PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A
MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Hunting Quanta

D. G. Kendall

Phil. Trans. R. Soc. Lond. A 1974 **276**, 231-266
doi: 10.1098/rsta.1974.0020

Phil. Trans. R. Soc. Lond. A. **276**, 231–266 (1974) [231]
Printed in Great Britain

Hunting quanta

By D. G. Kendall, F.R.S.
Statistical Laboratory, University of Cambridge

In loving memory of Noel Bryan Slater (1912–1973)

Thom considers that many of the dimensions of megalithic sites can be expressed in terms of a quantum of 5.44 ft (1.66 m). The problem of detecting a quantum when its size is not known in advance has been studied in pioneer papers by Broadbent (1955, 1956). With the greatly increased volume of evidence now available, a renewed attack on this problem seems called for, and a new approach, based on a Fourier analysis, will be outlined here. With samples of sufficient size, testing for the presence of a quantum is equivalent to testing a section of a realization of a Gaussian stochastic process, stationary under time shifts, for an unduly large supremum. If the test is not to be prejudiced by prior decisions about the fittings of 'eggs', 'fans', etc., then only data from circle-diameters can be used. This reduces the sample size, and so the asymptotic theory has to be supplemented by 'Monte Carlo' runs with simulated data (with and without a built-in quantum effect). It already seems clear that agreement on the data-set to be used in the analysis is the vital prerequisite for a decisive test of the quantum hypothesis.

1. Introduction

Although this is an archaeological lecture, spades will not appear in it, but if they were to do so I would make every effort to call them by their everyday name. There are two reasons why technical jargon is out of place on an occasion such as this. In the first place, those to whom it is unfamiliar may lose the thread of the argument, while secondly (and this is much more serious) others to whom it is only marginally familiar may be tempted to give to the argument greater weight than it deserves, for no better reason than that it is seductively incomprehensible.

The work of Professor Alexander Thom covers an enormous range, and raises a large number of statistical questions to most of which we can at present give no adequate answer. This is because of the imponderable influence of subjective selections which have to be made when choosing 'notches' to be associated with orientations, 'eggs' to be fitted to arrays of stones, or dimensions to be measured when dealing with monuments of complex shape. In order to take such selections fully into account one has to determine the universe of possibilities from which they have been drawn, perhaps with a due weighting for each item, so that one can assess the factor by which the significance-test probabilities must be diluted. If this is to be avoided, one is left with two alternatives. The first is, to review the thesis as an integral whole, and then to accept it as 'overwhelmingly convincing', or reject it as 'utterly preposterous'; both points of view are, perhaps, represented here today. Without committing myself to one or the other, I want to see what can be done by following the alternative path of picking out from the wealth of data collected by Professor Thom a particular group of observations which seems to be maximally free from the possibilities of conscious or unconscious prejudice, and then asking a straightforward 'yes' or 'no' question of this. We must of course be prepared to find that the data-set is too small to permit an unqualified answer – a constantly recurring situation in a branch of learning in which the data are in general limited by factors quite beyond our control.

2. The data

The data which I have chosen for this study are Thom's measurements of the *diameters* of *circles*. 'A circle is a circle is a circle' (if it *is* a circle). We may measure its diameter, or its radius, but if a unit of length underlies one set of measures then one half of that unit will underlie the others, and to the same extent. Thom has also been interested in the perimeters of the circles, and at first sight we might think that these are as rigidly linked to the diameters as the radii are, but this I consider to be a dangerous assumption. When *we* speak of the perimeter of a circle we mean its circumference in the specific sense of arc-length, but there is no guarantee that the constructors (as I shall call them) did so; they may have been interested rather in the perimeter of the polygon linking the stones which (for us) define the circle. It is plain that the ratio of the circular to the polygonal perimeter depends on the number of stones (i.e. on the number of vertices of the polygon), and as some stones may in individual cases have been lost, the danger in using perimeters is evident.

I am assuming without further inquiry that the circles listed as such in Tables 5.1 and 5.2 of Thom's book (Thom 1967) are 'true' circles, and that the diameters have been measured objectively, without reference to any prior opinions about the 'quantum'. Those who are unhappy about this assumption may feel, if the outcome of the present investigation should prove to be a positive one, that the next step ought to be a re-survey of the supposed circles, possibly by yet more sophisticated (and expensive) techniques (e.g. by aerial survey), and under the direction of an unprejudiced committee (that is, a committee with an unprejudiced distribution of prejudice).†

As those who are familiar with Thom's book will know, Tables 5.1 and 5.2 deal respectively with diameters 'known to ±1 foot or better', and 'diameters known with less accuracy'. If we take the *circles* only, then there are 112 in the first set and 57 in the second, making a total of 169 in all. These Tables also include 42 'eggs' (I use this term here to include ellipses, flattened ovals, and so on), but I shall make only occasional reference to these; if they are included then the total number of measurements rises to 211. The data are separated into two groups: Scottish (S) and English & Welsh (EW). When the topographical distinction is ignored, the two sets being pooled, I shall indicate this by the symbol SEW. Data-set 1 will be the accurate circle-diameters, data-set 2 will be all circle-diameters, and data-set 3 will be all circle-diameters together with those measured from the 42 'eggs'. The complete breakdown of the number 211 of all observations is shown in table 1 which follows.

