

# THE CYCLIC EFFECTS OF LINEAR GRADUATIONS PERSISTING IN THE DIFFERENCES OF THE GRADUATED VALUES

BY EDWARD L. DODD

*University of Texas*

**1. Scope of inquiry.** Slutsky [1] applied the moving sum, the repeated moving sum, and other linear processes to random numbers obtained from lottery drawings. But the graph of the *moving sum* becomes, when the vertical scale is changed in the ratio of  $n$  to 1, the graph of the *moving average*, the simplest form of *graduation*. When cyclic effects are studied, there is no essential difference between a moving sum and a moving average, nor between a general linear process with coefficients  $a_1, a_2, \dots, a_s$ , having sum  $A \neq 0$  and the corresponding *graduation*, with coefficients  $a'_i = a_i/A$ . Thus Slutsky's work throws considerable light upon graduation, although his main interest was in summation.

Slutsky found that the graphs of moving sums of random numbers bore strong resemblance to graphs of economic phenomena, such as [1, p. 110] that of English business cycles from 1855 to 1877. In fact, Slutsky regards the fluctuations in economic phenomena as due largely to a synthesizing of random causes.

In general the undulatory character of such values cannot be described as periodic; since the waves are of different length. But Slutsky found that, upon operating on random data having mean zero and constant variance, the resulting values approach a sinusoidal limit under certain conditions,—in particular, when a set of  $n$  summations by twos is followed by  $m$  differencings, and as  $n \rightarrow \infty$ ,  $m/n \rightarrow a$  constant. Romanovsky [2] generalized this result by taking successive summations of  $s$  consecutive elements of the data, with  $s \geq 2$ ; but required that  $m/n \rightarrow \alpha \neq 1$ . However, the cases which are of interest to me just now are those for which  $m = n - 1$  or  $m = n - 2$ ; and for these cases  $m/n \rightarrow 1$ . Romanovsky considers the case of  $m = n - 1$ ,—not, however, as leading to a sinusoidal limit,—and gives in formula (46) the value of a coefficient of correlation—which I deduce directly. From his formula (43) a corresponding coefficient of correlation can be obtained for the case of  $m = n - 2$ , as the sum of certain products. A more simple expression than this I need, which I obtain directly. In my treatment, these coefficients are the cosines of angles; and the ratio of such an angle to a whole revolution is an expected frequency of occurrence.

After setting forth in Section 2 some preliminary formulas, I treat in Section 3 the results of applying to random data an indefinite number  $k + 2$  of summations or averagings, followed by  $k$  differencings—the number of terms in a sum remaining fixed. In Section 4, however, only a few differencings are applied to a

graduation. In particular the Spencer 21-term formula is studied in some detail. In former papers [3, 4] I have dealt with the immediate effects of graduations upon random data.

The question to be considered in this paper is this: *Do the cyclic effects appearing in the graduated values persist in the successive differences? And, if so, do these affects fade out gradually or on the other hand, do they come to a rather abrupt termination?*

These differences of graduated values, indeed, up to the third, fourth or fifth are of considerable importance. Henderson [5] defines the smoothing coefficient of a given graduation as the ratio of the theoretical standard deviation of the third differences for the graduated values to that for the original values or data.

**2. Preliminary notions and formulas.** The data to be graduated will be supposed to be independent, or uncorrelated, or as Slutsky expresses it, "incoherent." This will imply that the expected value of the product of two different chance variates is the product of their expected values.

Now the operations of summing and differencing as used here are not inverse. To illustrate: Given as independent  $u, v, w, x, y, z, \dots$ . Summing by twos yields the sequence  $u + v, v + w, w + x, x + y, y + z, \dots$ . But the first differences of these numbers,  $w - u, x - v, y - w, z - x, \dots$  are alternately correlated, thus  $w - u$  is negatively correlated with  $y - w$ ;  $x - v$  with  $z - x$ , etc. Indeed, successive differencing following successive summing does not lead back to the original condition of incoherency. However, under certain conditions, the resulting coherency may be so slight that the final succession of numbers may have just about the same chaotic properties as the succession of data.

