ON THE INVESTIGATION OF HIDDEN PERIODICITIES WITH APPLICATION TO A SUPPOSED 26 DAY PERIOD OF METEOROLOGICAL PHENOMENA.

BY ARTHUR SCHUSTER, F. R. S.

1. OBVIOUS AND HIDDEN PERIODICITIES. A variable quantity may show periodic changes which become obvious as soon as a sufficient record has been obtained; such are the semi-diurnal changes of the tides, or the eleven years recurrence of sunspot maxima. We may call these *obvious periodicities*. Most often, however, small periodic variations are hidden behind irregular fluctuations, and their investigation then becomes a matter of considerable difficulty. The lunar influence on the daily variation of magnetic forces may serve as an example of such *hidden periodicities*. In the case of lunar effects the investigation of the periodicity is facilitated by our previous knowledge of the period; but additional difficulties arise when the periodic time is one of the unknown quantities. We possess a number of investigations dealing with a periodicity of various terrestrial phenomena, supposed to be coincident with that of the time of revolution of the sun round its axis. But although several authorities have considered the existence of such a period as proved, the scientific world has only reluctantly and very doubtfully accepted its reality. Nor can it be said that this scepticism is not justified, for no one has so far discussed the very essential question whether the results obtained may not be due to merely accidental circumstances.

It is the object of this paper to introduce a little more scientific precision into the treatment of problems which involve hidden periodicities, and to apply the theory of probability in such a way that we may be able to assign a definite number for the probability that the effects found by means of the usual methods are real, and not due to accident.

2. SUMMARY OF THE USUAL METHOD OF FINDING A HIDDEN PERIODICITY. If it be required to investigate a possible period of p intervals in a series of numbers t_1, t_2, t_3, etc., it is usual to solve the problem by some process analogous to the one which is briefly indicated in this paragraph. Let the numbers be arranged

according to the following scheme, where t_1' stands for t_{p+1}, t_1'' for t_{2p+1}, etc.

$$\left.\begin{array}{cccccccccc} t_1 & t_2 & t_3 & . & . & . & . & . & . & t_p \\ t_1' & t_2' & t_3' & . & . & . & . & . & . & t'_p \\ . & . & . & & & & & & & \\ . & . & . & & & & & & & \\ . & . & . & & & & & & & \\ . & . & . & & & & & & & \\ \overset{s-1}{t_1} & \overset{s-1}{t_2} & . & . & . & . & . & . & . & \overset{s-1}{t_p} \\ \hline T_1 & T_2 & T_3 & . & . & . & . & . & . & T_p \end{array}\right\} \quad (1)$$

T_1, T_2, etc., represent the sums of the vertical columns. The quantities T may be expressed by a periodic series of the form:

$$\begin{aligned} S = a_0 + a_1 \cos\theta + a_2 \cos 2\theta + \quad . \quad . \quad . \quad a_p \cos p\theta \\ + b_1 \sin\theta + b_2 \sin 2\theta + \quad . \quad . \quad . \quad b_{p-1} \sin (p-1)\theta \quad , \end{aligned}$$

where $S = T_1$ if we substitute $\theta = \frac{2\pi}{p}$, and generally S becomes T_q by the substitution of $\theta = \frac{2\pi q}{p}$.

The coefficients are determined by a well known process, which gives

$$\left.\begin{aligned} p\, a_0 &= T_1 + T_2 + T_3 + \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad T_p \\ \tfrac{1}{2} p\, a_1 &= T_1 \cos\theta + T_2 \cos 2\theta + T_3 \cos 3\theta + \quad . \quad T_p \cos p\theta \\ \tfrac{1}{2} p\, b_1 &= T_1 \sin\theta + T_2 \sin 2\theta + T_3 \sin 3\theta \quad . \quad . \quad T_p \sin p\theta \end{aligned}\right\} \theta = \frac{2\pi}{p} \quad (2)$$

a_0 is therefore equal to the mean value of all the quantities T or to s times the mean values of all the quantities t. If there is a well marked periodicity corresponding to p intervals, we should expect the value of $r_1 = \sqrt{a_1^2 + b_1^2}$ to have a markedly greater value than when no such periodicity exists, and we may take the quantity $\rho = \frac{r_1}{a_0}$, as a measure of the amplitude of the periodicity corresponding to p intervals. The quantity ρ is determined by the equation:

$$\frac{\rho^2}{4} = \frac{r_1^2}{4 a_0^2} =$$
$$\frac{(T_1 \cos\theta + T_2 \cos 2\theta + \;.\;.\;)^2}{(T_1 + T_2 + \;.\;.\; T_p)^2} + \frac{(T_1 \sin\theta + T_2 \sin 2\theta + \;.\;.\; T_p \sin p\theta)^2}{T_1 + T_2 + \;.\;.\;.\; T_p)^2} \quad (3)$$

3. PROBABILITY FOR DIFFERENT VALUES OF THE AMPLITUDES IF THE ORIGINAL NUMBERS ARE CHOSEN AT RANDOM. The first question which arises refers to the relative probability of

different values of ρ, calculated on the supposition that no true periodicity exists.

We may confine ourselves to such cases as may occur in nature, and fix our attention, say, on events like thunderstorms, earthquakes, or magnetic storms. Supposing it is required to investigate a possible period of p days, we should form the Table (1), entering for t the number of events which have happened on a particular day. Thus, if we are discussing a 30 day period of earthquakes, p would be equal to 30, and if 2 earthquakes occur on the the 33d day we should write $t_3' = 2$. The Table (1) being formed, and the values of ρ calculated, our problem may be stated thus: *What is the probability that ρ should lie between any two assigned values ρ_1 and ρ_2, on the supposition that the events are distributed quite at random?* Before proceeding to calculate this, I put the question in a rather more general form. The number p of intervals, into which the whole period is divided, may be chosen as large as we please; and each interval may be made as small as we like. It may be an hour, or a minute, or a second. If this process is carried sufficiently far, equations (2) become,

$$\begin{aligned} p\,a_0 &= n \\ \tfrac{1}{2}\,p\,a_1 &= \cos k\,t_1 + \cos k\,t_2 + \quad . \quad . \quad . \quad . \quad + \cos k\,t_n \\ \tfrac{1}{2}\,p\,b_1 &= \sin k\,t_1 + \sin k\,t_2 + \quad . \quad . \quad . \quad . \quad + \sin k\,t_n, \end{aligned} \tag{4}$$

where n is the total number of events and k stands for $\frac{2\pi}{T}$, T being the whole length of the period, and t_1, t_2, etc., the times of occurrence of successive events. (The quantities T_1, T_2 do not occur in the future investigation, and their confusion with T is therefore not possible.) Equation (3) becomes

$$\frac{n\,r_1}{2\,a_0} = \Big\{ (\cos k\,t_1 + \cos k\,t_2 + \quad . \quad . \quad . \quad \cos k\,t_n)^2 + (\sin k\,t_1 + \sin k\,t_2 + \quad . \quad . \quad . \quad \sin k\,t_n)^2 \Big\}^{\frac{1}{2}} \quad . \tag{5}$$

The meaning of the right-hand side of this equation is best illustrated by means of a diagram. On a circle with center at O, choose a number of points P_1, P_2, such that the angle between the lines $O\,P_1$ and $O\,P_2$ and a fixed direction are $k\,t_1$, $k\,t_2$, etc. If $O\,P_1$, $O\,P_2$ represent forces of equal intensity but different directions, the right-hand side of (5) gives the magnitude of the resultant. If the events may happen with equal probability at any time, the points P_1, P_2 will be distributed over the circle in such a manner that any direction of the line OP_1, OP_2, is equally probable. It has been shown by Lord Rayleigh,[1] in a paper "On the Resultant of a large number of vibrations of the same pitch and arbitrary phase," that the probability of the resultant having a value lying between s and $s + ds$ in that case is

$$\frac{2}{n} e^{-\frac{s^2}{n}} s\, ds\,, \tag{6}$$

n being the total number of vectors which are combined.

