4. Distribution of the values of $p(\lambda, t)$ given by equation (7) with various combinations of amplitudes and phases. *Symbols:* Blank $= p < 0$, dot $= (0 < p \leq 0.01)$, $1 = (0.01 < p \leq 0.05)$, $2 = (0.05 < p \leq 1.0)$, $3 = (1.0 < p \leq 1.5)$,... etc., $9 = (p > 4.0)$. (These values of p are *not* normalized to unity.) (a) Amplitude spectrum as in Fig. 2(a) and phases as in Table 2(a). (b) Amplitude spectrum varying randomly from cycle to cycle, but phases as in Table 2(a). (c) Amplitude spectrum as in Fig. 2(a), but phases varying randomly, cycle to cycle.

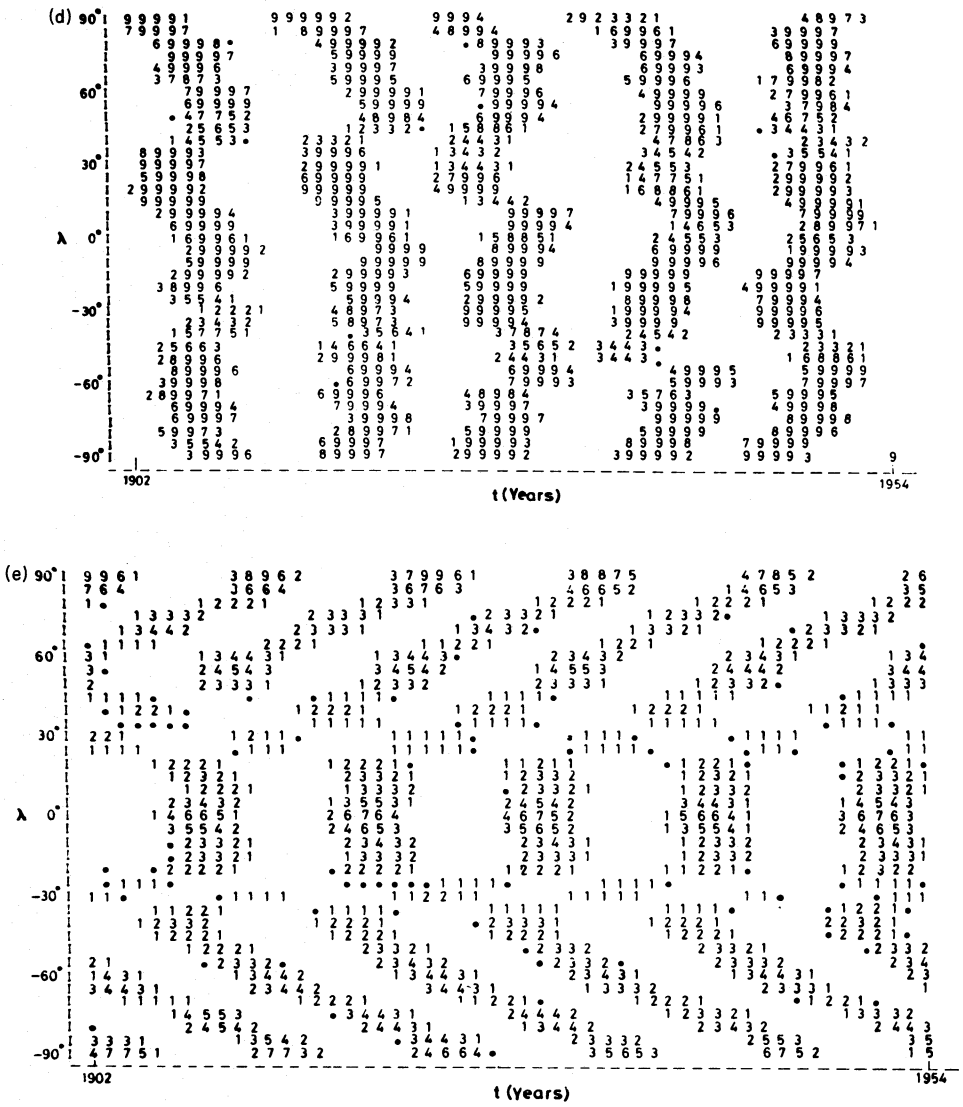

Figure 4 – *continued.* (d) Amplitude spectrum as well as phases constant from cycle to cycle, and phases as in Table 2(a), but amplitude spectrum arbitrary. (e) Amplitude spectrum and phases constant from cycle to cycle, and amplitude spectrum as in Fig. 2(a) but phase values arbitrary.

