On the Bartels Technique for Time-Series Analysis, and Its Relation to the Analysis of Variance

L. R. Hafstad

Journal of the American Statistical Association, Vol. 35, No. 210, Part 1. (Jun., 1940), pp. 347-361.

ON THE BARTELS TECHNIQUE FOR TIME-SERIES ANALYSIS, AND ITS RELATION TO THE ANALYSIS OF VARIANCE

By L. R. Hafstad
Department of Terrestrial Magnetism
Carnegie Institution of Washington

I. INTRODUCTION

It has long been recognized that there is difficulty in applying the usual statistical techniques when there is correlation between successive items in any given sample. This trouble arises in many different fields: in economic time series the index number for any one day, month, or year is dependent on that of the preceding time unit. In sociology the number of strikes in progress at any one time is dependent on the number at some immediately preceding time. In agricultural experiments the yield of any plot will be high or low with the fertility of the adjacent region. This characteristic is known by various names, among statisticians as serial correlation, among physicists as *nachwirkung* or persistence. Various devices are used by statisticians to circumvent it, but the problem is by no means solved. For this reason it may not be amiss to bring to the attention of statisticians an approximate intuitive device used on similar problems by physicists, with the hope not only that the method in its present form might prove of some value to statisticians, but also that they may be stimulated into a more critical study of the technique.

II. CORRELATION OF TIME SERIES

Simple examples showing the difficulties introduced by the application of formulas derived for random samples to data involving persistance are the correlations between time series. In a well-known paper Yule[1] gives as an example a correlation between Church of England marriages and standard mortality, for which a coefficient of $+0.9512$ with the very small standard error of 0.0140 was obtained. He points out that in spite of its "high significance" ($P \ll 0.01$) such a result in this case is obviously nonsense, and shows that in the series in question there exists a high degree of positive serial correlation, that is, successive observations are not independent but are related in such a way that high values are likely to be followed by other high values. He does not however suggest any way to obtain a valid measure of the significance of a correlation coefficient between series of this kind, con-

[1] U. Yule, *Journal Royal Statistical Society*, vol. 89, pp. 1–69, 1926.

cluding only that, in a series "in which successive terms are closely related to one another, the usual conceptions to which we are accustomed fail totally and entirely to apply."

The above example was deliberately chosen to develop a case of *reductio ad absurdum,* but in a serious work, *The Social Aspects of the Business Cycle,* we find a similar example. For the correlation coefficient between prosperity and mortality Thomas[2] is surprised to find the positive value +0.31 with a standard error of only 0.12 making the result "significant" ($P=0.01$) according to the usual rules. Such cases have led to the absurd situation in which the statistical tests for significance are often ignored and, as quoted by Mitchell,[3] " . . . an expert . . . (does) not consider of much value a correlation coefficient below 0.90."

To statisticians of the sociological-economic school the above situation apparently remains as an unexplained anomaly, for as recently as 1935 in discussing the correlation by Thomas above, Bartlett[4] suggests only that the result obtained must be due mainly to a particular coincidence with the influenza epidemic, and concludes that "If neither series is random no valid test . . . (for the statistical significance of correlations) . . . can be recommended."

Seeking a more general explanation of the above anomalous situation we recall that the formula for the correlation coefficient is derived on the assumption of independence of successive observations. To test our data for independence we may choose to use the fundamental proposition that if a variate is distributed with standard deviation $\sigma(1)$, then the mean of a random sample of h such variates is distributed with standard deviation $\sigma(h)=\sigma(1)/\sqrt{h}$. That is, for independent observations the ratio

$$d_h = h\sigma^2(h)/\sigma^2(1) \tag{1}$$

should be statistically constant and equal to unity if enough observations are used.

Taking successively $h=3, 5, 9$ for the marriage series in Yule's data we get the curve d_h in fig. 1. Obviously the requirement that $d_h=$unity is not satisfied, so the data are not independent. Similarly for Thomas' data we get for $h=5, 10, 15$ the curve d_h in fig. 2. Again it is clear that the data are not independent, and the usual formula for the correlation coefficient should not be used.

Thus, we are confronted with the problem of estimating in some way

[2] D. H. Thomas, *The Social Aspects of the Business Cycle* (C. P. Dutton, 1925).

[3] W. C. Mitchell, *Business Cycles* (National Bureau of Economic Research, 1927), p. 270.

[4] M. S. Bartlett, *Journal Royal Statistical Society,* vol. 98, pp. 536–543, 1935.

the number of independent components among any number of dependent ordinates.