In my paper [3, p. 262], I set forth a number of features on the basis of which a cycle length could be defined. One of these involves the frequency of maxima. Given independent chance variables, each subject to the same law of distribution,

$$(1) \quad P(x_i \leq x) = \Phi(x);$$

where  $\Phi(x)$  has a derivative  $\phi(x)$ . It is then easy to see that the expected relative frequency of maxima is  $1/3$ . That is:

$$(2) \quad P(x_{i-1} \leq x_i \leq x_{i+1}) = \int_{-\infty}^{\infty} [\Phi(x)]^2 \phi(x) dx = 1/3.$$

Now, for a given feature, a cycle length is defined as the reciprocal of the theoretic relative frequency. Then the cycle length here for maxima is three. It is well known that averaging tends to remove maxima. Thus, upon averaging or summing, the cycle length tends to increase. It is almost as well known that differencing tends to increase the frequency of maxima, and thus decrease cycle length. For if  $z_i = \Delta y_i = y_{i+1} - y_i$ , then between two maxima of  $y_i$ , there is at least one minimum (strong and weak) of  $y_i$ ; and following this minimum and before passing the next maximum of  $y_i$  there is at least one maximum of  $z_i$ . Successive differencing tends to reduce the cycle length of maxima from 3 to 2,

that is to make the graph a perfect zig-zag where positive and negative values of  $z_i$  alternate. A set of differencings following a set of summings may bring the cycle length from some fairly large number back to about 3, and thus restore something like the original chaotic appearance in the graph.

In dealing with the foregoing  $\Phi(x)$  or  $\phi(x)$  in (2), it was not assumed that the distribution be normal. But, in what follows, it will be assumed that

$$(3) \quad \phi(x) = \frac{1}{\sigma(2\pi)^{1/2}} e^{-(x-\mu)^2/2\sigma^2};$$

and, for convenience,  $\mu$  will be taken as zero—that is, the data will be supposed given as deviations from their theoretic mean. Actually, the data used by Slutsky and the data I have used belong to a rectangular distribution, as noted in my former paper. Nevertheless the close agreement between actual and expected results seems to indicate [3, p. 263] that the theory is in general applicable. It is well known that averaging of observations from non-normal distributions may lead rather quickly to an approximately normal distribution.

Given  $n$  real numbers,  $a_1, a_2, \dots, a_n$ , let

$$(4) \quad y_j = a_1x_i + a_2x_{i+1} + \dots + a_nx_{i+n-1}; \quad i = 1, 2, 3, \dots$$

Then  $y_j$  is the moving *sum* if each  $a_r = 1$ . Slutsky takes  $j = i$  or  $j = i + n - 1$ . Again,  $y_j$  is the moving *average* if each  $a_r = 1/n$ . For graduation in general, the condition  $\Sigma a_r = 1$  is imposed; and usually  $j = i + (n + 1)/2$ . If  $n$  is odd,  $y_j$  is thus associated with the middle  $x$ .

Under the assumption that the  $x$ 's are independent and normally distributed about mean zero, with constant variance, I have proven [3, p. 256]: The probability that for any specified  $j$ ,  $y_{j-1} < 0$ , and  $y_j > 0$  is given by  $P = \theta/360^\circ$ , where

$$(5) \quad \cos \theta = \frac{\sum_{r=1}^{n-1} a_r a_{r+1}}{\sum_{r=1}^{n-1} a_r^2}.$$

The expected relative frequency of up-crossings of the graph of the  $y$ 's through the zero base line is then  $\theta/360^\circ$ . That is:  $\theta/360^\circ$  is the expected relative frequency of a change in the sign of  $y$  from  $-$  to  $+$ ; also, of a change in sign from  $+$  to  $-$ .

But, as  $\Delta y_j = y_{j+1} - y_j$ , it follows that

$$(6) \quad \Delta y_j = b_1x_i + b_2x_{i+1} + \dots + b_nx_{i+n-1} + b_{n+1}x_{i+n},$$

where

$$(7) \quad b_1 = -a_1, \quad b_{n+1} = a_n, \quad b_r = a_{r-1} - a_r, \quad r = 2, 3, \dots, n - 1$$

and since a maximum for the  $y$ 's at  $y_i$  occurs when  $\Delta y_{i-1} > 0$ ,  $\Delta y_i < 0$ , it follows that the theoretic frequency therefor is  $\theta'/360^\circ$ , where

$$(8) \quad \cos \theta' = \frac{\sum_{r=1}^n b_r b_{r+1}}{\sum_{r=1}^{n+1} b_r^2}.$$

In a similar manner, by using *second* differences, we get the expected relative frequency  $\theta''/360^\circ$  for inflexional points, in specified direction. Moreover,  $\theta \leq \theta' \leq \theta'' \leq \dots \leq 180^\circ$ ; since inflections must be at least as frequent as maxima, etc.