From the *ad hoc* definition of 'B_φ'(θ, ϕ, t) it is clear that it may or may not represent the real toroidal component of the magnetic field in the Sun. However, in spite of its *ad hoc* definition, and its determination from the data *within a limited zone of latitudes*, 'B_φ'(θ, ϕ, t) yields the same set of 'SHF modes' and strikingly similar 'SHF spectra' (with respect to ν *as well as* with respect to l), as those given by the radial component of magnetic field ('B_r') derived from *full-disc-magnetograms* during 1960–85 (*cf.* Stenflo & Vogel 1986; Stenflo 1988).

Thus, it seems that the set of the SHF modes given by 'B_φ'(θ, ϕ, t) provides an alternative description of the same set of the Sun's global 'magnetic' oscillations as those described by $B_r(\theta, \phi, t)$ derived from the *magnetogram data.*

6 CONCLUSIONS AND DISCUSSION

(i) Sunspot activity may originate in the interference of the Sun's global (probably slow MHD) oscillations, with dominant contribution from axisymmetric modes of odd degrees up to $l \approx 13$ and periods ~ 22 yr.

(ii) The amplitudes and phases of the dominant oscillations remain approximately constant for at least five sunspot cycles, and perhaps even longer, but they have small systematic variations.

(iii) This unique combination of amplitudes and phases can account for (a) the clear migration of strong fields from middle latitudes to equator, and perhaps for (b) the apparently ambiguous migrations of weak field concentrations in higher latitudes.

(iv) The sunspot data can be used to study amplitudes and phases of the 'SHF modes' during the pre-magnetogram decades, at least in the low spatial and temporal frequencies. This will be useful even if the SHF modes do not represent global MHD oscillations.

If we reject the interpretation in terms of global MHD oscillations the only alternative 'explanation' for the butterfly diagrams is in the Babcock (1961) model which is