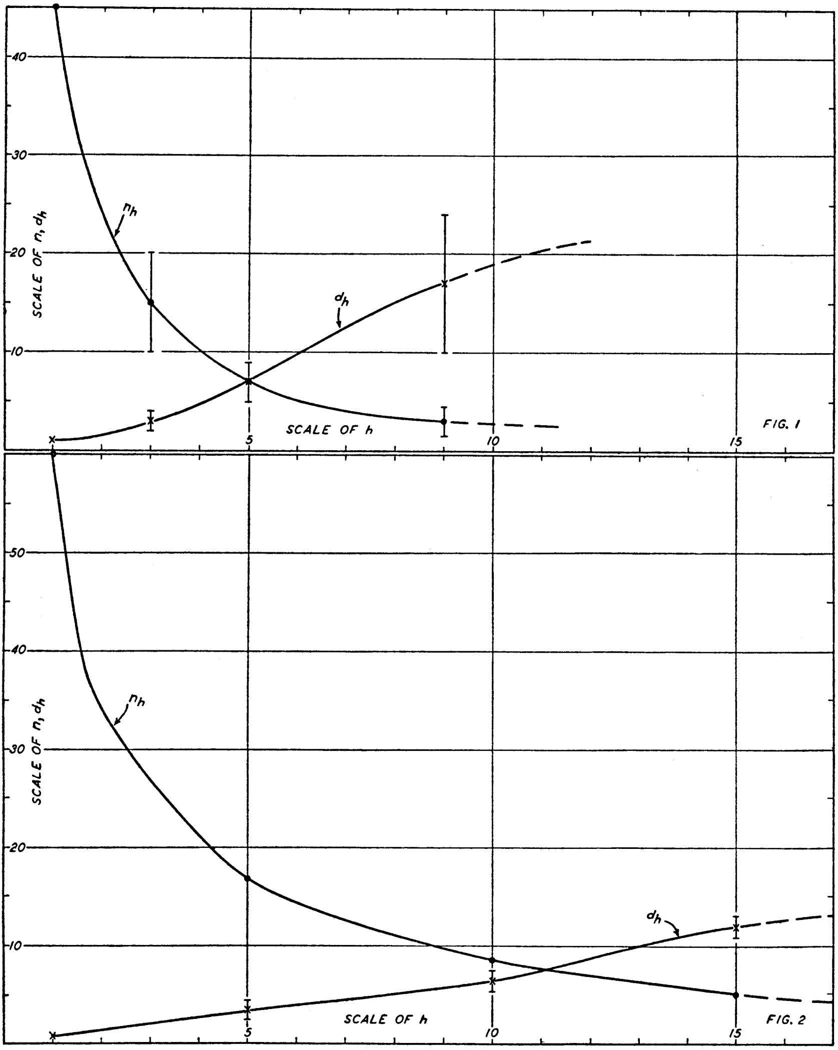

Fig. 1 (Upper panel).—Number of effectively independent observations in Yule's example.
Fig. 2 (Lower panel).—Number of effectively independent observations in Thomas' data.

III. THE BARTELS TECHNIQUE

A rigorous solution to the problem is difficult, but an approximation is available and is much used by physicists, who are convinced that "it is useless to labor over the farthings when the pounds are uncertain." The technique referred to was developed by Bartels[5] in the study of quasi-persistent periodicities, and can be explained as follows.

TABLE 1
ORDINATES OF A TIME SERIES

Year	Nine-Year Groups of Years						
	I	II	III	IV	V	VI	VII
0	133	133	116	125	94	99	102
1	94	136	101	109	92	99	119
2	79	155	97	131	85	110	94
3	83	95	102	126	77	132	101
4	61	59	104	110	83	100	100
5	77	61	98	81	108	80	101
6	93	83	107	92	87	58	107
7	97	106	88	90	84	38	108
8	120	121	128	93	88	49	104

Consider a set of observations like those in table 1. We may test these data for independence, as above, by means of the formula for d_h in eq. (1). The result is the curve in fig. 3. Again the requirement