If the foregoing formulas are applied to the identical "graduation"  $y_i = x_i$ , then  $\cos \theta = 0$ ,  $\cos \theta' = -1/2$ ,  $\cos \theta'' = -2/3$ . In fact,

$$(9) \quad \cos \theta^{(t)} = -t/(t+1).$$

This follows from the fact that the  $b$ 's and similar coefficients are the binomial coefficients; and

$$(10) \quad \sum_{r=0}^t {}_t C_r^2 = {}_2t C_t; \quad \sum_{r=0}^{t-1} {}_t C_r \cdot {}_t C_{r+1} = {}_2t C_{t-1}.$$

Thus repeated differencing leads toward the perfect zig-zag. An extension of this feature will be taken up in the next section.

**3. Repeated summing and differencing.** To indicate the *result* of the summing of  $n$  consecutive numbers in a sequence, I shall use the notation  $1^n$ . And the difference  $\Delta y_i = y_{i+1} - y_i$  will be indicated by  $-1, 0^{n-1}, 1$ . Thus if  $n = 3$ ,  $1^3$  and  $-1, 0^2, 1$  will stand respectively for

$$(11) \quad y_i = x_{i-1} + x_i + x_{i+1}; \quad \Delta y_i = -x_{i-1} + 0x_i + 0x_{i+1} + x_{i+2}.$$

If, now,  $z_i = y_{i-1} + y_i + y_{i+1}$ , then

$$(12) \quad z_i = x_{i-2} + 2x_{i-1} + 3x_i + 2x_{i+1} + x_{i+2}.$$

Since  $(n)$  is often used to indicate the *operation* of summing  $n$  consecutive numbers, we may write

$$(13) \quad (3)^2 = 1, 2, 3, 2, 1; \quad (n)^2 = 1, 2, \dots, (n-1), n, (n-1), \dots, 2, 1.$$

Then, for  $n \geq 2$ ,

$$(14) \quad \Delta(n)^2 = -1^n, 1^n; \quad \Delta^2(n)^2 = 1, 0^{n-1}, -2, 0^{n-1}, 1.$$

And, since the operations of summing and differencing are commutative, we are lead to

$$(15) \quad F_n^k = (-1)^k \Delta^k(n)^k = {}_k C_0, 0^{n-1}, -{}_k C_1, 0^{n-1}, {}_k C_2, 0^{n-1}, \dots, (-1)^k {}_k C_k;$$

as may be established by induction. For from the foregoing, it follows that

$$(16) \quad (-1)^k \Delta^k(n)^{k+1} = {}_k C_0^n, -{}_k C_1^n, \dots, (-1)^k {}_k C_k^n.$$

Then, since  ${}_{k+1}C_r = {}_k C_r + {}_k C_{r-1}$ , we conclude that

$$(17) \quad F_n^{k+1} = (-1)^{k+1} (n)^{k+1} = {}_{k+1}C_0^n, 0^{n-1}, -{}_{k+1}C_1^n, 0^{n-1}, \dots, (-1)^{k+1} {}_{k+1}C_{k+1}^n.$$

If now  $n \geq 2$ , then from (5) and (15) we find that

$$(18) \quad \cos \theta = 0; \quad \theta/360^\circ = 1/4.$$

Thus, the expected frequency of the changes in sign of  $\Delta^k(n)^k$  is the same as that for the raw or ungraduated data. Moreover, if  $n \geq 3$ , (8) leads to  $\cos \theta' = -1/2$ , found for the data. For, in this case, at least two zero coefficients intervene between any two non-zero coefficients. And thus

$$(19) \quad \cos \theta' = -\sum_{r=0}^k {}_k C_r^2 / 2 \sum_{r=0}^k {}_k C_r^2 = -1/2.$$

In fact, the same factor cancels from numerator and denominator as we take higher differences, if a sufficient number of zeros intervene. More explicitly stated, the formula (9) found for the data is valid also for  $\Delta^k(n)^k$ , provided  $n \geq t + 2$ .