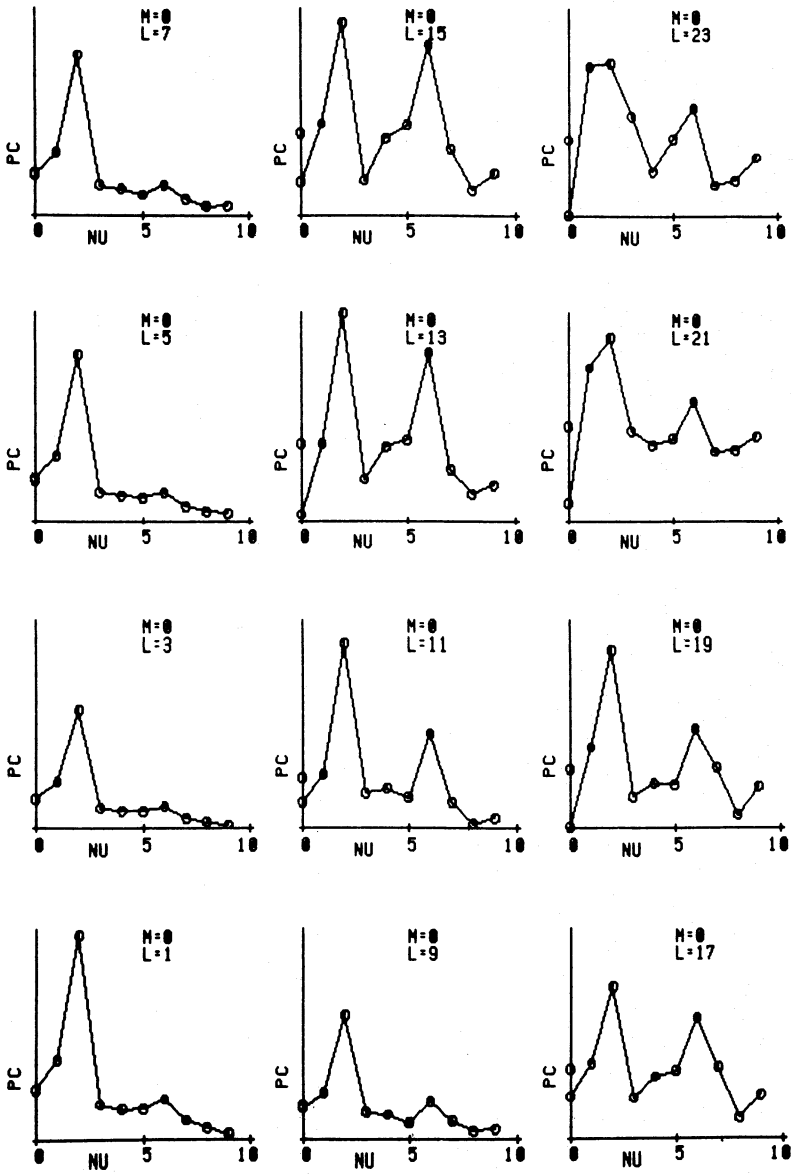

Figure 5. SHF amplitude $A(l, m, \nu)$ of 'B_ϕ'(θ, ϕ, t) as functions of ν (in units of $\frac{1}{44}$ yr^{-1}) for axisymmetric ($m=0$) modes of *odd* degrees up to $l=23$. (Symbols: PC = A, L = l, M = m, NU = ν and the dot on ordinate represents the average of A.)

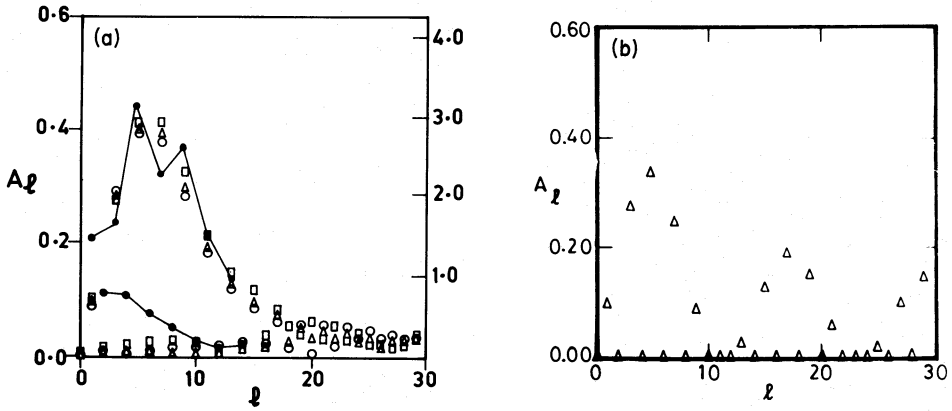

Figure 6. Amplitude $A_l = A(l, m, \nu)$ of 'B_ϕ'(θ, ϕ, t), as a function of l for $0 \le l \le 29$, $m=0$, $\nu = \frac{1}{22}$ yr^{-1} during 1902–33 (□), 1923–54 (○) and 1902–54 (△), all referring to the left-hand-side scale. The filled circles, connected separately for the odd and even values of l, represent amplitudes of $B_r(\theta, \phi, t)$ obtained by Stenflo (1988) from magnetograms during 1960–85 (referred to the right-hand-side scale). (b) Same as (a), obtained from a simulated data set for 1902–54 (△), with real values of t_i but random values of λ_i between ±30°.

Table 4. Values of SHF amplitudes $A(l, m, \nu)$ for $l=0$–9, $m=0$–2 and $\nu=0$–5 (Unit $=\frac{1}{44}$ yr^{-1}), obtained from analysis of 'B_ϕ' (θ, ϕ, t) during the whole sequence 1902–54.