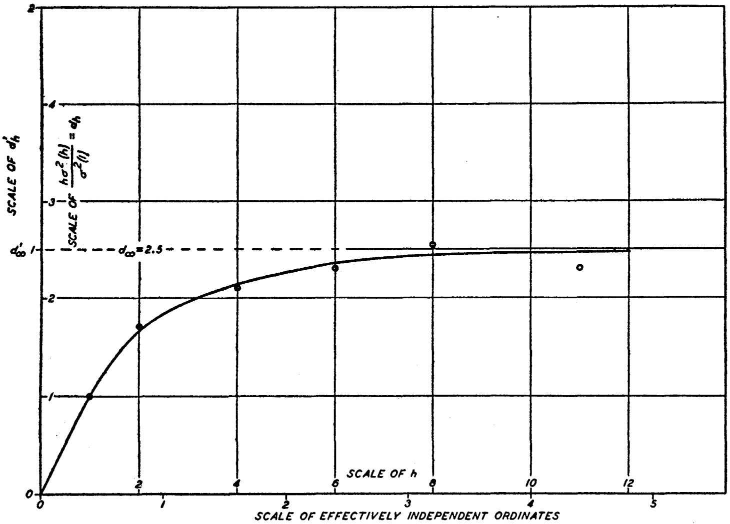

Fig. 3.—Number of effectively independent observations in Greenstein's data.

that $d_h = 1$ is not satisfied and the data are not independent. However, a peculiar feature of the curve may be noted. Above about $h = 5$ the value of the ratio is nearly constant and equal to (say) $d_\infty = 2.5$ in the

[5] J. Bartels, (a) Thesis (Göttingen, 1922); (b) *Zur Morphologie Geophysikalischer Zeitfunktionen*, Sitzungsberichte. Preuss. Akad. Wissensch., Phys.-Math. Klasse, pp. 504–522, 1935; (c) *Terrestrial Magnetism*, vol. 40, pp. 1–60, 1935; (d) "Verbogene periodische Erscheinungen" in Lubberger's *Wahrscheinlichkeiten und Schwankungen* (J. Springer, Berlin, 1937), pp. 66–73.

present case, so that for values H of h sufficiently large, the quantity $(H/d_\infty)\sigma^2(H)/\sigma^2(1)$ is unity. Comparing this with eq. (1) we see that $\sigma(H)$ is related to $\sigma(1)$ *as if* the averages of H successive ordinates were derived from H/d_∞ independent components. For this reason we call H/d_∞ the number of "effectively independent components among H successive ordinates." This process in effect calibrates our hitherto unknown scale of effectively independent components by means of the "calibration point" d_∞.

Bartels[5d] illustrates the meaning of the parameter d_h as follows. Consider a series of values $S_1 = a_1, a_2, a_3, a_4, \cdots$ taken at random from an infinite supply of values with standard deviation $\sigma(1)$. Divide this series into sets of h successive values and form averages for each set; the standard deviation for these averages should be statistically $\sigma(h) = \sigma(1)/\sqrt{h}$ if the original members of the sequence S_1 are independent. Now form a new series S_d from S_1 by repeating each value d times; for instance, with $d=4$, S_4 would be $a_1, a_1, a_1, a_1, a_2, a_2, a_2, a_2, \cdots$. The series S_d will still have the standard deviation $\sigma(1)$, but if it is divided into successive groups of $h = dh'$ (h' any integer)[6] the standard deviation for the average of these groups will be

$$\sigma_d(h) = \sigma(1)\sqrt{h'} = \sigma(1)\sqrt{(d/h)}. \tag{2}$$

Conversely, if only $\sigma_d(h)$ and $\sigma_d(1)$ were known for the series S_d, the number d of identical ordinates could be found by means of the formula

$$d = h\sigma^2(h)/\sigma^2(1). \tag{3}$$

For actual time series no rigorous treatment is possible, but so far as they may be considered analogous to the series S_d, differing only in the degree of correlation between successive items, the same considerations apply. This analogy is the justification for the use of the terms chosen to describe d_h and h/d_h in the general case.

The probable error in any estimate of the d_h of eq. (1) may be approximately determined as follows. The standard deviation in any estimate of $\sigma(1)$ made from m independent observations will be $\sigma(1)/\sqrt{(2m)}$, while that in the estimate $\sigma(h)$ will be $\sigma(h)/\sqrt{(2m/h)}$.

[6] For the case when h' is not an integer, $d=4$ and $h=3$, for instance, the groups would cut up the sequences of identical ordinates and the equation for d_h would not hold. This defect, which could be examined in detail, does not detract, however, from the illustrative value of the analogy, especially because its effect vanishes if h is large compared with d, which should be true in any case for this technique to be strictly applicable. Of course the value of d_h may depend on the way in which the groups of h intervals are divided; for instance, whether we begin with the values 1 to h, or 2 to $h+1$, etc. Such cases occur in geophysics and the possibility must always be kept in mind. This effect in itself may sometimes deserve special attention but in order to avoid secondary complexities, let us consider $\sigma(h)$ to be derived from all averages, obtained from all possible combinations, of h successive ordinates.