To make this more concrete, it may be noted that cycle lengths corresponding to  $t = 0, 1, 2, 3$ , and  $4$ , are respectively

$$(20) \quad 4, 3, 2.73, 2.60, 2.52.$$

From (15), we see directly that an element of  $\Delta^k(n)^k$  is correlated only with certain other elements which are at distances from it which are multiples of  $n$ .

Some of the foregoing results may be included in a theorem as follows: **THEOREM:** *Given a sequence of independent chance variates, each subject to the normal distribution (3) with mean zero. Upon this material, let  $k$  summings or averagings by  $n$  be performed and  $k$  differencings, in any order. Then the resulting sequence has something of the same chaotic nature as the data. In particular for  $n \geq 2$  the expected frequency of changes of sign is the same,—viz.,  $1/4$  for change from minus to plus and  $1/4$  for change from plus to minus. Moreover, as  $n$  is increased from 2 to 3, 4, 5,  $\dots$ , the expected frequency of other characteristics becomes the same, maxima and minima, points of inflection, etc., in accordance with (9).*

But, suppose now that after  $k + 1$  summings by  $n$ , only  $k$  differencings are performed. Is the resulting sequence almost chaotic? Hardly so. At least, it can be shown that changes of sign in each direction have no longer an expected frequency fixed at  $1/4$ ; but this expected frequency decreases as  $n$  increases. To show this, formula (5) is applied to (16); and setting in (10),  $C = {}_{2k}C_k$ ,  $C' = {}_{2k}C_{k-1}$  it follows that

$$(21) \quad \cos \theta = [(n - 1)C - C'] / nC = 1 - (2k + 1) / n(k + 1).$$

Then  $\cos \theta > 1 - 2/n$ ; and the cycle length for expected changes of sign in definite direction is somewhat greater than that obtained by setting  $\cos \theta = 1 - 2/n$ . For values of  $n$  not too small, we may write  $\cos \theta = 1 - \theta^2/2$ , approximately; and then approximately

$$(22) \quad \text{cycle length for definite change of sign in } \Delta^k(n)^{k+1} \text{ is } \pi\sqrt{n}.$$

If  $n = 9$ , this approximate length is 9.4, assuming  $k$  fairly large, whereas the more exact length is 9.2.

Consider now the result of summing  $k + 2$  times, and then differencing only  $k$

times. For this purpose, a few formulas for summing squares will be useful. By the method of differences it can be shown that if  $l = a + nh$ , and

$$(23) \quad T = a^2/2 + (a + h)^2 + (a + 2h)^2 + \dots + (a + \overline{n-1}h)^2 + l^2/2,$$

then

$$(24) \quad T = n(a^2 + al + l^2)/3 + (l - a)^2/6n.$$

Suppose, now, that  $a/n$  takes on the values  $0, {}_kC_0, -{}_kC_1, \dots, (-1)^k {}_kC_k$  in succession, while  $l/n$  takes on the values  ${}_kC_0, -{}_kC_1, \dots, (-1)^k {}_kC_k, 0$ . Let  $U$  be the sum of the  $(k + 1)$  values of  $T$  thus obtained. Then by (10).

$$(25) \quad U = n^3(2 {}_{2k}C_k - {}_{2k}C_{k-1})/3 + n \sum_{i=0}^{k+1} {}_{k+1}C_i^2/6.$$

$$(26) \quad U = \frac{n^3 (k + 2)(2k)!}{3 k!(k + 1)!} + \frac{n}{6} {}_{2k+2}C_{k+1}.$$

Now, by applying to (16) one more summation by  $n$ , there are formed  $(k + 2)$  arithmetic progressions of  $(n + 1)$  terms each, alternately increasing and decreasing. The maximum and minimum terms at the juncture of the progressions are to be split into two halves to apply (23). Then the sum of the squares of these coefficients is given by (26). This forms a denominator for (5).

To obtain the numerator for (5) we note that from  $ab = [a^2 + b^2 - (a - b)^2]/2$  it follows that if

$$(27) \quad V = a(a + h) + (a + h)(a + 2h) + \dots + (a + \overline{n-1}h)(a + nh);$$

then, from (23),

$$(28) \quad V = T - nh^2/3 = T - (l - a)^2/3n.$$

If now  $W$  is the sum of such  $V$ 's, reference to the last terms of (24) and (26) shows that

$$(29) \quad W = U - (n/3) {}_{2k+2}C_{k+1}.$$

And hence, from (5),

$$(30) \quad \cos \theta = \frac{(k + 2)n^2 - 4k - 2}{(k + 2)n^2 + 2k + 1}.$$

Then

$$(31) \quad \cos \theta > \frac{n^2 - 4}{n^2 + 2};$$

but only slightly greater when  $k$  is large. Again

$$(32) \quad \cos \theta > 1 - 6/n^2;$$

but only slightly greater when  $n$  is not small. In this case,  $\cos \theta = 1 - \theta^2/2$ , approximately. And thus, approximately, for large  $k$ , and for  $n$  not small

$$(33) \quad \text{cycle length for definite change of sign of } \Delta^k(n)^{k+2} = 1.81n.$$