Table 1. The data-sets

	'good' circles (5.1)		all circles (5.1 and 5.2)		all data	
Scotland	S_1	(66)	S_2	(109)	S_3	(127)
England and Wales	EW_1	(46)	EW_2	(60)	EW_3	(84)
U.K.	SEW_1	(112)	SEW_2	(169)	SEW_3	(211)

† One contributor to the discussion asked if my analysis of the quantal effect could be conflated with the circle-survey procedure in one monster 'least squares' analysis, starting from Thom's individual unadjusted measures and sightings. I think he underestimated the technical difficulties of this (for my technique is not of 'least squares' type), and also failed to see how much less convincing (or perhaps how dangerously overconvincing) such an immensely complicated analysis would be. I feel myself that it is better for the argument to be broken down into small digestible morsels which can be contemplated one at a time.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Such a symbol as SEW_{2-1} will denote the less accurately measured circles only, for Scotland and England and Wales combined, and so on. *Our primary data will be that in the set* SEW_2. It must be emphasized that 169 is, statistically speaking, a small number. We may indeed learn something from this sample, but if we ask too many questions of it, we must expect to get some peculiar answers reflecting the peculiarity of the unique sample. The most straightforward procedure is to agree on a small number of simple and natural questions, and then to see what the data-set has to offer in respect of these. I shall be asking, in essence, the following question: is it reasonable to suppose that the circle-diameters could have arisen from a smooth (but not necessarily uniform) distribution over the range which they cover (from 10 ft, up to several hundreds of feet), *or* is it more reasonable to suppose that (apart from a small residual error) they are whole-number multiples of a basic unit (to be called, if it exists, the *quantum*)?

I shall work with the statute foot (≈ 0.305 m) as current unit of length throughout, because all of Thom's measurements are expressed in terms of this. By 'small' (with reference to the error component) we mean (*a*) small relative to the value of the quantum (for otherwise the second interpretation would be stupid), and (*b*) small in the sense that it might well have arisen from (i) errors made by the constructors; (ii) errors made by the surveyors; (iii) errors associated with the finite sizes of the stones themselves; (iv) errors associated with the re-erection of fallen stones etc.

As has very properly been pointed out by Professor Huber, a positive answer in favour of the second (the quantum) alternative would really amount merely to a rejection of the 'smooth distribution' hypothesis, and we can only accept this as a 'proof' that a quantum actually exists *if no other natural alternative hypothesis is available.* (For an analysis of what is 'natural', see the remarks of J. M. Hammersley in the printed discussion following Thom's 1955 paper.) [Professor Huber, in the present discussion, advanced one such further hypothesis, and although it seemed to me artificial rather than natural, I have subsequently examined it and will report briefly on it in the Appendix at the end of this paper.]

Perhaps it should be mentioned, even though it is obvious, that the hypothesis of the existence of a unit of length does not necessarily imply the existence of an invar rod, or a length of whalebone. The unit could indeed have been the height of a man (for example), and need not as such necessarily be excessively variable. If we can recruit guardsmen to within a fraction of an inch, I suppose that the equivalent task might have been achieved in neolithic times, if the matter were thought to be important enough. If we were to find evidence for a quantum approximating to some human dimension, that would not of necessity imply that the constructional work was in fact done 'anatomically', although it might be taken to indicate that at a much earlier stage this *may* have been so. After all, many of our English units (foot, span, cubit, etc.) have an anatomical reference with an interesting history behind it, but we do not therefore assume that there was at one time a 'standard human foot' kept in the Tower of London. If a unit were employed in the construction of these monuments, it could have been 'stored' and 'transmitted' from one generation to another, and transferred from one location to another, in a variety of ways. The primary question is not *how* measurements were made, but *whether* they were made. Of course the technique used (a whalebone rod being one extreme example, and human 'pacing' another) does have implications for the error terms, and to this we shall return later [see the Appendix].

MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES
THE ROYAL SOCIETY A
PHILOSOPHICAL TRANSACTIONS OF

3. The basic principles of quantal analysis

We now turn to the analysis of the circle-diameters. Here I will permit myself one or two mathematical formulae; as in this meeting we are apparently expected to be well read in the Chinese language, a little trigonometry may perhaps also be presumed to be common ground.

In its simplest form, Thom's quantum hypothesis is that an observed circle diameter X can be considered to be composed of two additive components:

$$X = Mq + \epsilon. \tag{3.1}$$

Here q is the quantum, in statute feet (or twice the quantum, if we ought to have been reckoning with radii), M is a whole number, and ϵ denotes the error component as at (i) to (iv) above in § 2. In a more refined analysis one might wish to break ϵ down into two independent components ϵ_1 and ϵ_2, where ϵ_1 is the cumulative effect of the errors to be attributed to each individual version of the M quanta making up X, so in average magnitude proportional to the square-root of M, whence to the square-root of X (this arises particularly in connexion with the 'pacing hypothesis'), and where ϵ_2 represents the variation in average quantum size from one circle to another (and in particular from one locality to another), and is thus in average magnitude proportional to M itself, and so to X itself. We shall, however, largely ignore the possibility of M (i.e. X) dependence in the error term, and leave that more technical matter to some other occasion.

Thom also worked for a time with a hypothesis slightly more complicated than (3.1) in that a constant $c = q\beta/(2\pi)$ was added to the right-hand side. I have examined this possibility also, as my analysis conveniently allows one to test the hypothesis $\beta = 0$ against $\beta > 0$. Like Thom, I found that the data did not support $\beta > 0$ (or rather, did not object to the simpler hypothesis $\beta = 0$), so I shall not go into such questions in any further detail today. (See, however, the Appendix.)

We have to test the hypothesis (3.1) against the alternative that the observed diameters X are better described by a smooth continuous distribution over the full range. This is a well-defined problem, and it is somewhat older than is generally believed. In 1815 Prout advanced such a proposal in respect of the atomic masses of the chemical elements, and a century later a statistician (von Mises, in 1918) tackled the matter with the aid of the statistical techniques available at that time. However, the problem was killed stone dead by Aston one year later, when he was able to announce that the atomic masses of all *isotopes* (save hydrogen) measured up to the date just mentioned were exactly quantal, the quantum being $\frac{1}{16}$th of the atomic mass of oxygen. (The anomalous position of hydrogen does not concern us here, but it is sobering to reflect that neither von Mises nor Aston can have realized all the consequences implicit in it.)

When Thom's interpretation of the metrology of the ancient British circles came to the notice of statisticians in the mid-1950s (Thom 1955), what happened was that Broadbent in two notable papers (1955, 1956) attacked the whole question *de novo*. Whether he was familiar with von Mises's earlier work I do not know, but I suspect that he was not; at all events Broadbent's analysis was entirely different in style and vastly more penetrating, and his work has dominated that of all others for the last 20 years. My own approach, worked out specifically for the present occasion, owes something to von Mises and something to Broadbent, and I should like to express my admiration for the pioneer work of these two writers. In its final form, however, my technique is quite different from either of its predecessors, although not all the points of difference will be apparent today, when emphasis on mere matters of technology would be quite out of place. Just because Broadbent's work has been used so extensively by archaeologists, however, it seems necessary to point out that devices introduced by him cannot necessarily be facilely transferred to my analysis (and vice versa).