It is a simple matter to pass from this result to the solution of our problem. From (5) and (6) it follows that the probability of the value of $\frac{n r_1}{2 a_0}$ lying between $\frac{n}{2}\rho$ and $\frac{n}{2}(\rho + d\rho)$ is

$$\frac{n}{2} \rho\, e^{-\frac{n\rho^2}{4}} d\rho \tag{7}$$

and this is therefore also the probability that the quantity $\frac{r}{a_0}$, which we have taken as the measure of the amplitude of the periodicity, lies between the values ρ and $\rho + d\rho$. The expectancy for $\frac{r_1}{a_0}$ is

$$\frac{n}{2}\int_0^\infty \rho^2 e^{-\frac{n\rho^2}{4}} d\rho = \sqrt{\frac{\pi}{n}} = \frac{1.77}{\sqrt{n}}\,. \tag{8}$$

Similarly the expectancy for $\left(\frac{r_1}{a_0}\right)^2$ is

$$\frac{n}{2}\int_0^\infty \rho^3 e^{-\frac{n\rho^2}{4}} d\rho = \frac{4}{n}\,. \tag{9}$$

The probability that the value of $\frac{r_1}{a_0}$ exceeds ρ is

$$\frac{n}{2}\int_\rho^\infty \rho\, e^{-\frac{n\rho^2}{4}} d\rho = e^{-\frac{n\rho^2}{4}}\,. \tag{10}$$

Our result may now be expressed as follows:

If a number n of disconnected events occur within an interval of time T, all times being equally probable for every event, and if the

[1] *Phil. Mag.* Vol. X, p. 73 (1880) II.

frequency of occurrence of these events is expressed in a series of the form

$$a\left(1+\rho_1 \cos 2\pi\frac{t-\tau_1}{T}+\rho_2 \cos 4\pi\frac{t-\tau_2}{T}+\ldots\ldots+\rho_p \cos 2p\pi\frac{t-\tau_p}{T}\right),$$

the probability that any coefficient ρ *has a value lying between* ρ *and* $\rho+d\rho$ *is*

$$\frac{n}{2}\rho e^{-\frac{n\rho^2}{4}} d\rho$$

and the expectancy for ρ *is* $\sqrt{\frac{\pi}{n}}$.

In proving the proposition, it was assumed that the number p of intervals into which the period T is subdivided, is very large; but this condition is not essential. To suit accurately the process employed in actual calculations we should have to consider the vectors OP_1, OP_2, to be confined to fixed directions forming angles $\frac{2\pi}{p}$ with each other. But it follows directly from the method employed by Lord Rayleigh in his proof that his results must apply to this case also if p is a multiple of 4. Further, the expression for the expectancy of ρ^2 can be shown to be the same for all values of p. It is not necessary to inquire whether equation (7) also holds in the most general case, when p, for instance, is an odd number, because the process of calculation illustrated by the tabular arrangement (1) and the result (2) is justified only when p is so large that a further increase in p would not produce any material change in the value of the coefficients of Fourier's series. We may therefore accept equations (7), (8), (9) and (10) as applicable to all cases which concern us.

Equation (8) gives the expectancy of ρ, *i. e.*, its mean value when a great number of cases are treated. It is seen to vary inversely as the square root of the total number of events. Thus if 10,000 events are subjected to the Fourier analysis, the expectancy for the coefficient ρ is 0.0177, and the probability that ρ is greater than the expectancy $e^{-\frac{\pi}{4}}$ is 0.456. The probability that ρ is greater than k times the expectancy is found from (10) to be $e^{-\frac{\pi k^2}{4}}$

To facilitate the application of our result to individual cases, I have calculated in Table I the function $e^{-\frac{\pi k^2}{4}}$ for different values of k,

TABLE I.

k	$e^{-\frac{\pi k^2}{4}}$	k	$e^{-\frac{\pi k^2}{4}}$
0.1	0.9922	1.4	0.2145
0.2	0.9691	1.6	0.1339
0.3	0.9318	1.8	0.07850
0.4	0.8819	2.0	0.04321
0.5	0.8217	2.5	0.00738
0.6	0.7537	3.0	0.0008514
0.7	0.6806	3.5	6.631×10^{-5}
0.8	0.6049	4.0	3.487×10^{-6}
0.9	0.5293	4.5	1.238×10^{-7}
1.0	0.4559	5.0	2.967×10^{-9}
1.2	0.3227		

The use of the table is as follows: Find the coefficients of Fourier's series; a_n, b_n being two corresponding coefficients and a_0 the constant term, calculate $\rho = \frac{\sqrt{a_n^2 + b_n^2}}{a_0}$; next calculate the expectancy (ϵ) for ρ, from the formula $\epsilon = \frac{1.77}{\sqrt{n}}$ where n is the total number of observations, and form the ratio $k = \frac{\rho}{\epsilon}$. The above table will then give the probability that the quantity ρ is still greater than the one found.

Thus, if for 10,000 observations the quantity ρ is found to be 0.035, k would be 2, and we should find that in one case of 23 a still higher value would be obtained for ρ, if the events take place at random. Such a value for ρ would not justify us therefore to consider a real periodicity as proved, although we might be encouraged to continue the investigation by taking an increased number of events into account. If, on the other hand, the quantity ρ is equal to 3 or 4 times the expectancy, the table shows that there is a reasonable ground for supposing the events not to be distributed at random.

Our equations will also allow us to fix beforehand the number of events we must take into account in order to discover a periodic effect of a given magnitude. This is best illustrated by an example.

Let us wish, for instance, to discover whether there is a lunar period in magnetic storms. What is the number of magnetic storms we must take into account in the calculations, if the periodic term is of such magnitude that the number of magnetic storms occurring within a certain time near the maximum should bear the ratio $1 + \lambda$ to the number of storms occurring on the average during the same time throughout the lunation? To be reasonably certain of the effect, the fraction λ should be equal to at least three times the expectancy calculated on the supposition of an arbitrary distribution; putting therefore $\lambda = 3\sqrt{\frac{\pi}{n}}$ we find for the number of storms required

$$n = \frac{q\,\pi}{\lambda^2} = \frac{28.3}{\lambda^2}.$$

Thus if λ is to be one per cent, n must be at least 28,000. This shows how futile it is to attempt to discover small periodic effects unless a great quantity of material is at our disposal.