Now for a quotient x/y, we know that when x and y are independent, the standard errors are related by the equation

$$[\sigma_{x/y}/(\bar{x}/\bar{y})]^2 \doteq (\sigma_x/\bar{x})^2 + (\sigma_y/\bar{y})^2 \tag{4}$$

as can be found in any treatment of the propagation of error. So here' the mean percentage error in d_h will be $\sqrt{[a^2+b^2]}$, where a and b are the mean percentage errors in the numerator and denominator. For large values of h the mean error in an estimate of $\sigma(1)$ will be negligible in comparison with the mean error in the estimate of $\sigma(h)$, wherefore the mean percentage error in d_h will be practically that in $\sigma^2(h)$. For small values of h the mean percentage errors in the estimates of $\sigma^2(h)$ and $\sigma^2(1)$ may be nearly equal, in which event the above argument does not apply. In this circumstance we may set an upper limit to the error in d_h by assuming the errors in the estimates of $\sigma^2(h)$ and $\sigma^2(1)$ to be always operating in the same direction, giving $\sqrt{[a^2+b^2]} = \sqrt{[a^2+a^2]} = 2a$, or twice the error in either one alone.

The error in d_∞ of course depends on the method of extrapolation. Its error must therefore be judged from the behavior of curves such as that given in fig. 1 where a constant value of d_∞ will be attained only if the effect producing the persistence is constant. However, these approximations are adequate for the case in hand, for the quantity $h\sigma^2(h)/\sigma^2(1)$ is itself used as a correction and its error must be a second order effect.

IV. APPLICATION

Adopting Bartels' concept of "effectively independent components" we may now reconsider the examples discussed in section II. In Yule's data the total number N of items may be divided into groups of h items each. Within each of these groups there are h/d_∞ "effectively independent components" (section III) giving

$$n_h = (N/h)(h/d_\infty) = N/d_\infty \tag{5}$$

as an estimate of the total number of independent components on which the correlation coefficient might be based.[7] From fig. 1, though no accurate estimate of d_∞ may be made due to the large error in extrapolating d_h, we may safely conclude that $d_\infty > 15$ and that $n \not> 3$, so that instead of being based on the 45 recorded items, the correlation coefficient was actually based on only three independent components, whence applying the usual t test we find the significance level at $P = 0.2$, wherefore the correlation coefficient no longer need be considered significant.

[7] Of course more rigid tests would consider both series, which may have different degrees of persistence. It would be interesting to consider whether the number of effectively independent *pairs* of values might be obtained from the equivalent numbers of independent components in the two series.

Similarly, for Thomas' example, we find from fig. 2 for the mortality series that there are less than five independent pairs of observations so that the "significance" does not even reach the level of $P=0.6$, and we are no longer obliged to find excuses for the result.

Thus, bearing in mind only the fact that the standard error of a correlation coefficient is determined by the number of *effectively independent* observations, we see that Bartels' technique provides us with a safeguard such that the correlation coefficient becomes harmless, even in inexperienced hands. Furthermore we find that, contrary to the conclusions of Yule, the ordinary concepts of statistics do apply and the usual tests for significance remain valid.

V. THE RELATION BETWEEN BARTELS' TECHNIQUE AND THE ANALYSIS OF VARIANCE

In the above examples Bartels' technique was applied to series of numbers that are related in time. Since the mathematical operations apply only to the numbers, being quite independent of what the numbers represent, the same technique may be applied to series of numbers related in space or in any other natural way.

A typical example is given by the so-called station-year rainfall record.[8] In this case we have, let us say, 30 years of observations at each of 40 neighboring stations. Such a record has been regarded as equivalent to 1200 years of observations at a single station under the same general climatic conditions. Accordingly, in the desire to know the frequency of occurrence of storms above some specified amount of rainfall, and particularly to obtain an estimate of the reliability of this frequency, it has been customary to assume that a storm magnitude which has been reported (say) 12 times in the entire series of observations would have a frequency of occurrence of once in 100 years, with the implication that this value is fairly reliable, being based on 12 observations.

A procedure of this kind is hardly valid, for it is probable that these 12 observations were recorded during only 3 or 4 separate storms, and quite possible that all were recorded at the 12 different stations during a single storm. In fact, by applying the chi-test to such data we shall usually find that the means for the stations are much more alike than would be expected for random data. This, however, tells us only what we already suspect; namely, that a single storm may be hitting several stations simultaneously. Similarly, from the analysis of variance, we

[8] Miami Conservancy District, Engineering Staff, *Storm Rainfall of Eastern United States* (Dayton, revised ed., 1936). See however, K. C. Clarke, *Transactions American Geophysical Union*, 1938. See also H. Wold, *Stationary Time Series* (Upsalla, 1938).

should conclude only that there are unexpectedly large differences between the means of the years, with correspondingly small difference between the means for the stations.