The spirit of the present analysis can be very simply conveyed by the following 'thought experiment'. Suppose for the moment that we are interested in comparing the data with a *particular* suggested value q for the quantum. As Broadbent was the first to point out, and as I shall stress over and over again today, this is *not* the situation in which we actually find ourselves, but we adopt it temporarily for convenience of exposition.

This being so, we construct the following piece of apparatus, in structure simple enough to have been devised and built by many of the ancient civilizations which have been described to us by earlier lecturers. It consists of (1) a wheel and (2) a long transparent tape. The wheel is to have a circular perimeter (*not* diameter!) of exactly q feet. One point of the perimeter of the wheel (to be called the zero-point) is to be clearly marked, and one end of the tape is to be attached to this point. The tape is to carry a scale (e.g. of feet), with origin at the zero-point.

Suppose that we have a set of N measured diameters of ancient circles, or chemical atomic masses, or any other data which it is desired to test for quanticity *with quantum* q. Take each measurement X_j ($j = 1, 2, \ldots, N$) in turn, and lay out on the tape a length X_j (starting always at the zero-point, that is, at the fixed end of the tape), marking the tape boldly at the other end of the measured length X_j. Notice that the segments of length $X_1, X_2, \ldots$ are not laid 'end to end' along the tape, but partially overlap because they all start at the same point (the zero-point).

Now wind the tape round and round the wheel until one can proceed no further. (Which way round is irrelevant.) If the data were *exactly* quantal, *with quantum* q (so that all the ϵ's in formula (3.1) would be equal to zero), then it will I hope be obvious that on looking through the layers of transparent 'wound-up' tape we should see all the 'marks' lying on top of one another and coinciding with the zero-point. (If there is a 'β effect', then the marks would still coincide, but would no longer lie on top of the zero-point. However, we are not going into that matter today.)

Next suppose that the data are not quantal, but that instead they are smoothly distributed over the whole range of possibilities. Then of course the marks seen on the wound-up tape will be more or less uniformly but randomly distributed 'higgledy-piggledy' all around the wheel. Furthermore, much the same thing will happen even if the data *are* quantal, provided that the true quantum is neither q nor a rational multiple of q.

Finally return to the situation where the data are quantal, with quantum q, but suppose that the ϵ effect is no longer negligible. Then the marks now seen on the tape will lie, not exactly together, but still near to one another and near to the zero-point, and the amount of scatter which they show about the zero-point will increase with an increase in the statistical order of magnitude in the error-component ϵ, until, with a really large error-component, we shall once again have a uniform random distribution of marks all around the wheel.

Clearly in testing for quanticity with quantum q, we must assess how far the marked points on the wound-up tape are clustered round the zero-point.

Now most of the audience today, perhaps, are familiar with only one law of statistical distribution, namely that called Gaussian. But the Gaussian law has to do with errors scattered over the whole line from $-\infty$ to $+\infty$, and here we clearly have to do with an error-distribution confined to the perimeter of a circle. Statisticians know quite well how to modify the Gaussian distribution in order to produce its 'circular equivalent'. This time suppose that we have an *endless* (doubly infinite) transparent tape with a 'central point' marked on it and fixed to the wheel at the zero-point. Suppose further that this new tape is finely covered with narrow marks that are heavily clustered *on the tape* near to its central point, and less and less clustered as one moves away from it in either direction, in such a way that the density of marks at a particular place *on the tape* is proportional to the density of a particular Gaussian law (with given standard deviation σ) at that distance from the centre point.

This being arranged, we wind our new tape around the wheel (divested now of its old single-ended tape, which has served its purpose for the moment). We are in a new situation, however, because our new tape has two infinite 'arms', and what we are required to do is to wind one 'arm' infinitely often round the wheel in a clockwise direction, and the other arm infinitely often round the wheel in a counterclockwise direction. (For this to be physically possible the two arms of the tape have to be capable of passing through one another, but we are not to worry about that. Just imagine that you are wrapping the longest possible scarf round your neck on the coldest possible night.)

It will be clear, I hope, that after this new experiment has been carried out we shall see, on looking through the infinitely many layers of transparent tape, a distribution of marks *on the perimeter of the wheel* which are densest near the zero-point, and which fall off in density steadily and symmetrically as we move in either direction away from the zero-point towards its opposite or 'antipodal' point. How close the clustering is to the zero-point will of course depend on the magnitude of σ. If σ is very small, the concentration will be immense. If σ is very large, then we shall have a very nearly uniform distribution of marks round the whole perimeter of the wheel.

For given σ, *the density distribution of marks on the wheel follows what is called the circular Gaussian distribution* (or sometimes, as here, the *lumped Gaussian*).

The advantage of this construction is, of course, that it provides us with an adjustable standard, against which to assess the degree of concentration of marks about the zero-point in our first 'thought experiment' (that one using the real data, and the single-ended tape). If the concentration near the zero-point is high, then we shall be able to describe it quantitatively by quoting the σ value which best yields a matching distribution of marks in the associated lumped Gaussian distribution.

The lumped Gaussian distribution appears in the work of Broadbent, and in a sense it is the generating element therein. However, there is another frequently used error-distribution on the circle, first introduced by von Mises in the paper cited above (1918), but sometimes erroneously attributed to others. I will try to give an intuitive idea of this distribution by an appeal to the science of archery.

Suppose we have a skilled archer equipped with a good bow, and best quality arrows, and that he is trying to hit the conventional circular target embellished with the usual coloured rings concentric with a 'bull' occupying the centre of the target. If the range is short, he will of course hit the bull every time, but let us suppose that it is so large that he will only rarely hit the bull, and will sometimes miss the target altogether. As we are supposing that he is a first-class shot, with the best possible weapons, we may suppose that his strikes will be symmetrically disposed around the target, with highest density at the bull, and a density falling off steadily as one moves out from the bull to the successively larger surrounding circles. It will be a plausible assumption, and not in fact far from the truth, to take the law of distribution of arrow-strikes on the target to be the two-dimensional circular Gaussian law, centred on the bull.