4. Occurrence of Events in Groups. The expectancy of amplitude may be increased considerably if the events do not take place at random, but are apt to occur in groups. Thus, for instance, if any one wanting to study a small periodic variation in the number of sunspots, were to count each spot as a separate "event," the average amplitudes of the periodic series would be found considerably in excess of our calculated expectancy on account of the tendency of sunspots to form groups. The following reflection will show this to be the case. It is clear that if, in our previous deductions, we consider each event to be entered in the tabular arrangement (2) as m instead of as 1, the quantity we have called ρ would not be altered, while the total number of events would be m times as much as before, and the expectancy obtained from (8) would be reduced in the ratio $\sqrt{m} : 1$, which is not the correct value. We may, however, generalize our equations, so as to be applicable to this case. If the events occur in groups of m, the probability that the quantity ρ lies between ρ and $\rho + d\rho$ becomes equal to

$$\frac{n}{2m} e^{-\frac{n\rho^2}{4m}} \rho \, d\rho$$

and the expectancy becomes

$$1.77\sqrt{\frac{m}{n}}\ .$$

More generally still, if there are n_1 groups of m_1 events, n_2 groups of m_2 events, etc., the total number being N, all quantities n_1, n_2, etc., being large, the probability that the coefficient lies between ρ and $\rho + d\rho$ is

$$\frac{N^2}{2(n_1 m_1^2 + n_2 m_2^2 + \dots)} e^{-\frac{N^2 \rho^2}{4(n_1 m_1^2 + n_2 m_2^2 + \dots)}} \rho \, d\rho$$

and the expectancy is

$$\frac{\sqrt{\pi(n_1 m_1^2 + n_2 m_2^2 + \dots)}}{N} . \qquad (11)$$

5. PERIODICITIES IN THE DAILY VALUES OF FLUCTUATING QUANTITIES. We have so far considered periodicities which may appear in the occurrence of detached events, each event being considered as of equal magnitude. But there is another class of periodicity which requires a somewhat different treatment. I mean a periodicity in *magnitude* of a quantity which recurs at equal intervals, the former case being a periodicity of *occurrence* of quantities of equal magnitude. Thus we may wish to investigate lunar periodicities in the daily average of barometric pressure, or of the daily mean of magnetic declination. The calculation of Fourier's coefficients is carried out exactly as in the former case. In the tabular form (1), the quantities t_1, t_2, denote now the daily values which it is intended to analyze, and equations (2) hold as before. Writing again $r_1 = \sqrt{a_1^2 + b_1^2}$, we deduce from (2)

$$\tfrac{1}{2} p r_1 = \left\{ (T_1 \cos\theta_1 + T_2 \cos 2\theta + \dots)^2 + (T_1 \sin\theta_1 + T_2 \sin 2\theta + \dots)^2 \right\}^{\frac{1}{2}}.$$

The quantity on the right-hand side is the sum of vectors T_1, T_2, etc., acting in directions which form angles θ_1, $2\theta_1$, etc., with some fixed direction, and remembering that T_1, T_2, etc., is each made up of a sum of quantities t_1, t_2, etc., the problem to be solved is equivalent to the following:

A number of vectors of varying magnitudes act in p fixed directions, forming angles equal to $\frac{2\pi}{p}$ with each other. The number of vectors in each direction being equal to s, what is the probability that the resultant exceeds any given value R? It will be sufficient to consider the case that the probability of positive and negative vectors is the same in all directions. If, further, $\frac{2\pi}{p}$ is a submultiple of a right angle, the result may be written down at once from Lord

Rayleigh's investigation. If there be n_1 vectors of magnitude a_1, n_2 vectors of magnitude a_2, etc., the probability that the resultant vector has a magnitude intermediate between R and $R + dR$ is

$$\frac{2}{n_1 a_1^2 + n_2 a_2^2 + \ldots} e^{-\frac{R^2}{n_1 a_1^2 + n_2 a_2^2 + \ldots}} R\, dR \quad .$$

It follows that the probability that r_1 has a value intermediate between ρ and $\rho + d\rho$ is

$$\frac{p^2}{2\,(n_1 a_1^2 + n_2 a_2^2 + \ldots)} e^{-\frac{p^2 \rho^2}{4\,(n_1 a_1^2 + n_2 a_2^2 + \ldots)}} \rho\, d\rho.$$

The expectancy of r_1 is $\frac{1}{p}\sqrt{\pi\,(n_1 a_1^2 + n_2 a_2^2 + \ldots)}$

and the expectancy of r_1^2 is $\frac{4}{p^2}(n_1 a_1^2 + n_2 a_2^2 + \ldots)$.

If the vectors are distributed according to the law of errors, so that the number which have a value intermediate between β and $\beta + d\beta$ is $\frac{2hN}{\sqrt{\pi}} e^{-h^2\beta^2} d\beta$, N being the total number of vectors and h a constant, it follows that

$$n_1 a_1^2 + n_2 a_2^2 + \ldots = \frac{2hN}{\sqrt{\pi}} \int_0^\infty x^2 e^{-h^2 x^2} dx = \frac{N}{2h^2} \; .$$

Hence the probability that the coefficients in Fourier's series have a value intermediate between ρ and $\rho + d\rho$ becomes

$$\frac{p^2 h^2}{N} e^{-\frac{p^2 h^2 \rho^2}{2N}} \rho\, d\rho \; .$$

The expectancy of the coefficients becomes

$$\frac{1}{ph}\sqrt{\frac{N\pi}{2}}$$

and the expectancy of the square of the coefficients

$$\frac{2N}{p^2 h^2}.$$

It has been assumed that Fourier's analysis has been applied to the quantities T in (1), but if, as is more rational, we treat directly the quantities t, we must write N for p in the preceding results.

The results of this paragraph are summed up as follows: Let a number of quantities $t_1, t_2 \ldots t_n$ be treated by Fourier's analysis, the quantities all being independent of each other and distributed according to the law of errors, so that the probability that any quantity t has a value intermediate between β and $\beta + d\beta$ is

$$\frac{2h}{\sqrt{\pi}} e^{-h^2 \beta^2} d\beta .$$

The probability that any coefficient of the periodic series has a value intermediate between ρ and $\rho + d\rho$ will then be

$$N h^2 e^{-\frac{N}{2} h^2 \rho^2} \rho \, d\rho .$$

The probability that the coefficient exceeds ρ is

$$e^{-\frac{N}{2} h^2 \rho^2}$$

and the expectancy for the coefficient and its square is respectively

$$\frac{1}{h}\sqrt{\frac{\pi}{2N}} \quad \text{and} \quad \frac{2}{N h^2} .$$

If the law which regulates the distribution of magnitude of the quantities t is not the law of error, the preceding results will also give the proper values for the expectancy, provided we substitute

$$N h^2 = 2\,(n_1 a_1^2 + n_2 a_2^2 + \ldots) .$$

6. LIMITATION OF THE PRECEDING RESULTS. The results of the preceding paragraph are deduced on the supposition that the values of the fluctuating quantity on successive days are quite independent of each other. This is seldom the case. If we were to take quantities, like the average height of the barometer during 24 hours on successive days, and were to investigate possible periodicities of these daily averages, our previous results could not be applied, for the barometric pressure on any one day is not independent of the pressure on the previous day, a high barometer being more likely to be followed by a high than by a low barometer. The effect of such regularities must be taken into account, and their effect will generally be to diminish the amplitudes of the shorter periods. When there is no connection between the individual quantities, all coefficients of Fourier s series are equally probable, but any regularity will favor certain periods as against others If we draw a curve at random on a sheet of paper, we cannot assign any definite value to the probability that a coefficent of the Fourier

series should exceed a given value, unless we take account of the particular bias of a person, which determines the average slope he gives to the lines. If, on the other hand, we were to rule a number of closely adjacent vertical lines, and place a point at random on each, the continuity being destroyed, the results obtained in the preceding paragraphs will hold, because the successive points of the curve are now independent. Some regularities nearly always exist, even if they do not appear at first sight; and it is of the greatest importance to be clear as to their effect whenever periodicities are to be looked for. Let us take as an example the case of sunspots, and admit, for the sake of argument, that there is no regularity at all in their distribution; that, for instance, a certain number of sunspots appear on the average every year, but that their appearance on a particular day is purely regulated by the laws of chance, and that the life of all sunspots is the same. The latter fact introduces a regularity. If the number of sunspots appearing on successive days were analyzed by Fourier's series, the period which is equal to the life of a sunspot would disappear, and shorter periods would all be reduced in magnitude, but not to an equal amount, so that the result might show periodicities which are caused by the fact that all spots have the same length of life. If, as is the case in reality, the lives of sunspots are not the same, yet follow some law of distribution round an average value, investigations on sunspot periodicities are affected in as far as the periods approximately equal to the average life are reduced in amplitude, and that by contrast, therefore, periods which are decidedly longer will seem to be increased.