We desire, therefore, to segregate in some way the independent events, or at least to find out how many there are. We are thus again faced with the problem of estimating the number of independent components among an arbitrary number of related observations. Resorting to the Bartels' technique, we may adapt it to an $r \times s$ table of a station-year record of storms by setting the rows, or years, end to end successively and grouping the resulting series in the manner described earlier, using $h=2, 3, 4$, etc., until the ratio d_h seems to have reached a constant value. The number d thus arrived at will be an estimate of the number of stations hit by a single storm. As a matter of fact, the number d thus estimated will usually be too small, and we shall ordinarily be safe in saying that the actual number of stations affected is still larger, for if a single storm hits several stations, geographically adjacent, these stations will in general not all be entered in adjacent columns in the table, thus effectively reducing the amount of persistence, and its index, the number d. Similarly for storms at the edge of the area, the boundary effect will tend to reduce the number of stations affected by a single storm. Thus for a station-year record of storms we may safely conclude that the actual average number of stations hit by a single storm is larger than d, and the effective length of record even shorter than that given by a simple application of the Bartels method.

In the above example the correlation existed between adjacent items in the rows, but examples are easily found in which the persistence is mainly in the columns. Thus Fisher[9] shows a table of the frequencies of rain at different hours in different months, with successive hours in vertical columns. Clearly the fact that it is raining at any given time is not independent of whether it was raining one hour earlier, and persistence is to be expected. Applying the usual formula for d_h it is found that $d_{24}=13.1$, so that only about two of each days' observations are effectively independent. Cases can also be found in which persistence has been introduced into both rows and columns. For these, special methods will have to be developed.

Thus we find ourselves in a situation which may be likened to the following game. Starting with a set of random numbers such as may be taken from Tippett,[10] we may allow an assistant (or nature) to perform secretly certain operations on these numbers. The problem for the statistician is to test the resulting numbers in various ways and to de-

[9] R. A. Fisher, *Statistical Methods for Research Workers* (Blackie & Son), table 44.
[10] L. H. C. Tippett, *Tracts for Computers* No. 15 (Cambridge, 1927).

duce from his tests what operations have been performed. The analysis of variance is a first test, for it may help to tell us whether the numbers are still random. Bartels' point of view enables us to go one step further and specify in certain cases the groupings used in the unknown operations.

There are many interrelations between the Bartels' parameters and other statistical conceptions, for example, the Lexis theory of supernormal and subnormal dispersion; hence one might expect that his parameters may also be expressed in the language of the analysis of variance. This is easily demonstrated by considering the mathematical formulation of the analysis of variance as given for example by Irwin.[11] It will be found that Bartels' parameters d_{rows} and $d_{\text{cols.}}$ are essentially the ratios of the "column mean square" to the "total mean square," and the "row mean square" to the "total mean square," respectively.

Thus Bartels' parameters may be obtained from the usual analysis of variance tables, provided only that the "total mean square," usually omitted, is also computed, and it should be possible to devise tests of significance based on analysis of variance methods which can be carried to any desired degree of refinement.

VI. HARMONIC ANALYSIS

Among economists the use of harmonic analysis is viewed with a healthy skepticism, largely because by its use apparent "real" periods are too readily found. Col. M. C. Rorty,[12] for example states:

> . . . The fudamental defect in the harmonic analysis is that it will resolve any ordinary business time series into definite regular periodicities, regardless of whether any real periodicities exist or not. . . . Probably the best indirect proof of the lack of value of the harmonic analysis is to create an artificial time series by throwing dice and then to analyze this series with and without assumptions as to lag.

Among physicists, on the other hand, harmonic analysis continues to be regarded as one of the most powerful and dependable tools for the study of time series of any kind. This curious diversity of opinion seems to be founded in the fact that Rorty is speaking of the harmonic analysis of Schuster,[13] whereas physicists are now using a modern form developed mainly by Bartels. This technique has been devised for the express purpose of investigating hidden periodicities in geophysical time series, which are usually encumbered with random fluctuations,

[11] J. O. Irwin, *Proceedings Royal Statistical Society*, vol. 94, pp. 284–300, 1931.
[12] W. C. Mitchell, *Business Cycles* (National Bureau of Economic Research, 1927), p. 261.
[13] A. Schuster, *Terrestrial Magnetism*, vol. 3, pp. 13–41, 1898.

cycles of changing phase or period, and persistence (serial correlation) between successive items, like the refractory time series encountered in sociological and economic investigations. Since the publication of Bartels' paper of 1935 his technique has been the standard procedure among physicists, yet the writer has been informed by "professional statisticians" that the Bartels' methods, being based on harmonic analysis, are "old-fashioned," and "not accepted by statisticians," and (by implication) that were physicists "properly trained" they would solve their problems by the use of the analysis of variance.