l	m	$\nu=0$	$\nu=1$	$\nu=2$	$\nu=3$	$\nu=4$	$\nu=5$
0	0	0.141E−2	0.276E−2	0.336E−2	0.529E−2	0.468E−2	0.746E−2
1	0	0.239E−1	0.378E−1	0.960E−1	0.166E−1	0.144E−1	0.146E−1
1	1	0.905E−2	0.468E−2	0.104E−1	0.871E−2	0.683E−2	0.126E−1
2	0	0.353E−2	0.664E−2	0.819E−2	0.968E−2	0.990E−2	0.166E−1
2	1	0.152E−2	0.229E−2	0.876E−2	0.533E−2	0.356E−2	0.547E−2
2	2	0.816E−2	0.589E−2	0.445E−2	0.483E−2	0.572E−2	0.425E−2
3	0	0.705E−1	0.110E+0	0.280E+0	0.477E−1	0.419E−1	0.416E−1
3	1	0.496E−2	0.231E−2	0.343E−2	0.235E−2	0.292E−2	0.504E−2
3	2	0.876E−3	0.113E−2	0.637E−3	0.647E−3	0.969E−3	0.893E−3
4	0	0.494E−2	0.109E−1	0.951E−2	0.519E−2	0.942E−2	0.177E−1
4	1	0.145E−2	0.140E−2	0.848E−2	0.515E−2	0.337E−2	0.563E−2
4	2	0.169E−2	0.111E−2	0.784E−3	0.889E−3	0.997E−3	0.639E−3
5	0	0.101E+0	0.155E+0	0.393E+0	0.679E−1	0.600E−1	0.557E−1
5	1	0.327E−2	0.218E−2	0.253E−2	0.216E−2	0.182E−2	0.210E−2
5	2	0.728E−3	0.570E−3	0.391E−3	0.296E−3	0.602E−3	0.427E−3
6	0	0.584E−2	0.139E−1	0.668E−2	0.777E−2	0.841E−2	0.134E−1
6	1	0.108E−2	0.125E−2	0.625E−2	0.372E−2	0.258E−2	0.462E−2
6	2	0.435E−3	0.292E−3	0.188E−3	0.274E−3	0.158E−3	0.238E−3
7	0	0.101E+0	0.151E+0	0.385E+0	0.723E−1	0.633E−1	0.499E−1
7	1	0.154E−2	0.197E−2	0·361E−2	0.255E−2	0.199E−2	0.154E−2
7	2	0.602E−3	0.292E−3	0.348E−3	0.236E−3	0.409E−3	0.307E−3
8	0	0.445E−2	0.115E−1	0.263E−2	0.159E−1	0.904E−2	0.381E−2
8	1	0.587E−3	0.244E−2	0.412E−2	0.216E−2	0.223E−2	0.346E−2
8	2	0.234E−4	0.158E−3	0.148E−3	0.173E−3	0.182E−3	0.287E−3
9	0	0.753E−1	0.109E+0	0.292E+0	0.641E−1	0.577E−1	0.385E−1
9	1	0.597E−3	0.136E−2	0.328E−2	0.171E−2	0.193E−2	0.183E−2
9	2	0.464E−3	0.256E−3	0.347E−3	0.230E−3	0.290E−3	0.297E−3

schematic and in the turbulent dynamo models (e.g. Yoshimura 1975), which are kinematical. In their dynamical version also the latter models require an unreasonable assumption about the radial gradient of the Sun's internal rotation (Gilman & Miller 1981). On the contrary the simple explanation in terms of global oscillations is energetically *plausible* and will be *inherently dynamical if confirmed.*

Secondary processes (like magnetic buoyancy and other instabilities) will be needed for production of magnetic flux tubes that are manifested as sunspots and surrounding activity.

The approximate equality of the periods for all the relevant values of l may be due to the presence of some single 'forcing' oscillation (Alfvén 1943; Dicke 1977). Alternatively the internal magnetic field (and the plasma density) may be so structured as to yield periods in the neighbourhood of 22 yr for all slow MHD modes of free oscillations.

To decide whether the oscillations are 'forced' or 'free' it will be necessary to determine theoretically the oscillation modes allowed by the equations of magneto-hydrodynamics. This will require details of the Sun's internal magnetic field which we do not know.

The analysis should be extended to earlier as well as later sunspot cycles, and also to other parameters like sunspot areas.

The mathematical and quantitative description of the butterfly diagrams derived in this analysis will provide useful constraints for any theoretical modelling of the solar cycle.

REFERENCES

Alfvén, H., 1943. *Arkiv. Mat. Astr. Fys.*, **29A,** No. 12.
Babcock, H., 1961. *Astrophys. J.*, **133,** 211.
Becker, U., 1955. *Z. Astrophys.*, **37,** 47, and references therein.
Bracewell, R. N., 1988. *Mon. Not. R. astr. Soc.*, **230,** 535.
Dicke, R. H., 1978. *Nature,* **276,** 676.
Gilman, P. A. & Miller, J., 1981. *Astrophys. J.*, **46,** 211.
Gough, D., 1986. *Nature,* **319,** 263.
Howard, R., 1974. *Sol. Phys.*, **38,** 283.
Makarov, V. I. & Sivaraman, K. R., 1983. *Sol. Phys.*, **85,** 215, 227.
Makarov, V. I. & Sivaraman, K. R., 1986. *Soln. Dann.*, No. **9**, 64.
Snodgrass, H. B., 1987. *Sol. Phys.*, **110,** 35, and references therein.
Stenflo, J. O., 1988. *Astrophys. Space Sci.*, **144,** 321.
Stenflo, J. O. & Vogel, M., 1986. *Nature,* **319,** 285.
Yoshimura, H., 1975. *Astrophys. J. Suppl.*, **29,** 467.