7. Optical Analogy. Regularities like those discussed in the preceding paragraph will have the effect that the expectancy of the values of the Fourier coefficients depends to some extent on the period; but there will not in general be any well-defined maxima *for particular periods*, unless there is some definite periodic cause. The problem, which, so far, has only been treated by the laws of probability, must now be approached from a different point of view. Let $f(t)$ be any variable function of the time, and consider the integrals

$$A = \int_{t_1}^{t_1+T} f(t) \cos kt\, dt \ , \quad B = \int_{t_1}^{t_1+T} f(t) \sin kt\, dt \ . \qquad (12)$$

The quantity $R = \sqrt{A^2 + B^2}$ will depend on the values of k, t_1, and T; but supposing t_1 is altered while k and T remain the same, the values of R calculated for a great many values of t_1 chosen at random, will in all cases which we are now considering fluctuate about some mean value R'. At any rate, we may exclude from our discussion any case in which this is not true. This value of R' may depend on k, and the problem with which we are concerned consists in determining in what way it does depend on k, and particularly whether there are any well-defined maxima for certain periods. It will be seen that Fourier's analysis here serves the same purpose as the prismatic analysis of a luminous disturbance. The irregular fluctuations of light give continuous spectra, and complete irregularity means equal amplitude for all periods. "Lines" or "bands" are produced by greater or smaller regularity in the luminous disturbance. This optical analogy is a very important one, and we may translate some well-known optical theorems into useful propositions concerning the general analysis of fluctuating quantities. If, for instance, in optics we are dealing with a "double line"—*i. e.*, a superposition of two nearly equal periodicities —we know that the separation of the lines depends on "resolving power." The resolving power is proportional to the quantity T in the above equations, and just as a spectoscope of low resolving power is insufficient to separate two lines which are close together, so shall we be unable to distinguish between two periodicities of different frequencies, unless the time limits are sufficiently extended. In optics we seldom use a resolving power lower than that required to separate the two sodium lines. To accomplish such a separation the quantity T must include 1,000 periods; that is to say, in the case of a 26 day period we should have to take into account a series of observations extending over not less than 70 years.

It is easily seen that the number of lines on a grating determines the optical resolving power exactly in the same way as the number of periods taken into the account in investigations like the above.

8. THE PERIODOGRAM. It is convenient to have a word for some representation of a variable quantity which shall correspond to the "spectrum" of a luminous radiation. I propose the word *periodogram*, and define it more particularly in the following way Let

$$\tfrac{1}{2}\,T a = \int_{t_1}^{t_1+T} f(t)\cos kt\,dt\,, \quad \tfrac{1}{2}\,T b = \int_{t_1}^{t_1+T} f(t)\sin kt\,dt\,. \qquad (13)$$

where T may for convenience be chosen to be equal to some integer multiple of $\frac{2\pi}{k}$, and plot a curve with $\frac{2\pi}{k}$ as abscissæ and $r=\sqrt{a^2+b^2}$ as ordinates; this curve, or, better, the space between this curve and the axis of abscissæ, represents the periodogram of $f(t)$. A few examples may be given in illustration. The periodogram of the sound emitted by an organ-pipe or a violin string consists of a series of equidistant "lines." A noise would be represented by a periodogram showing a broad band. The periodogram of sunspots would show a "band" in the neighborhood of a period of eleven years, while the periodogram of tides would have a line coincident with the lunar month. The periodogram of temperature has long lines for the year and the day, and shorter lines for their submultiples.

The periodogram as defined by the equation (13) will in general show an irregular outline, and also depend on the value of t_1. In the optical analysis of light we are helped by the fact that the eye only receives the impression of the average of a great number of adjacent periods, and also the average, as regards time, of the intensity of radiation of any particular period. If the value of r in the periodogram shows maxima, this may be due to accidental circumstances, and we must find the easiest methods of separating the accidental from the real periodicities.

9. Separation of Accidental from Real Periodicities. If we were to follow the optical analogy we should have to vary the time t_1 in equations (12) continuously and take the average value of r obtained in this way for each value of k. By repeating the process for different values of k we should ultimately be able to decide whether there is any real periodicity; but this would involve an almost prohibitive labor. The following considerations simplify the investigation. Give to t_1 the successive values, o, T, $2T$, etc., up to $n\,T$, and call the corresponding values of a, b, r: a_1, b_1, r_1, a_2, b_2, r_2, etc. The quantities r may now be taken to be vectors having components a and b, any angle θ defined by $\tan\theta=\frac{b}{a}$ will be equally probable for all vectors, if there is no real periodicity and it the value of T is chosen sufficiently large. This last condition is rendered necessary by the regularities alluded to in § 6. If, for instance, we were to investigate barometric heights, and T were to be taken equal to one day, while $\frac{2\pi}{k}$ when chosen equal to one month, successive values of $\frac{b_1}{a_1}$, $\frac{b_2}{a_2}$, would have a tendency to be nearly equal

and not altogether independent of each other, on account of the persistent states of high or low barometers. But we have no reason to suspect any connection between the barometric heights after a time interval of, say, one year, and if T is therefore put equal to one year, the independence of successive values of θ would be secured. This being the case, we can apply results already obtained, for the magnitude of r will in all cases depend on some law of probability. If therefore n represents a very large number, and we form the vector $R = \sqrt{A^2 + B^2}$ from the equations

$$\frac{n}{2} TA = \int_0^{nT} f(t) \cos k t \, dt \, , \quad \frac{n}{2} TB = \int_0^{nT} f(t) \sin k t \, dt \, , \quad (14)$$

we may consider R to be the resultant of all the vectors $r_1, r_2, \ldots r_n$, and if amongst the latter there are n_1, of magnitude a_1, n_2 of magnitude a_2 .... the expectancy for R and R^2, according to § 5, is

$$\sqrt{\frac{\pi}{4(n_1 a_1^2 + n_2 a_2^2 + \ldots)}} \quad \text{and} \quad \frac{1}{n_1 a_1^2 + n_2 a_2^2 + \ldots} .$$

But the quantities n_1, n_2, etc., will vary proportionally to n. As the law according to which A and B varies, with increasing values of n, must be the same as that of the variations of R, we have the following two important propositions:

1) If $f(t)$ is a function of t which fluctuates about some mean value in an irregular fashion, the integrals

$$\int_0^T f(t) \cos k t \, dt \quad \text{and} \quad \int_0^T f(t) \sin k t \, dt$$

will with increasing values of T fluctuate about some average value which increases as $\sqrt{T}$.

2) Taking $R = \sqrt{A^2 + B^2}$ when A and B are defined by (14), and writing R' for the mean value of R, if different values of t_1 are taken, the probability that any particular value exceeds λR is $e^{-\frac{\pi\lambda^2}{4}}$. This last result follows from the investigation in § 5.

The condition under which these results hold is that the values of $f(t)$ and $f(t + T)$ are entirely independent, where, however, T may be as large as we please. If there is a true period $\frac{2\pi}{k}$ contained in $f(t)$, this condition does not hold. By writing $f(t) = \cos k t$, it is easily seen that the above integrals will fluctuate about some average value which increases as T instead of as $\sqrt{T}$, and we have, therefore, here a criterion to decide between accidental and real periodicities. To decide between the two, it would be necessary to form the integrals

$$\int_0^T f(t) \cos k t \, d t, \int_T^{2T} f(t) \cos k t \, d t, \int_{2T}^{3T} f(t) \cos k t \, d t \quad . \qquad (15)$$

and by successive additions calculate the values of

$$\int_0^T f(t) \cos k t \, d t, \int_0^{2T} f(t) \cos k t \, d t, \text{ etc. up to } \int_0^{nT} f(t) \cos k t \, d t.$$

If these integrals increase on the whole proportionally to $\sqrt{n}$, it would show that the successive values of (15) are wholly independent of each other; but if there is a more rapid increase, it would tend to show that $f(t)$ contains some true periodicity $\frac{2\pi}{k}$.