This gives for  $n = 10$  a cycle length of 18.1; whereas, if  $\cos \theta$  is taken as the right member of (31), the cycle length is 18.2.

Thus, if a  $(k + 2)$ -fold summation or averaging of random data is followed

by only  $k$  differencings, the resulting graduation or linear processing  $z = \Delta^k(n)^{k+2}$  is decidedly not as chaotic as the data; as seen from (31) and (33). But further,  $\Delta z = \Delta^{k+1}(n)^{k+2}$ ; and thus from (22) the cycle length for the expected maxima of  $z$  is about  $\pi\sqrt{n}$ .

Now Slutsky [1, p. 109] distinguished conspicuous waves from inconsequential "ripples." On this basis, the frequency of significant cyclical features for a chance variable, such as  $z$ , would be less than the frequency of the maxima. It is not so clear that the frequency of significant features of a chance variable will be *greater* than that for changes of sign in definite direction. That turned out to be true for graduated values such as discussed in my earlier paper [3, p. 262]. If this be also valid for  $z$ , we would expect that conspicuous "waves" of  $\Delta^k(n)^{k+2}$  would have average length between  $\pi\sqrt{n}$  and  $1.81n$ , except for small values of  $n$  and  $k$ .

**4. Graduations or linear processes and their successive differences.** If double summation by  $n$  is followed by a single differencing, the result—as indicated in (14)—is, for  $n = 3$ ,

$$(34) \quad y_i = -x_i - x_{i+1} - x_{i+2} + x_{i+3} + x_{i+4} + x_{i+5}.$$

Then

$$(35) \quad y_{i+3} = -x_{i+3} - x_{i+4} - x_{i+5} + x_{i+6} + x_{i+7} + x_{i+8}.$$

Thus  $y_i$  and  $y_{i+3}$  are negatively correlated; since  $x_{i+3}$ ,  $x_{i+4}$ , and  $x_{i+5}$  appear in each, but with sign changed. This would seem to tend to make maxima alternate with minima at distances of about 3; or at distances of  $n$ , in the general case (14). Here, following Slutsky and Romanovsky, the coefficient of correlation  $r_p$  between elements at a distance of  $p$  is taken as

$$(36) \quad r_p = E(x_r \cdot x_{r+p}) / E(x_r)^2.$$

Using computed averages, instead of expected values, Alter [6] recommends a "correlation periodogram," in which  $r_p$  is the ordinate for abscissa  $p$ .

Moreover, we would expect a graduation (4) with coefficients  $a_i$  proportional to the ordinates  $y$  of the sinusoid  $y = \sin(\alpha + 2\pi x/p)$  taken for  $x = 1, 2, 3, \dots$  to impress upon random data oscillations with maxima separated from minima by about  $p/2$ . But such  $a_i$ , as well as those in (34), have abrupt endings which introduce noticeable alterations. More satisfactory results come from tapering ends, such as appear in damped vibration, with coefficients about proportional to  $e^{-c|x|} \cos 2\pi x/p$  or to  $e^{-c|x|} \sin 2\pi x/p$ . H. Labrouste and Mrs. Labrouste [7] give a powerful operator of this description.

Slutsky (loc. cit. pp. 119–123), Yule [8], and Walker [9] make use of damped harmonic vibration to explain the creation of cycles; while Bartels [10] approaches by a different method the oscillations that do not last.