That a barrier between the two groups exists is clear from even a casual inspection of the literature on time series presented by exponents of the respective schools. Both schools accept as fundamental the work of the physicist Schuster, which was in turn based on the work of the physicist Rayleigh. After this time, however, further work by the investigators of the physical school seems to be largely ignored by the statistical school and we find two separate lists of standard references covering the same subject, with little duplication of authors. The statistician apparently reads papers by Moore, Crum, Mitchell, R. A. Fisher, Roos, Snedecor, etc., and adopts Slutzky[14] as a classic reference. The physicist reads papers by von Laue, Einstein, Taylor, Stumpff, and Pollack, and adopts the paper of 1935 by Bartels as his reference text. The statistician's development is based on the mathematical treatments of probability distributions, serial correlations, etc. The physicist's approach is by analogy with problems already solved in connection with the kinetic theory of gases, diffusion problems, Brownian motion, turbulence theory, and physical optics. Perhaps it is the frequent assumption of familiarity with the analogy which tends to make the work of the physical school relatively inaccessible to the pure statisticians. Conversely the use of such terms as "graduality" and "fluency"[15] as introduced by Slutzky for instance, make the statistician's work difficult reading for the physicist.

There is another factor that probably contributes to the segregation of the two schools. It seems that writers on sociological and economic time series have convinced themselves that the methods used in physics are valid only for pure sine wave phenomena, whereas these are seldom encountered in economic series.

[14] E. Slutzky, *Econometrica*, vol. 5, pp. 105–146, 1937.

[15] "The unconnected random waves are usually called irregular zigzags. A correlation between the items of a series deprives the waves of this characteristic and introduces into their rising and falling movements an element of graduality . . . ."

"We must distinguish between the *graduality* of the transitions and their fluency. We could speak of the absence of the latter property if a state of things existed where there would be an equal probability for either a rise or fall after a rise as well as after a fall . . . (Slutzky[14].")

To the economist, apparently quite certain of the meaning of "reality," Rorty's statement[12] quoted above, for example, constitutes damning evidence; to the physicist, grown accustomed to seeing even "solid rock" represented by ψ-waves and matrix elements, it is no evidence at all. The harmonic analysis gives a mathematical model of a time series just as the Schrödinger equation gives a mathematical model of an atom. The time series behaves *as if* it were composed of harmonics just as the atom behaves *as if* it were composed of ψ-waves, and the economist has quite as much right to deal with harmonic components that may never be isolated as a physicist has to deal with an equally intangible ψ-wave or matrix element.

The point is—to put it bluntly—if a period of any conceivable shape exists, then the harmonics that it may be composed of will also exist, and a test for the "reality" of the harmonics is automatically a test for the reality of the original period. That a wave of special shape should be indistinguishable from the sum of its harmonics is a totally irrelevant question.[16] Bartels' analysis of the 27-day recurrence phenomenon[5c] in magnetic activity gives a beautiful example of a case in which no simple physical meaning can be attached to the harmonic components.

Statements such as that by Rorty are made probably not so much in protest against the *use* of harmonic analysis as against the *conclusions* which are often obtained by its use. In this he is ably and effectively seconded by Bartels, in whose hands periodogram analysis has been freed from many of the objections that previously were entirely valid. For example, regarding Schuster's methods, Slutzky[14] in 1937, states:

> . . . we must give up his criterion when we remember that among his assumptions is that of independence of successive observations.

This is true of Schuster's early work, but the technique of Bartels was developed in order to avoid just this assumption.

In Schuster's method, $C(f_1)$ is the amplitude of a sine wave of a certain frequency f_1 obtained by a harmonic analysis of the total observed time series. For obtaining an estimate of E, which Schuster called "expectancy," he used his "periodogram" and assumed E equal to an average for all frequencies f (supposed to be free of actual periods) of the amplitudes $C(f)$. Schuster based this choice of E on the assumption that E is independent of the frequency. This equipartition of the variance is true for random series in which successive values are

[16] Such an ambiguity in interpretations is a commonplace occurrence in modern physics. See, for example, the discussion of the coupled pendulum given by C. G. Darwin, *The New Conceptions of Matter* (Macmillan, 1931), p. 170.

independent, but untrue for actual time series in which successive values tend to be alike (persistence tendency), as has often been pointed out by Bartels. To avoid this difficulty, which may lead to underestimating E many fold, Bartels devised an estimate of E on the basis of harmonic analyses for the same frequency f_1 as that to be tested, in suitable subdivisions of the entire series; thus the persistence tendency affects E in much the same way as C, tending to cancel its effect in the ratio. It will be instructive to consider an actual example in order to demonstrate his method.