There is another method of securing the same object. It has been shown that if we form the periodogram as defined in the § 8 by calculating

$$r = \frac{2}{T} \left\{ \left(\int_{t_1}^{t_1+T} f(t) \cos k t \, d t \right)^2 + \left(\int_{t_1}^{t_1+T} f(t) \sin k t \, d t \right)^2 \right\}^{\frac{1}{2}}$$

for different adjacent values of k, the quantities r will fluctuate about some mean value r' so that the probability of r being greater than $\lambda r'$ is $e^{-\frac{\pi \lambda^2}{4}}$, the condition being that there is an equal probability for all values of k within the range considered. The chances that r is greater than four times its mean value are exceedingly small, as shown by table (1); and if the periodogram shows a sudden elevation at any point corresponding to a particullar value of k which is greater than 4 times its average value, we may conclude with reasonable certainty that $f(t)$ contains a periodic term, having a period $\frac{2\pi}{k}$. The second method, although not so direct as the first, will be more easy to apply when we are looking for variations, the periodic times of which are not accurately known. We must in any case include different values of k into our calculations, and we need not extend the time limits as much as would be necessary if we were to apply the first method. It must be noted that the values of k should not be taken to lie too close to each other, as otherwise the values of r would not be independent, for the integrals

$$\int_0^T f(t) \cos k t \, d t \text{ and } \int_0^T f(t) \cos k' t \, d t$$

will not differ much from each other if $k T$ differs from $k' T$ by less than about 45°. To secure complete independence, it will be better to let $(k' - k) t$ be as much as 90°.

10. EXAMPLES. We possess a number of investigations on hidden periodicities which allow us to apply the second test explained in the last paragraph. Professor Balfour Stewart[1] has published some calculations made in conjunction with Mr. William Dodgson on periodical variations supposed to be common to solar and terrestrial phenomena. Their method consisted in finding, by means of a neat and well chosen system of calculation, a numerical value for the inequality of 27 closely adjacent periods between 23.5 and 24.5 days. Table II gives their result for the temperature ranges at three stations during 16 years. By temperature range is meant the difference between the daily indications of the maximum and minimum thermometers. The numbers are not exactly the coefficients of the corresponding term in the Fourier expansion, but are approximately proportional to them, and for our present purpose may be taken to represent the ordinates of the periodogram. A glance at the table will show that the distribution of the figures is very much what would be expected on the theory of chance. The mean ordinates found from the table are, 3639, 3740, and 3117, and the maximum ordinates are equal to 1.6, 1.5, and 1.7 times the mean ordinates respectively, while the minima are equal to 0.38, 0.48, and 0.46 times the mean ordinate. Reference to Table I will show that there is nothing in these figures to indicate any true periodicity, as, on the theory of chance, one case out of every 13 should give an amplitude more than 1.8 times the mean amplitude, and one in every 8 one smaller than 0.4 times the mean one. The other tables given in the same paper show similar variations, and if we look at the results obtained by the authors, keeping in mind the variability of the inequalities which may be expected by the rules of chance, we come to the conclusion that there is no evidence either in the temperature range or in the declination range of any periods in the neighborhood of 24 days.

[1] *Proceedings Royal Society* XXIX (1879), p. 303.

TABLE II.

Exact period in days	Magnitude of inequality Kew	Utrecht	Toronto
23·5400	2100	1922	2934
23·5729	3093	3080	1422
23·6057	4700	3950	2252
23·6386	4025	3980	4446
23·6715	1386	3030	3939
23·7043	1887	2624	5246
23·7372	3910	2166	3638
23·7700	3915	1780	2622
23·8029	3140	3992	2148
23·8357	2771	4540	3337
23·8686	4234	4578	3422
23·9014	5921	4624	2906
23·9343	5518	3878	3098
23·9671	2374	2572	2772
24·0000	3912	2586	3428
24·0329	5135	3958	2863
24·0657	4516	2984	1678
24·0986	2157	3394	3902
24·1314	2378	5392	3216
24·1643	3795	5690	3360
24·1971	3926	3350	4274
24·2300	3043	2520	2728
24·2628	2520	4342	2377
24·2957	3004	5802	3258
24·3285	4302	5572	3601
24·3614	4761	5146	2400
24·3943	5824	3832	2906
Mean:	3639	3740	3117

As a second example I take Unterweger's[1] attempt to prove variations in sunspot activity having periods of 28, $30\frac{1}{5}$, and 36 days. The process employed is similar to that of Balfour Stewart. Twenty different periods, called trial periods, varying between 24 and 37 days are taken, and their amplitudes are found to be as follows: 8.4, 12.5, 7.3, 9.8, 12.6, 13.3, 17.1, 6.9, 12.2, 6.0, 17.4, 19.1, 13.9, 15.5, 8.0, 11.6, 15.9, 12.6, 20.8, 12.8.

It is argued that the highest amplitudes, 17.1, 19.1, and 20.8, stand out sufficiently above the rest to give evidence in favor of a true periodicity; but as the mean of the above number is 12.7, and as by the probabilities an amplitude equal to more than twice the mean ought to occur in about one case out of every 23, it is seen that the figures are just such as we should expect by the laws of chance.

11. LENGTH OF RECORD NECESSARY TO ESTABLISH PERIODICITIES. It follows from Table II that, if fluctuations are of a

[1] *Denkschrift d. math.-na. urw. Classe d. kais. Akad. d. Wissenschaften* (Wien), Vol. LVIII.

purely accidental character, the ordinate of the periodogram would only once in 300,000 cases rise to four times its mean value. The abscissæ, as has been pointed out, must be taken sufficiently far apart for the ordinates to be independent of each other. If we adopt the limits given at the end of § 9, it would follow that two periods T' and T, for which the amplitudes are calculated, should be sufficiently far apart to satisfy the equation

$$\frac{1}{\tau'} \sim \frac{1}{\tau} > \frac{1}{4T}$$

where T is the whole time. If, for instance, the observations taken during one year are treated, and the periods surrounding 26 days are considered, it is found that the difference between two periods should amount to at least 0.54 days, if the amplitudes found are intended to be independent of each other. There may, of course, be other reasons for taking the periods nearer together. If a periodicity having an amplitude b is to be separated from amongst other irregular variations, it would follow that b must be at least equal to four times the mean amplitude to afford reasonable security against deception by accidental circumstances. As the mean height of the periodogram has been shown to vary inversely as the square root of the time space considered, we have the following rule for separating accidental and real periodicities:

If the record of a number (n) of days has been subjected to analysis by Fourier's theorem, and the mean amplitude of the periodogram is found to be a, the number of days (N) required to establish with reasonable certainly a true periodicity of amplitude b is

$$N = \frac{16\, a^2 n}{b^2}$$

If a probability of one in a thousand is considered a sufficient guard against accidental periodicities, the number N may be reduced by half.