Now the common graduation formulas have coefficients not conforming strictly to damped vibration, as the tapering ends vibrate more quickly. However, these ends have little more than a smoothing or stabilizing effect. Furthermore,

the coefficients for first differences are likely to conform to something like  $e^{-c|x|} \sin 2\pi x/p$ . Some *experimental* evidence will be presented for the following conclusion:

*If the coefficients  $a_i$  of a graduation or linear process (4) appear to conform roughly to equidistant ordinates of a damped vibration,  $\pm e^{-c|x|} \cos 2\pi x/p$  or  $\pm e^{-c|x|} \sin 2\pi x/p$ , with changes of sign at intervals of  $p/2$ , then when this process (4) is applied to independent chance data having zero mean and constant variance, there is a tendency for the graduated or processed values to change sign at intervals of about  $p/2$ .*

A number of standard graduations have first and second differences—see (6), (7)—which bear a decided resemblance to damped vibrations, while the third or fourth differences have only moderate, if any, cyclic appearance. This is especially true of those graduations which are constructed by applying three summings—the number of terms in a sum being in general different—and a fourth

TABLE I

*Coefficients ( $\times 350$ ) for Spencer 21-term graduation and for first four differences. Also theoretical cycle lengths for change in sign in values obtained from random data*

		Cycle Length
Grad.	$\frac{+ \quad 6, 18, 33, 47, 57, 60, 57, 47, 33, 18, 6}{-1, 3, 5, 5, 2} \quad \frac{2, 5, 5, 3, 1}{}$	10.7
1 <sup>st</sup> D.	$\frac{+1, 2, 2, 0 \quad 3, 10, 14, 15, 12, 8, 3}{- \quad 3, 8, 12, 15, 14, 10, 3} \quad \frac{0, 2, 2, 1}{}$	7.0
2 <sup>nd</sup> D.	$\frac{+ \quad 2, 3, 5, 4, 3 \quad 3, 4, 5, 3, 2}{-1, 1, 0 \quad 1, 4, 7, 6, 7, 4, 1} \quad \frac{0, 1, 1}{}$	5.5
3 <sup>rd</sup> D.	$\frac{+1, 0 \quad 1, 1, 4, 3, 3 \quad 1 \quad 2, 1, 2, 1}{- \quad 1, 2, 1, 2, \quad 1 \quad 3, 3, 4, 1, 1} \quad \frac{0, 1}{}$	3.2
4 <sup>th</sup> D.	$\frac{+ \quad 1, 1, 1 \quad 1 \quad 0 \quad 1 \quad 4 \quad 4 \quad 1 \quad 0 \quad 1 \quad 1, 1, 1}{-1 \quad 1 \quad 3 \quad 3 \quad 0 \quad 2 \quad 0 \quad 3 \quad 3 \quad 1} \quad \frac{1}{}$	1.6

process with negative coefficients. This is, indeed, a favorite form of graduation, with which are associated the names of Woolhouse, Spencer, Higham, Kenchington, Henderson, etc. The Spencer 21-term formula, for which some features have already been described, [3, p. 262], will now be examined, with special reference to its differences. Cycle length for change of sign is one-half that for change from minus to plus.

In the graduation formula, itself, there are 11 positive coefficients, centrally located, and relatively large as compared with the negative coefficients. This 11 is close to 10.7 the theoretical cycle length for changes of sign of  $y_r$  — 4.5, the difference between the graduated value  $y_r$  and its mean—the arithmetic mean of 1, 2, . . . , 9. The structure of the first and second differences also



matches closely the corresponding cycle lengths. In the third differences, there is a break at the center; but still there appears considerable regularity. But among fourth differences, the zigzag is the prominent feature. Now the theorem of Section 3 does not really apply to the Spencer formula, with its two summations by fives and one summation by sevens, and another process. But it is not surprising that the cyclicity ceases after passing the third differences.

As a basis for comparing observed values with expected values, the tenth digits in the 600 logarithms from  $\log 200$  to  $\log 799$  were taken as a random set of numbers. These 600 numbers had been given a Spencer 21-term graduation [3, pp. 261–262], yielding 580 graduated values. From these the 579 first differences were found, the 578 second differences, etc. These numbers, 580, 579, . . . , were multiplied respectively by the expected relative frequencies of change in sign of  $y_r - 4.5$ , of  $\Delta y_r$ ,  $\Delta^2 y_r$ , etc., as found by use of (5), (8), and similar expressions to form the following table.