In table 1 a series of numbers was given, representative of an economic time series. To illustrate Bartels' approach, suppose that these represent yearly values and that we wish to test for the reality of the nine year period frequently reported in the literature of economics.

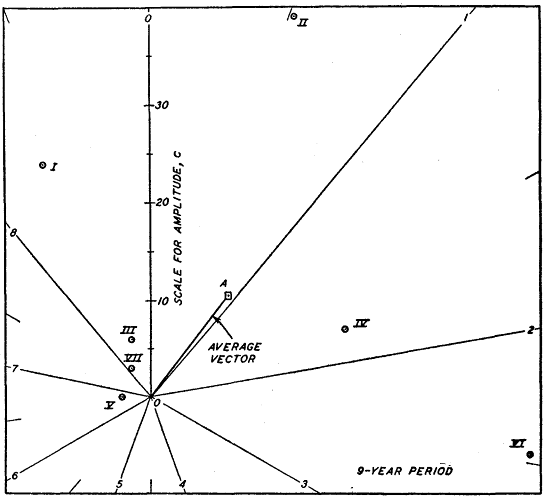

Fig. 4.—Harmonic dial for the data of Table I

The values are arranged in seven groups of nine years each; the groups are numbered I to VII, the years in each interval from 0 to 8. Each group is harmonically analyzed to yield a sine wave,

$$c \sin (t + a)$$

with time t reckoned from the middle of year 0 of group I, and increasing by 2π in nine years. The results are plotted in the "harmonic dial" of fig. 4 which replaces the usual phase diagram. The nine year sine wave for each group is represented by the points numbered I to VII, or by the vector connecting the origin 0 with these points. This vector indicates by its length the amplitude c, and by its direction the phase a which is taken as zero for the direction from zero toward the right, increasing counter-clockwise. The scale for the amplitude c is indicated; the time when each sine wave reaches its maximum can be read from the scale around the edge of the dial where the figures 0 to 8 indicate the midpoints of each nine year group. The average sine wave of a nine year period, obtained by a harmonic analysis of the total $7\times 9 = 63$ years is indicated by the "average vector" OA. Now A is the mass center of the "point cloud" formed by the seven dots for the individual years, and this property of A may be used to test whether OA can be regarded as real, or whether it can be explained by a chance occurrence. The average vector OA has an amplitude of 13.1 units, and the root-mean-square amplitude of the seven individual waves is 25.2 units. Under random conditions the expectancy for the average amplitude of the seven individual waves is $25.2/\sqrt{7} = 9.5$ units. Among several possibilities for the test under consideration Bartels chooses two:

(a) We suppose that the seven dots are chosen at random from a point cloud of a normal distribution, with root-mean-square amplitude 25.2 units and centered at O. What is the probability that the average vector OA ($=13.1$ units) exceeds its expectancy ($=9.5$ units) in the ratio $k = 13.1/9.5 = 1.38$? The answer to this question, familiar in Brownian motion and random walk problems, is[17]

$$P_a = e^{-k^2} = 1/7. \tag{6}$$

(b) We suppose that the seven dots are chosen from a point cloud centered at A with root-mean-square distance from A to be calculated in the usual way as $\sqrt{(25.2^2 - 13.1^2)} = 21.6$. The expectancy for the distance of the mass center of the seven dots from A is then $21.6/\sqrt{7}$ or 8.2. The probability that this mass center is more than $OA = 13.1$ units distant from A is with $k = 13.1/8.2$,

$$P_b = e^{-k^2} = 1/13. \tag{7}$$

Thus both of these formulations lead to probabilities for chance occurrence of the order $P = 1/10$, which is hardly sufficient to warrant the contrary assumption of reality. As pointed out by Bartels, even these estimates of P may be too favorable to the assumption of reality because the data may be affected by "quasi-persistence."

[17] This calculation is made under the assumption that 25.2 is the root-mean-square amplitude for a long series of random points, of which these seven may be a random sample. Actually the standard deviations of samples of seven random points are not constant, but fluctuate from one sample to another, and the calculation made here does not take these fluctuations into account. While this might be done the result of such a calculation could hardly be called a refinement because it would remain subject to fluctuations. What is more, the phenomenon of "quasi-persistence" may be, and probably is, present, so it seems satisfactory merely to regard P_a itself as subject to fluctuations.