12. Spurious Periodicities. It can not be too often insisted upon that whenever Fourier's theorem is applied to finite intervals of time, the resulting periodic series gives correct values only within that interval. In consequence, the analytical calculation may give periodicities not inherent in the function $f(t)$ at all, but due to the discontinuities at the limits. Those familiar with the theory of optical instruments will be aware of the fact, that when a homogeneous vibration is examined by means of a spectroscope, the prin-

cipal line has a number of companions on either side. These companions are assigned to diffraction effects; but any one unacquainted with the theory of undulations might be misled to believe that the luminous body emitted light which is not altogether homogeneous. We may call such periodicities "spurious." They are due to the above-mentioned discontinuity at the limits, and may occur in all problems in which Fourier's analysis is applied. In order to illustrate the bearing of this on the examination of hidden periodicities, I will take a strictly periodic function $\cos q t$, and show that when, analyzed by Fourier's theorem within a finite range, it will, in addition to the true period $\frac{2\pi}{q}$, show certain other "spurious" periods. In order to examine the amplitude of a possible period $\frac{2\pi}{k}$ in $\cos q t$, we calculate the value of $r = \sqrt{a^2 + b^2}$ where

$$\frac{1}{2} T a = \int_0^T \cos q t \cos k t \, dt$$

$$\frac{1}{2} T b = \int_0^T \cos q t \sin k t \, dt \; .$$

If the time T includes n periods equal to $\frac{2\pi}{k}$ it follows that

$$\frac{1}{2} T a = \frac{2q}{q^2 - k^2} \sin \alpha \cos \alpha \qquad (16)$$

$$\frac{1}{2} T b = \frac{2k}{k^2 - q^2} \sin^2 \alpha \qquad (17)$$

when α is written for $\pi n \frac{q-k}{k}$.

Hence $r = \frac{2}{q+k} \frac{\sin \alpha}{\alpha} (q^2 \cos^2 \alpha + k^2 \sin^2 \alpha)^{\frac{1}{2}}$.

The value of r is small except when q and k are nearly equal, and in that case we may with sufficient accuracy write

$$r = \frac{\sin \alpha}{\alpha} \; .$$

As r has several maxima besides the principal one for which $\alpha = o$, *i. e.* $q = k$, we have here something exactly analogous to the diffraction images in spectroscopes. The maxima of amplitude take place when $\tan \alpha = \alpha$. At the first maximum, which is the only one which need be considered, $a = 1.43\,\pi$. If τ is the time of the true periodicity, τ' that of the most important spurious period, it is found by substitution for the value of a in terms of τ and τ' that

$$\tau' = \tau\left(1 \pm \frac{1.43}{n}\right).$$

Thus, for instance, if we were to discuss a year's record of tidal observations in order to see what periodicities there are, the known period of 29 53 days would give spurious periods, the principal ones of which are obtained by putting in the above equations $n = 12$, because a year contains nearly 12 complete periods. We should thus find these spurious periods to have lengths of 26.10 and 32.96 days. As periods of about 26 days are habitually put down to solar rotation, we might be misled to believe in an influence of solar rotation on tides. The spurious periods are easily recognized by the fact that they depend on the time space included in the calculations, and they approach the true period more and moie as that time space is extended. If in equations (16) and (17) k and q are nearly equal, we obtain as a first approximation $\frac{b}{a} = -\tan a$. Hence, if in that case the value of $\cos q\,t$ is expressed in terms of Fourier's series between the limits $t = o$ and $t = \frac{2\,\pi\,n}{k}$, the first periodic term is represented by

$$\frac{\sin a}{a}\cos(k\,t + a) \qquad (18)$$

where $a = \pi\,n\,\frac{q - k}{k}$.

13. THE "SMOOTHING PROCESS." A few words should be said on the common practice of "smoothing down" an irregular series of numbers before submitting them to periodic analysis. This is done by forming a new series, taking successive and overlapping means of, say, 4 or 5 numbers. The process is only justified if the second series is so regular that the periodicities which were hidden in the original series now become obvious. But whenever this is not the case, so that Fourier's analysis has to be applied, the labor spent in the process is wasted, and its effect often very deceptive. In order to determine the result of smoothing on the coefficients of Fourier's series, let us begin by taking a periodic function $\cos k\,t$. The process of taking overlapping means is equivalent to substituting for $\cos k\,t$ an expression formed from it by

taking at each time t the average value of the function in the interval $t-\tau$ to $t+\tau$, when τ is a constant. This average value is given by:

$$\frac{1}{2\tau}\int_{t-\tau}^{t+\tau}\cos kt\,dt=\frac{1}{k\tau}\sin k\tau\cos kt\ .$$

The result is again a periodic function, but with an amplitude diminished, in the ratio $\frac{\sin k\tau}{k\tau}$.

In the general case, where the function $y=f(t)$ need not be periodic, we may substitute by Fourier's theorem,

$$f(t)=\frac{1}{\pi}\int_0^\infty dk\ \int_0^T f(\lambda)\cos k(t-\lambda)\,d\lambda$$

where the limits o and T have been chosen, so as to make the problem correspond to the actual process used. We now form a new variable

$$y'=\int_{t-\tau}^{t+\tau}f(t)\,dt\ .$$

Performing the integration, we find

$$y'=\frac{1}{\pi}\int_0^\infty\frac{\sin k\tau}{k\tau'}\,dk\ \int_0^T f(\lambda)\cos k(t-\lambda)\,d\lambda\ . \qquad (20)$$

This equation is approximate only owing to the fact that when t is smaller than τ or greater than $T-\tau$, the integration involves values of t for which the equation (19) does not hold; but if T is large compared to τ, the error introduced is negligible. The result shows that the periodogram is reduced everywhere in the ratio $\frac{\sin k\tau}{k\tau}$, the period being $\frac{2\pi}{k}$. The process of smoothing, therefore, has completely destroyed periods equal to $k\tau$ or submultiples thereof. This, no doubt, was the object of those who employed it; but they do not seem to have noticed that the other coefficients are also affected in a manner which might easily lead to a belief in imaginary periodicities. To show this by an example, take the case that a 26 day period is looked for and the material treated as explained in § 1, after taking overlapping means of 5 successive numbers. If there is no true period, the expectancy for the coefficients of Fourier's series is the same, and therefore no regularity is to be expected in the numbers which we have called T. But the process of smoothing reduces the expectancy of the first coefficients in the ratio of 0.94, and the others successively in the ratios 0.77, 0.54, 0.27, 0.04. The consequence is that, a fictitious appearance of regularity might be introduced into the numbers T, prominence being given to the periods of 26 days as compared with that of its submultiples. There is little doubt that the regularities artificially introduced by the smoothing process have been the cause of frequent mistakes.

14. ELIMINATION OF SECULAR VARIATIONS. Very considerable labor has sometimes been spent in eliminating secular variations and other known periodicities before the hidden periodicities are searched for. We may reasonably ask the question, what object is thereby gained? It is one of the great advantages of Fourier's analysis that each of its terms is independent of the others; and if we wish to determine any particular coefficient, it is unnecessary to begin by eliminating the others. The analysis itself performs that process in the best possible way, if the coefficients are obtained by arithmetical calculations. In some cases, however, when mechanical processes are employed, it may be better to get rid of *known* variations before the unknown ones are searched for; and this is particularly the case if the former are large compared to the latter. The best method of procedure must be settled in an individual case; but the general rule may be given, that it is the best to eliminate as few variations as possible, and to carry out the elimination at as late a stage of the computation as possible. Known variations may be got rid of at the end by expressing them separately in a periodic series. Thus a uniform change, such as is often assumed in the case of secular variations of terrestrial magnetism, may be expressed by $-\frac{\beta t}{T}$ where β is the change taking place in the time T. Expressed in a periodic series between the limits of time $t = 0$ and $t = T$ we have

$$-\frac{\beta t}{\tau} = -\frac{\beta}{2} + \frac{\beta}{\pi}\left\{\sin kt + \frac{1}{2}\sin 2kt + \frac{1}{3}\sin 3kt + \ldots\right\}$$

where for shortness k is written for $\frac{2\pi}{\tau}$.