If we fix our attention on one of the more extreme surrounding coloured rings, then we will expect the arrows which strike it to lie roughly uniformly around that ring.

But now suppose that a steady left-to-right wind is blowing, and that our archer, skilled as he is in all other respects, has not yet learned that one must allow for this. Then of course the pattern of his arrow-strikes will be shifted to the right of the bull, and the strikes on the particular coloured ring which interests us will no longer be uniformly disposed around that ring, but will have a higher density on its right-hand side, and a lower density on its left-hand side. *The angular distribution of strikes on the ring*, if one idealizes the situation by making the ring very thin in a radial sense, *will be a von Mises distribution.*

Those who wish to see a formula explicitly giving the von Mises density can easily be satisfied, for it is, when we work in radian units of angular measure on the ring,

$$K(k)\, e^{k\cos\theta} \quad (-\pi < \theta < \pi). \tag{3.2}$$

Here k is positive, and it is large when the cross-wind is high and the asymmetry of the distribution greatest (maximum concentration at the right-hand extreme point, labelled here as $\theta = 0$). When there is little or no cross-wind then k is small and close to zero, and the distribution is nearly uniform on the ring. The constant $K(k)$ need not detain us; we know quite well what it is, and it is in any case uniquely determined by the requirement that the density at (3.2) must sum up to unity when we integrate round the circle.

It is a very curious and not quite adequately explained matter of *fact* (Stephens 1963) that if we take any one lumped Gaussian distribution, characterized by the appropriate value of σ, then we can find a von Mises distribution with appropriately corresponding k such that the two distributions are almost identical. Some care is necessary in giving numerical effect to this, because we have described the von Mises distribution in angular terms, and the lumped Gaussian distribution in linear terms. The relationship has to take into account the fact that the perimeter of our ring/wheel is 2π in the one case, and q in the other case; there is no difficulty in dealing with this, but it should not be forgotten.

Figure 1 shows a collection of lumped Gaussian distributions, and figure 2 a collection of von Mises distributions. Finally figure 3 shows a matched pair of distributions (one lumped Gaussian, one von Mises); they are not quite identical, but it will be evident that a very large sample of observations would be required to discriminate between them.

For analytical, computational, and statistical purposes it is sometimes the lumped Gaussian distribution which is the more convenient, and sometimes the von Mises. As they are practically indistinguishable, we shall have (and will accept) the option of using sometimes one, and sometimes the other, and we shall change horses more than once in crossing the broad stream that lies before us.

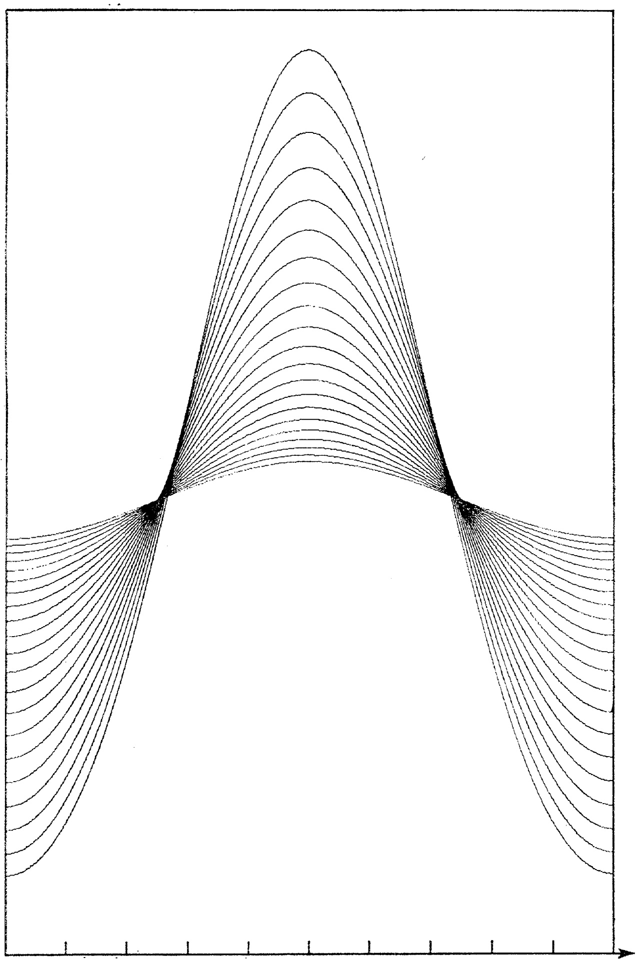


Figure 1. Some 'lumped Gaussian' distributions ($\sigma = 0.20$ to 0.40).