The most abrupt change in frequency or cycle length appears to occur in passing from third to fourth differences. In Table I, this is seen in the configura-

TABLE II

*Comparison of expected changes of sign with observed changes for a Spencer 21-term graduation*

	<i>Expected Number of Changes from - to +</i>	<i>Observed Number of Changes from - to +</i>
Graduated values—4.5.....	27.2	27
First differences.....	41.3	42
Second differences.....	52.9	48
Third differences.....	90.4	74
Fourth differences.....	176.7	146

tion of positive and negative terms, and in the drop from 3.2 to 1.6 in cycle length; and in Table II in the corresponding increase in expected sign changes from 90.4 to 176.7. More spectacular is the increase in the number of zigzags represented by  $-$ ,  $+$ ,  $-$ ,  $+$ . Among the third differences, there were found only 13 instances of four successive terms with signs as just indicated, whereas among fourth differences there were found 75 such instances. For random material, about 36 such zigzags would be expected—decidedly more than found among the third difference, and decidedly less than found among the fourth differences.

The Spencer 21-term graduation appears to be fairly representative of commonly used graduations as regards regularity or irregularity in the distribution of positive and negative coefficients among the differences. For graduations with a much larger number of terms, the alternation of sign in fourth differences may not be so rapid, as, e.g. in the 35-term 5th degree parabolic graduation which Macaulay [11] calls No. 18. On the other hand, for a formula with non-tapering ends, such as the 13-term formula which Macaulay gives [11,

p. 64], the coefficients appearing in the differences are more irregular, especially at the ends. While the Spencer formula is fairly representative, different formulas have distinguishing features. If it is desirable to form an idea of what a given formula will do to random data, a table like Table I can be constructed.

**5. Summary.** When upon independent chance data, summing; averaging or some more general graduation process is used, the graduated values tend to assume a wavy configuration. These waves often seem to have a fair amount of regularity or cyclicity. The first differences usually, and often other differences of the graduated values, are decidedly cyclic. But, as we go in turn to the higher differences, the cyclicity may weaken. Indeed there may be a return to something like randomness. And subsequent differencings may tend to set up zigzags.

If  $(k + 2)$  successive summings by  $n$  have been performed on independent chance data, with  $n$  not too small, say  $n \geq 5$ —then  $k + 2$  differencings will just about bring back the original chaotic or random condition. But with only  $k$  or  $(k + 1)$  differencings, a definite cyclicity remains, at least theoretically, in the expected values.

In the case of the Spencer 21-term graduation, the coefficients for the successive differences indicate the appearance of cyclicity in first, second, and third differences.

#### REFERENCES

- [1] EUGEN SLUTZKY, "The summation of random causes as the source of cyclic processes," *Econometrica*, Vol. 5 (1937), pp. 105-146. This supplements an earlier paper (1927) in Russian.
- [2] V. ROMANOVSKY, "Sur la loi sinusoïdale limite," *Rendiconti del Circolo Matematico di Palermo*, Vol. 56 (1932), pp. 82-111.
- [3] EDWARD L. DODD, "The length of the cycles which result from the graduation of chance elements," *Annals of Math. Stat.*, Vol. 10 (1939), pp. 254-264.
- [4] EDWARD L. DODD, "The problem of assigning a length to the cycle to be found in a simple moving average and in a double moving average of chance data," *Econometrica*, Vol. 9 (1941), pp. 25-37.
- [5] ROBERT HENDERSON, *Graduation of Mortality and Other Tables*, Actuarial Society of America, New York, 1919.
- [6] DINSMORE ALTER, "A group or correlation periodogram, with applications to the rainfall of the British Isles," *Monthly Weather Review*, Vol. 55 (1927), pp. 263-266.
- [7] H. AND MRS. LABROUSTE, "Harmonic analysis by means of linear combinations of ordinates," *Terr. Mag. and Atmos. Elec.*, Vol. 41 (1936), pp. 17-18.
- [8] G. UDNY YULE, "On a method of investigating periodicities in disturbed series, with special reference to Wolfer's sunspot numbers," *Phil. Trans. A*, Vol. 226 (1927), pp. 267-298.
- [9] SIR GILBERT WALKER, "On periodicity in series of related terms," *Roy. Soc. Proc.*, Ser. A, Vol. 131 (1931), pp. 518-532.
- [10] J. BARTELS, "Random fluctuations, persistence, and quasi-persistence in geophysical and cosmical periodicities," *Terr. Mag. and Atmos. Elec.*, Vol. 40 (1935), pp. 1-60.
- [11] F. R. MACAULAY, *The Smoothing of Time Series*, Publication of the National Bureau of Economic Research, No. 19, New York, 1931.