If, therefore, the uncorrected figures for, say, the magnetic declination give a series

$$\begin{aligned} a_0 + a_1 \cos kt + \quad & a_2 \cos 2kt + \quad \ldots\ldots \\ + b_1 \sin kt + \quad & b_2 \sin 2kt + \quad \ldots\ldots \end{aligned}$$

we correct for the effect of secular variation by leaving the *a* coefficients as they stand, and subtracting $\frac{\beta}{n\pi}$ from the *b* coefficients. This method of treating the problem has not only the advantage of greater simplicity, but it gives us also a clearer idea as to magnitudes and uncertainties of the corrections we apply.

15. MEANING OF THE TERM "PERIOD." Some confusion has arisen owing to a certain vagueness in the use of the term "period." If a quantity varies according to the symbolical expression $\cos kt$, it is generally agreed to call $\frac{2\pi}{k}$ the "period" of the variation. Strictly speaking of course, the quantity is periodic, not only in a time $\frac{2\pi}{k}$ but also in a time $\frac{4\pi}{k}$, $\frac{6\pi}{k}$, etc.; yet no one would call the multiples of $\frac{2\pi}{k}$ the "period" of $\cos kt$. In investigations of more complicated periodicities the term is, on the contrary, often applied in a loose way, and we meet with statements affirming, for instance, a periodicity of 26 days, "the variable having two maxima and two minima within the range of that period." This ought to be called a 13 day period as distinguished from a 26 day period. The matter is, I think, o greater importance than might at first sight appear. In a complicated subject a clear nomenclature helps towards clear ideas. I think, therefore, it would be well to apply the term diurnal "period" solely to a periodic change which goes through one cycle in 24 hours, and to distinguish it, therefore, from the semi-diurnal or ter-diurnal periods. If we wish to have a name for the complete change including all periods which are submultiples of the principal one, it would be better to use a more general term such as diurnal "oscillation" or diurnal "variation."

16. DISTINCTION BETWEEN CAUSE AND EFFECT. The confusion alluded to in the last paragraph has, like others in this subject, arisen from an insufficient distinction between the analytical representation of a certain variable in terms of a periodic series and the causes, which may perhaps quite indirectly have produced the periodicities. The apparent cause of the tides, for instance, is the revolution of the moon in one lunar day; but the forces which cause the tides have a period of half a lunar day, and this is the period of the tides. No one confuses the time of revolution of the disk of a siren with the note given out by it, and similarly we should draw a clear distinction between the time of revolution of the sun or

moon and the periodic times of any variables which possible may be due to solar or lunar rotation.

I am at present only concerned with the best methods of discovering periodicities, and not in finding their causes; but it may be worth while to lay stress on the fact that periodicities are sometimes produced by a combination of circumstances, and that when we have discovered a periodic effect, it is not necessary to ascribe it to something which "revolves" in the time of the period which has been found. Thus periods approximately equal to those ascribed to solar rotation may be produced by a combination of an annual and a monthly period. Let an effect depend, for instance, on the moon's declination (δ), in such a way that it is proportional to cos δ, but is also dependent on the sun's position with respect to the equator, and may be analytically expressed by

$$(a + b \cos D) \cos \delta$$

when a and b are constant, and D is the sun's declination. A simple trigonometrical transformation changes the expression to

$$a \cos \delta + \tfrac{1}{2} b [\cos (\delta + D) + \cos (\delta - D)] ,$$

and the last two terms represent periods of $2\pi\left(\frac{1}{d} \pm \frac{1}{D}\right)$. If t and T denote the length of the lunar day and solar year respectively, the length of the periods into which the whole effect resolves itself contains terms having periodic times θ given by $\frac{1}{\theta} = \frac{1}{t} \pm \frac{1}{T}$. Substituting $t = 27.3$, $T = 365.2$; θ becomes equal to 25.4 and 29.5 days respectively. If the lunar effect is a fortnightly one, the smaller value for θ would be 13.16, or half of 26.32. A period of this kind might easily be mistaken for one due to solar rotation.

17. THE 26 DAY PERIOD. This period has already been alluded to. As it would be a matter of some importance to establish its reality, we may illustrate some of the results obtained by a short reference to the principal researches on the subject, amongst which Hornstein's papers deservedly take the first place. In an investigation published in 1871[1] Hornstein analyzes the records of the magnetic elements. He groups, for instance, the daily values of the declination at Prague in 1870 in the manner explained in § 1. Taking 15 trial periods, the results are collected in a table which is here reproduced (Table III). The first column gives the number of days

[1] *Wiener Ber.* LXIV, p. 62 (1871).

in the period chosen, and the second column gives the representation of the value of declination, neglecting all terms in Fourier's series except the first.

TABLE III

Period	Declination
16 days	$12^\circ\ 1'.11 + 0'.03\ \sin(x + \ldots)$
17	$1\ .09 + 0\ .12\ \sin(x + \ldots)$
18	$1\ .11 + 0\ .07\ \sin(x + \ldots)$
19	$1\ .07 + 0\ .19\ \sin(x + \ldots)$
20	$1\ .10 + 0\ .12\ \sin(x + \ldots)$
21	$1\ .13 + 0\ .19\ \sin(x + \ldots)$
22	$1\ .12 + 0\ .13\ \sin(x + \ldots)$
23	$1\ .09 + 0\ .10\ \sin(x + \ldots)$
24	$1\ .16 + 0\ .10\ \sin(x + \quad)$
25	$1\ .20 + 0\ .172\ \sin(x + 3^\circ\ 22')$
25.5	$1\ .16 + 0\ .336\ \sin(x + 70^\circ\ 0')$
26	$1\ .17 + 0\ .616\ \sin(x + 123^\circ\ 4')$
26.5	$1\ .23 + 0\ .696\ \sin(x + 174^\circ\ 18')$
27	$1\ .21 + 0\ .660\ \sin(x + 217^\circ\ 40')$
28	$1\ .24 + 0'.281\ \sin(x + 326^\circ\ 18')$

If the amplitudes in the table are examined, it is found that it is on the average $0'.12$, while the period 26.5 days gives a value which is more than 5 times as great. According to Table I, there is here a very strong evidence that this periodicity is not due to mere accident, and a further confirmation may be found in the gradual change of phase as the trial periods gradually increase from 25 to 28 days. For, according to (18), the expression for the first term of Fourier's series for a trial period $\frac{2\pi}{k}$ is

$$\frac{\sin a}{a} \cos(k t + a)$$

if the true period is $\frac{2\pi}{q}$ and $a = \pi n \frac{q-k}{k}$ where n is the total number of periods included in the time interval. In the present case $n = 14$. Putting $\frac{2\pi}{k} = 26.5 \quad \frac{2\pi}{q} = 25$, we find $a = 151^\circ$ and $\frac{\sin a}{a} = 0.16$, while the difference in phase in Hornstein's table which should be equal to a is 171° and the ratio of amplitude 0.24. Considering the superposition of accidental variations, these numbers are in good agreement. But a closer examination somewhat weakens the argument. In order to see how far lunar effects might have something to do with the periodicity found, I have extended Hornstein's calculations to the periods of 29, 30, and 31 days. I find for the amplitudes of the first terms of Fourier's series 0.105, 0.189, and 0.157, which is decidedly higher than the numbers given by Hornstein's for the periods below 24 days. I was then struck by the fact that the amplitudes given by Hornstein were not, when casually examined, borne out by his material.

I recalculated, therefore, some of the coefficients, and found for the amplitude of the periods of 26, 26.5, and 27 days the values 0′.54, 0′.61, and 0′.58, which are decidedly smaller than those given by Hornstein, while for the period of 21 days I obtain a value .295, which is decidedly larger than that of Hornstein. Hornstein states that he has obtained the amplitudes of all periods up to 24 days inclusive by a "graphical process," and it would seem, therefore, that the process must have given too small values. The corrected numbers raise the ordinates of the mean periodogram, and weaken considerably the evidence in favor of a true periodicity.