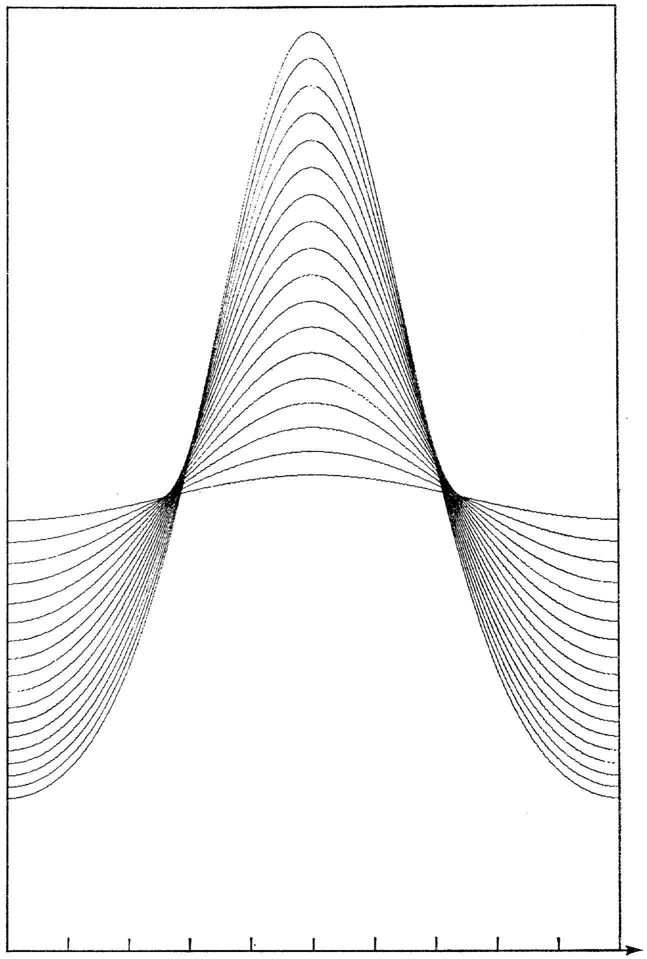


Figure 2. Some 'von Mises' distributions ($k = 0.05$ to 0.90).

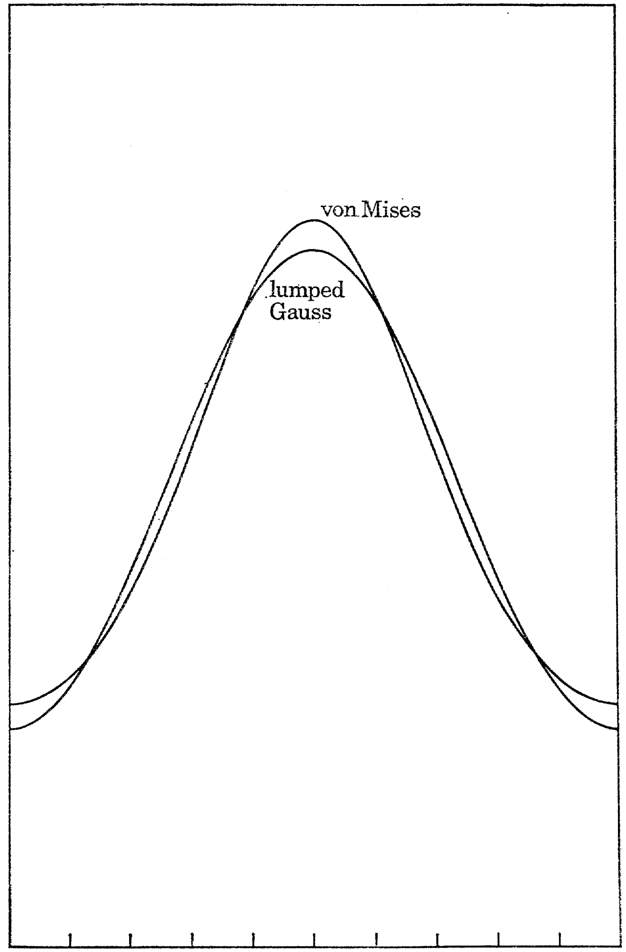


Figure 3. A matched pair of lumped Gaussian and von Mises distributions.

4. The cosine quantogram

We have remarked that position on the wheel/ring carrying our markers can in a natural way be transformed into angular terms, using angular displacement θ in radians (measured from the zero-point) as a position-indicator on the circumference. An equally good position-indicator for our purposes is $\cos\theta$, which ranges from -1 to $+1$; it takes the value $+1$ at the zero-point, and falls off steadily to -1 as one proceeds away from the zero-point *in either direction*. Thus a natural (and in relation to the von Mises distribution, *the* natural) measure of clustering of marks around the zero-point will be the sum $\Sigma \cos\theta$, normed in some appropriate way so as to take into account the size N of the sample of measured values. If we follow this idea to its logical conclusion, we are led to the view that the appropriate 'score' for high clustering, in relation to the prescribed quantum q, is

$$\phi(\tau) = \sqrt{\left(\frac{2}{N}\right)} \sum_{j=1}^{N} \cos(2\pi X_j \tau), \tag{4.1}$$

where $\tau = 1/q$ is the reciprocal of the quantum, measured in ft^{-1} (one can think of τ as a frequency measure). To understand this formula, notice that if as in (3.1) we have $X_j = M_j q + \epsilon_j$, then

$$\cos(2\pi X_j \tau) = \cos(2\pi\epsilon_j \tau),$$

and as ϵ_j is the error in linear arc-length on a wheel of perimeter q, so $2\pi\epsilon_j\tau$ is the corresponding error in angular terms (measured in radians). If all ϵ values are zero, then we shall have the large value $\sqrt{(2N)}$ for $\phi(\tau)$. If, however, the errors ϵ are very large, or if the quantum has been wrongly chosen, or if there is no quantal effect at all in the data, then the contributions of the cosines in (4.1) will largely cancel out through interference, and $\phi(\tau)$ will be small, and may be either positive or negative.

The purpose of the factor $\sqrt{(2/N)}$ is to allow for the dependence of the statistical 'size' of $\phi(\tau)$, when either there is no quantum, or when the quantum has been chosen incorrectly, on the size N of the sample. In fact, for not too small N, and for not too large q-values (i.e. for not too small τ-values), it can be shown that $\phi(\tau)$ *for fixed* τ is very nearly an ordinary Gaussian random variable with zero mean and unit standard deviation, so that a value in excess of $+4$ would only occur with a probability of about 0.000032.

We now have to make two exceedingly important points, both due to Broadbent, although he made them with reference to his own rather different method of analysis.

First, we do *not* know *a priori* what the quantum q is supposed to be (equivalently, we do not know what τ is supposed to be). Thus we have to consider *not* the single number $\phi(\tau)$, computed from the observations, but rather *the range of values of* $\phi(\tau)$ *as* τ *is varied*. We do not, however, have to consider the whole range $0 < \tau < \infty$, for this would correspond to $\infty > q > 0$, and clearly microscopically small and cosmologically large values of q are irrelevant. Thus the range must be truncated in some way, at each end. Indeed, if we do not do this, disaster will ensue, as is easily seen. For suppose we take q to be smaller and smaller; ultimately we will arrive at $q = 0.1$ ($\tau = 10$), and we will then certainly discover a quantum, because none of Thom's observations were quoted to more than one place of decimals! On the other hand, if we take q to be larger and larger, ultimately we will reach a situation in which (3.1) will be satisfied quite adequately by taking $M_j = 0$, and $\epsilon_j = X_j$, and once again we will have found a (ridiculous) quantum.