Thus our task would be finished, except for the fact that applying the tests of significance available in the literature at the time he made his study, Greenstein[18] finds from the *very same data*, which give the number of business failures per 10,000 business concerns for the years 1867–1929, the probability of a chance occurrence for a 9.14 year period to be only $P=1/1790$. Furthermore, by refinements (locating the period more accurately as 9.43 years), Greenstein reduces this value to $P=1/30700$ by Schuster's method, and to 1/5520 by Fisher's.[19]

While Greenstein himself is properly skeptical of the reality of the nine year period, the mathematical result obtained cannot be dismissed lightly as one of those cases in which a wrong result unfortunately creeps into the literature. This work was done under the direction of the late Henry Schultz at Chicago, and C. F. Roos[20] has stated that this paper by Greenstein is a "lucid description" of the Schuster technique for periodogram analysis. The fact seems to be that Greenstein used the best technique generally available to economists at the time he made his study.

Now Greenstein used Fisher's formula for his significance test, an this formula is derived on the assumption of *independence of successive ordinates*. That this condition is not satisfied by Greenstein's data is at once apparent from his periodogram in fig. 5, for to an unbiased observer this curve is not so much characterized by high values near nine years as by low values below six years. This is just as is to be expected if serial correlation is present, for this effect suppresses any high frequency oscillations, since successive values tend to be alike. We may therefore suspect that Bartels' result is the more nearly correct, especially since it agrees with our common sense judgment; but it would be of interest to attempt to correct for persistence in Fisher's formula so as to bring the results into better agreement. Thus we are again confronted with the problem of estimating in some way the number of independent components among any number of related ordinates.

Returning to Greenstein's data, since we have shown in fig. 3 that $d_\infty=2.5$ for this case, we divide the total number of items (66) by $d_\infty=2.5$ to obtain the number of "effectively independent items" (26). Using this value in Fisher's formula, we get

$$P_g = \tfrac{1}{2}(26 - 2)(1 - g)^{\frac{1}{2}(26-2)} = 1/10 \tag{8}$$

and the embarrassing discrepancy between the results of Bartels' test and Fisher's test disappears. Thus both Bartels' and Fisher's formulas

[18] B. Greenstein, *Econometrica*, vol. 3, pp. 170–198, 1935.
[19] R. A. Fisher, *Proceedings Royal Society*, vol. 75, pp. 54–59, 1929.
[20] C. F. Roos, *Econometrica*, vol. 4, pp. 368–381, 1936.

give reasonable results when discreetly used. It may not be amiss however to point out that, as used by Greenstein, any desired increase in the significance of his nine year period could be obtained with Fisher's formula merely be recording monthly, daily, or hourly values of business failures over the time interval in question, for although Fisher's treatment remedied one defect in Schuster's method, he neglected to

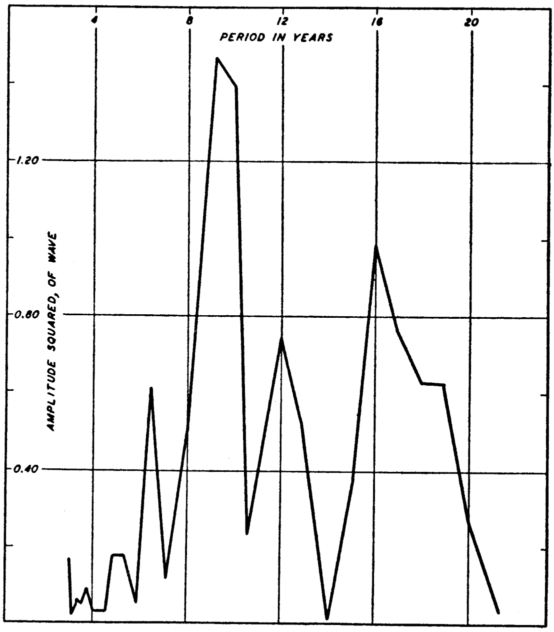

Fig. 5.—Periodogram for percentage ratios of business failures to total number of business concerns in United States, 1867–1932 (after Greenstein).

consider the overwhelming effect of a persistence tendency on the periodogram and its consequent effect on tests for reality.

It is a pleasure for the writer to acknowledge his indebtedness to Dr. W. Edwards Deming without whose continuous help and encouragement this paper would never have been written. The writer wishes also to express his gratitude to Mr. Stephan and many other friends and colleagues with whom he has discussed various aspects of the problems considered.