It should be said that an absolute check of Hornstein's calculations is not possible, because he does not state how he has eliminated the secular variation, and there is some indirect evidence that he has unknowingly strengthened the 26.5 days period by his treatment of it. This only confirms what has been pointed out in § 14, that it would be better to eliminate such variations *after* Fourier's analysis has been applied. We may pass more quickly over the remainder of Hornstein's paper. The declination at Vienna during the same year shows a variation for a period of 26 days, which is little more than double that found for a period of 24 days, and no certain conclusions can be based on so slight a preponderance. The results for the inclination are equally undecisive, while those for the horizontal intensity give a mean ampitude of 7.7 units of the fourth decimal place, while the greatest among 15 amplitudes is 17.0. Here the results are, therefore, entirely such as we may expect to be due to accidental variations.

In a subsequent paper Hornstein endeavors to prove the existence of a 26 day period in the daily variation of barometric pressure, but I must express my opinion that he has failed to establish his point. His method of procedure consists in forming the series of numbers T (see (1) § 1), the amplitude of the diurnal period having been grouped in periods of 24, 25, 26, 27, and 28 days. Instead, however, of applying Fourier's analysis to the numbers T, he sums up the figures irrespective of sign, and thus obtains what he considers to be a measure for the amplitude of the period. The figures found are: 222, 298, 490, 245, 226, for the five periods respectively, the largest number belonging to the 26 day period. Even if we could accept Hornstein's method of deducing the amplitude, the figures would not prove much; for the largest of them is less than twice as great as the average. But a closer inspection shows that the 26 day series gives a large value, because it contains the highest and lowest numbers, viz., + 85 and —58; but these appear on *two successive days*. If, therefore, Fourier's analysis were applied, we should get a large value for the amplitude, not of the 26 day period but of the two day period, and it is practically certain that the close juxtaposition of such a high and low value can only be due to accident. Hornstein specially remarks that Fourier's series, when applied to his numbers, would be misleading; but there is no reason for this statement beyond the fact that Fourier's series would not support the 26 day period. I have calculated the first coefficients of the series, and find them to be equal to 146, 176, 109, 155, 132, for the five periods respectively, so that the 26 day period now gives the smallest instead of the largest value.

Hornstein's work was soon followed up by Liznar, Müller, and others, who adopted very much the same method of procedure, taking trial periods of 24, 25, 26, 27, and 28 days, and calculating the amplitude of the first term of Fourier's series. Table IV gives a summary of the principal results obtained.

Periods	I	II	III	IV	V	VI	VII	VIII	IX
24	0.0812	0.0463	0.4404	2682	1724	0.0401	0.0476	24.48	5.85
25	0.1001	0.2042	0.9627	2873	6173	0.2101	0.1196	36.23	36.39
26	0.2262	0.3258	0.5491	3295	3239	0.2327	0.1416	44.22	41.27
27	0.1920	0.1678	0.9822	4278	2393	0.1295	0.1006	31.54	28.35
28	0.0528	0.1620	0.3100	2080	1288	0.1007	0.0445	15.52	10.64

The numbers given refer to the amplitudes where the periods are those stated in the first column. The variables in the different vertical columns are as follows:

I. The amplitude of daily variation of declination at Vienna (1882—1884), calculated by taking the difference of the observation at 2 p. m. and 8 a. m. on each day.[1]

II. The same for Kremsmünster.

III, IV, V. The daily variations according to Liznar of declination, horizontal intensity and vertical intensity at St. Petersburg.

VI, VII. The easterly and westerly disturbances at Vienna.[2]

VIII, IX. The disturbances of horizontal and vertical intensity at St. Petersburg.[3]

[1] LIZNAR. *Wiener Ber.* Vol. 94, p. 834 (1887).
[2] LIZNAR. *Wiener Ber.* Vol. 91, p. 474 (1885).
[3] MUELLER. *Bulletin of the Akademie of St. Petersburg.* 1886.

An inspection of the numbers leads to the following conclusions:

1. Each column by itself is not sufficient to prove the existence of a 26 day period, the ratio of the greatest to the smallest amplitudes being in no case greater than one might expect from the theory of chance.

2. There is, however, the significiant fact that in six out of the nine columns the greatest amplitude falls on the 26 day period, and in no case does it fall either on the 24 or the 28 day period. It is difficult to believe that this is due to accident.

As regards the length of the most probable period which seems indicated in the above tables, Adolph Schmidt[1] calculates it to be 25.87 days.

One of the most striking arguments in favor of a periodicity connected with the above is that derived from von Bezold's calculations on the occurrence of thunderstorms. The tables given do not allow us to apply the results of the previous pages; but the fact that, according to von Bezold, no trial period has given him amplitude similar in magnitude to that of 25.84 days, together with the similarity in the numbers obtained separately from two different and independent time intervals, renders it unlikely that the results are due to mere chance. But it should be understood that it is really the first submultiple of the 25.84 day period; *i. e.*, a period of 12.92 days, which gives the exceptionally large amplitude.

Prof. Frank H. Bigelow[2] has discussed a supposed connection between solar rotation and meteorological phenomena, in a series of papers.

The period he adopts is 26.68 days, which differs materially from that arrived at by Liznar, Müller, and von Bezold, but more nearly agrees with that deduced by Hornstein for the magnetic declination. Unfortunately, Professor Bigelow does not, as far as I know, give anywhere sufficient details to allow us to apply our methods of testing their reality. The curves he gives in support of his views would, however, imply that it is the fourth or fifth submultiple of his period, rather than the period itself, which gives the largest effect. The general result of a critical examination of the published investigations on the 26 day period leads me to think that, although the magnetic elements and the occurrence of thunderstorms seem to be affected by a period of 26 days and of its

1 *Wiener Ber.* XCVI, p. 989 (1887).

2 *Meteorological Journal*, September, 1893; and other publications.

first submultiple, the subject requires a good deal of further study before we can be sure as to the exact nature of the period. Even though it may be considered as proved, it must not be necessarily assumed that it is due to solar action.

If it was a question merely of magnetic disturbances, there does not seem to be any great improbability, however, that some periodicity may be connected with the sun's rotation about its axis, especially at times of great sunspot activity. Groups of spots have been observed to persist for several rotations. If a large group is likely to be accompanied by a magnetic disturbance, that disturbance may easily be repeated after a complete revolution of the sun. The result of such an action would not be a "homogeneous" period, or a "line" in the periodogram as we have called it, but rather a broad band having its center at a period coincident with the average period of revolution of a sunspot. It would seem therefore that the most promising line of investigation would be to determine the shape of the mean periodogram taking account of a sufficiently long time interval. I am at present engaged in treating the Greenwich observations of magnetic declination from this point of view.

18. Conclusion. The importance of calculating the mean periodogram has been pointed out in the last section, and, quite independently of any research as to particular periodicity, I believe that great interest attaches to it in many meteorological phenomena. The periodogram, for instance, of the changes of barometric pressure would seem to me likely to give important information. It has been shown that if the height of the barometer on one day were perfectly independent of that on the previous day, all periods would be equally probable, and the mean periodogram would be a straight line. In virtue of the persistence in the duration of high and low barometers, the mean periodogram will show maxima and minima, and nothing is known as to their position. It must be of interest to find out whether different localities show any marked differences in the periodogram, and it is almost certain that places which lie near a track along which frequent cyclones are passing will show characteristic differences in the periodogram. Similarly periodograms of temperature are likely to prove of importance.