So much is easy to appreciate, once it has been pointed out. It is not quite so easy to decide how to choose the appropriate range, $\tau_0 < \tau < \tau_1$, or equivalently, $q_0 < q < q_1$, where $\tau_0 = 1/q_1$, $\tau_1 = 1/q_0$, although it is not hard to fix some bounds; the difficulty is in agreeing on precise ones. For the sake of argument we might choose $q_0 = 2$ ft (the dimensions of a single stone) and $q_1 = 10$ ft (the diameter of the smallest circle). This gives $\tau_0 = 0.1$ and $\tau_1 = 0.5$. For technical reasons it turns out to be convenient to use instead

$$\tau_0 = 0.09 \quad \text{and} \quad \tau_1 = 0.59,$$

thus giving $\tau_1 - \tau_0 = 0.50$. In what follows, this has been the basic choice of τ_0 and τ_1, although many computations have been performed with both wider and narrower ranges than this. The computer program implementing my procedure will of course accept any (τ_0, τ_1) values one chooses, perhaps at the cost of a greater expenditure of computer time; with a wide range of τ values it is necessary to compute and plot a very large number of points $(\tau, \phi(\tau))$ in order to be sure of picking up all the 'peaks' mentioned below. For this reason it is preferable, if a wide τ range has to be searched, to break this up into several subranges and then to search each one separately.

We are now in a position to appreciate Broadbent's second point, which is (in the present context) that we must look at the plot of $\phi(\tau)$ (plotted vertically) against τ (plotted horizontally) over the range $\tau_0 < \tau < \tau_1$, and pick out the principal 'peak'. But this plot is the plot of a sum of cosines, and will be wildly fluctuating; peaks will in fact abound. Which are we to choose? There is a latent difficulty of choice here, which Broadbent solved in one way, and I in another (see the appendix), but supposing (to make matters as simple as they could possibly be) that there is one overwhelmingly dominant peak, how are we to decide whether or not we have just got this big peak 'by chance'? On the face of it, if we find a peak greater than $+4$, we might say that the anti-quantum hypothesis has been rejected at the 0.0032 % level, but as Broadbent pointed out this is totally inadmissible because we *chose* the biggest peak in the range, so made it too easy for ourselves to get such a big one. (The position would of course be logically utterly different if the quantum q, and so the value of τ, were specified by our hypothesis *in advance*.) Thus the fact that we are choosing the highest peak in the whole of the τ interval requires us to water down the naïve significance level by a very considerable factor, and it is not altogether easy to say, by how much.

There are various ways of tackling this difficulty. For example (and here, in an aside for statisticians, jargon will creep in) we can take note of the fact that for not too small N and not too small τ_0, the process $\{\phi(\tau):\tau_0 < \tau < \tau_1\}$ is asymptotically a section of a stationary normalized Gaussian stochastic process for which the autocorrelation function can be estimated near $u = 0$ by setting $\rho(u) = \phi(|u|)/\sqrt{(2N)}$. Unfortunately, while $\rho(u)$ for the relevant small u can be (and has been) estimated in this way, we do not have a good usable theory for the distribution of the supremum of such a process. We could, of course, see how rapidly $\rho(u)$ approaches zero as u increases, and so estimate crudely how many 'independent pieces of information' are represented by the section (τ_0, τ_1) of the stochastic process. Alternatively we could invoke a limit theorem of Cramér concerning the supremum functional for stationary Gaussian processes. But we can only estimate $\rho(u)$ near to $u = 0$; also we do not know what are the circumstances in which Cramér's limit theorem can safely be used for a given finite N, and finite values of τ. (Here the aside to specialists ends.)

We are therefore driven, as was also Broadbent driven in his earlier but rather different investigation, to what are called *Monte Carlo techniques*. Despite the proximity of this building to a great palace devoted to the 'doctrine of chances', it will not be reasonable to suppose that everyone present is equally familiar with Monte Carlo procedures, so I will give a brief sketch of these in so far as they relate to our problem.

5. Monte Carlo, and 'Mount Thom'

Let us first take a look at a typical example of the plot of $\phi(\tau)$ against τ; I shall call such a plot a *cosine quantogram* (c.q.g.). That shown in figure 4 is taken over the very wide range

$$\tau_0 = 0.05, \quad \tau_1 = 2.55,$$

and here for once we are using the large data-set involving 211 measurements ('eggs and all'). As remarked above, such 'wide' plots are not very convenient and it is better to scan a wide τ range piece-by-piece, but it will shortly appear why we have deliberately taken a single wide range for our introductory example. Several large peaks will be observed. Of two on the right-hand side, the larger has

$$\tau = 0.05+(7.8/10)\ (2.55-0.05) = 2.00, \quad \text{so} \quad q = 0.5,$$

and the large peak near the middle of the range has

$$\tau = 0.05+(3.8/10)\ (2.55-0.05) = 1.00, \quad \text{so} \quad q = 1.0,$$

while similar measurements and calculations show that for the very large peak on the left-hand side of the range (indicated by vertical and horizontal markers),

$$q = 5.442.$$

Now the first two of these peaks obviously arise from the circumstance that when the less accurately measured circles are included, a number of the measurements will have been rounded by Thom to the nearest foot, or the nearest half-foot. It is exceedingly satisfactory that the c.q.g. has picked up these 'factitious' quantal effects, and it is also a very striking fact that the major peak, that at Thom's 'megalithic fathom' value $q = 5.44$, is actually more prominent than the two factitious peaks of whose reality we can be in no doubt.

We can eliminate the factitious peaks in either of two ways; by reducing τ_1 to a smaller and more appropriate value (and so just leaving the factitious peaks 'off the screen'), or we can *unround the data*, adding a random fraction of ± 0.5 to all measurements quoted to the nearest foot, and a random fraction of ± 0.25 to all measurements quoted to the nearest half-foot. The second alternative is to be preferred, and in order to be quite systematic about this I have also added a random fraction of ± 0.05 ft to all the 'accurate' measurements (i.e. those quoted to the nearest 0.1 ft); thus for an 'accurate' circle X is replaced by $X+0.05U$, where U is chosen randomly (separately for each such X) from a uniform distribution over the range $(-1, 1)$, and similarly in the other cases. The program I have written allows one to choose the values of U once and for all, or separately and independently each time the data are re-analysed, and so 'the effect of unrounding' can be studied by varying the U values.

The reader is now invited to look at figure 5, which shows a cosine quantogram exactly similar to that in figure 4, save that the set of *unrounded* data has been used. The peak at $\frac{1}{2}$ ft has now collapsed, likewise that at 1 ft, but the remaining peak at 5.44 ft is as prominent as ever. The 'ladder' running up the middle of the cosine quantogram is in effect a scale of units with

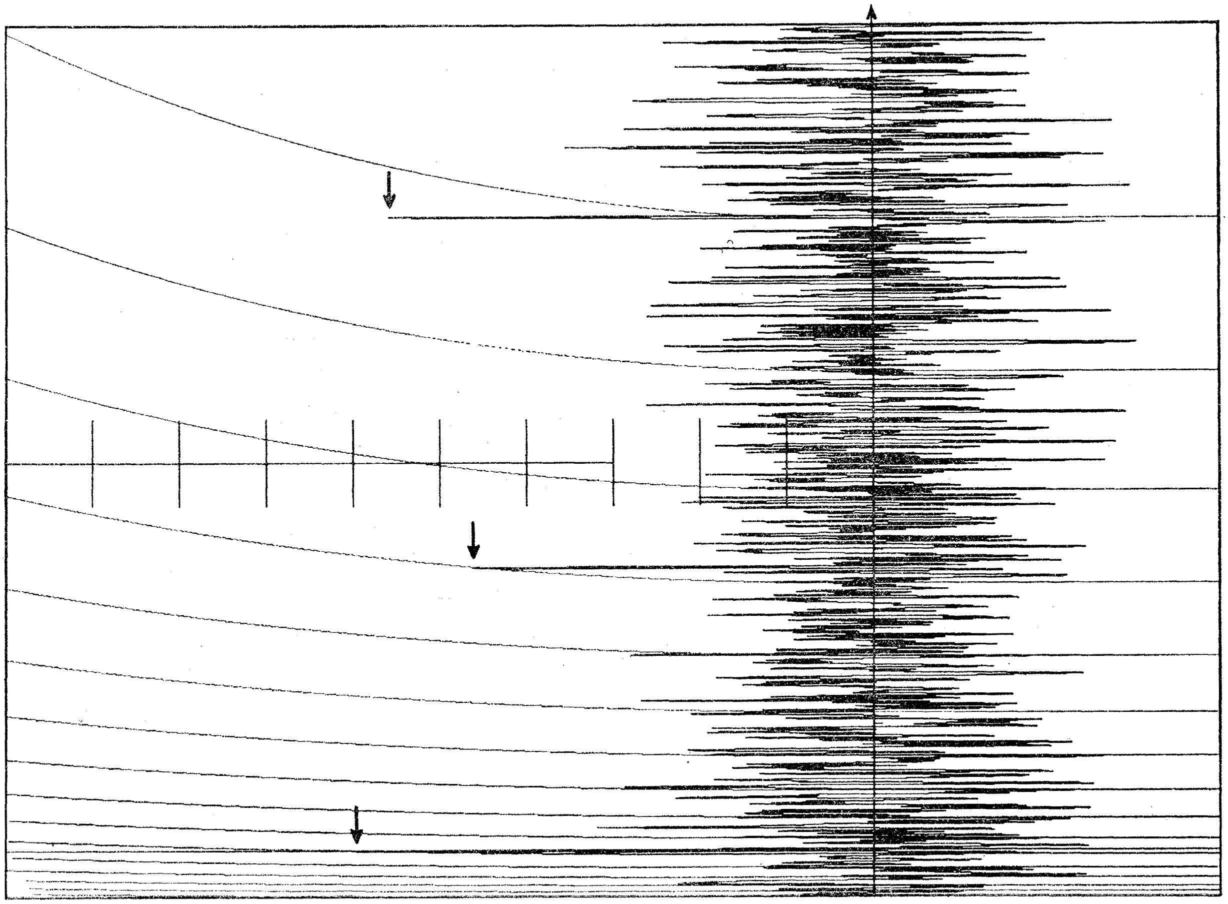


FIGURE 4. Cosine quantogram for SEW_3 (raw data) ($\tau_0 = 0.05$, $\tau_1 = 2.55$).

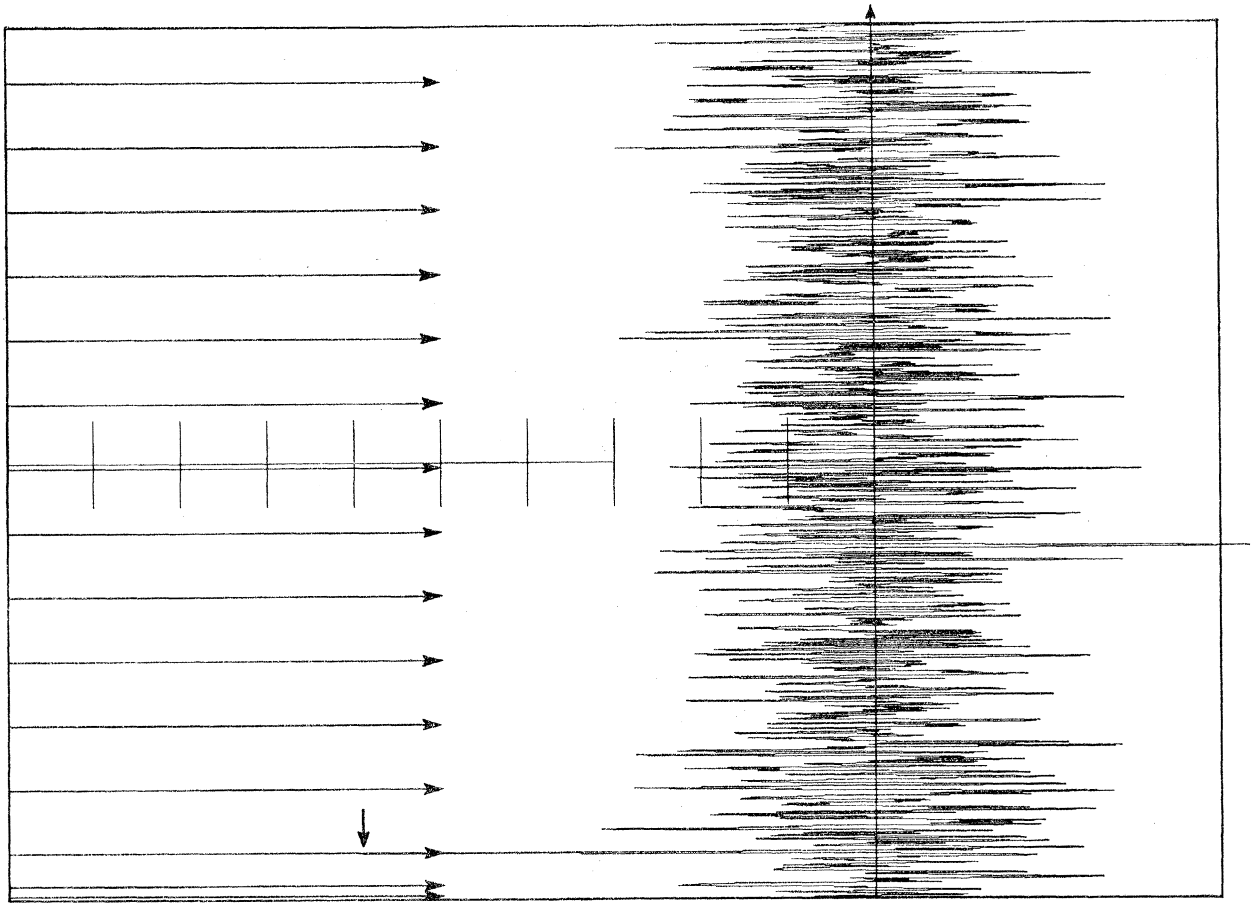


FIGURE 5. Cosine quantogram for SEW_3 (unrounded data) (τ_0, τ_1 as before).

respect to which the standard deviation of a single value of $\phi(\tau)$, that is, a single ordinate of the cosine quantogram, is 1 unit. Thus the peak at 5.44 ft is immensely significant, judged as a peak *by itself*, but the question is rather, is it significant when due allowance is made for the enormous number of peaks present, any one of which might have been chosen?

The various *vertical lines* on the new c.q.g. in figure 5 indicate the integer multiples and submultiples of 5.442 ft, and they are drawn automatically to enable the Thom peak and its under- and over-tones to be identified visually without having to go through a laborious procedure of measuring τ and converting to q. The *curves* (very steep, and on the left almost vertical straight lines) on the c.q.g. in figure 4 are *isopithons* (which I name after Πειθώ: the goddess of Persuasion); they represent loci of constant relative support in the sense of the theory of likelihood (Edwards 1972). The method of construction and use of these will be presented elsewhere (see the appendix). What is relevant here is that the two 'factitious' peaks at 1 ft and $\frac{1}{2}$ ft both lay to the right of and below the isopithon through the peak at 5.44, and this fact *of itself* would have suggested the priority of the 5.44 ft peak.

From analytical considerations it can be shown that large values of $\phi(\tau)$ (arbitrarily close to $\sqrt{(2N)}$) will occur for arbitrarily large τ. The isopithons show that these are to be regarded as dominated by 'moderate' peaks at 'moderate' τ values.

A glance at either of these two figures is sufficient to indicate that we here have to do with a 'signal:noise ratio' problem; we need to have some way of picking up genuine signals of quanticity from the general background of 'grass' in which they grow. As already remarked, the cosine quantograms in figures 4 and 5 cover much too wide a span of τ values, and were chosen to do so simply to illustrate the power of our technique in picking up factitious peaks like the ones at 1 ft and $\frac{1}{2}$ ft. But now we shall revert to the 'official' τ range,

$$0.09 < \tau < 0.59,$$

and we shall also revert to the 'official' data set, SEW_2, which, it will be remembered, contains only 169 observations.

Before leaving figure 5, it will be convenient to deal briefly with a question which may have formulated itself in the minds of some of those here: why do we work with cosines only, in (4.1) above; do not the *sines* also contain information of value? The answer to this question is readily given. If we define $\psi(\tau)$ exactly as at (4.1), but with the cosines everywhere replaced by sines, and if we then define

$$U(\tau) = \tfrac{1}{2}\{\phi(\tau)^2+\psi(\tau)^2\} \quad (\tau_0 < \tau < \tau_1), \tag{5.1}$$

then we obtain in $U(\tau)$ a quantity which asymptotically has a negative-exponential distribution with unit parameter, and which 'peaks' at a value $\tau = 1/q$ when the observations follow a modified law,

$$X_j = \beta+M_j q+\epsilon_j \quad (0 \leqslant \beta < q). \tag{5.2}$$

Another version of my computer program plots $U(\tau)$ against τ, complete with appropriate isopithons etc., (the *modular quantogram*) and also prints out an estimate of β, and an indication of whether it is significantly different from zero. The estimate of β not being significant for the Thom data, this aspect of the problem is not entered into any further here, but in the mathematical account of this investigation more details will be given (see appendix).

How are we to decide whether the peak at 5.44 ft in figure 5 is 'significant'; more exactly, how likely is it that we should find a peak as high or higher than this, *with its τ value within the agreed range* (0.09, 0.59), by the mere operations of chance?

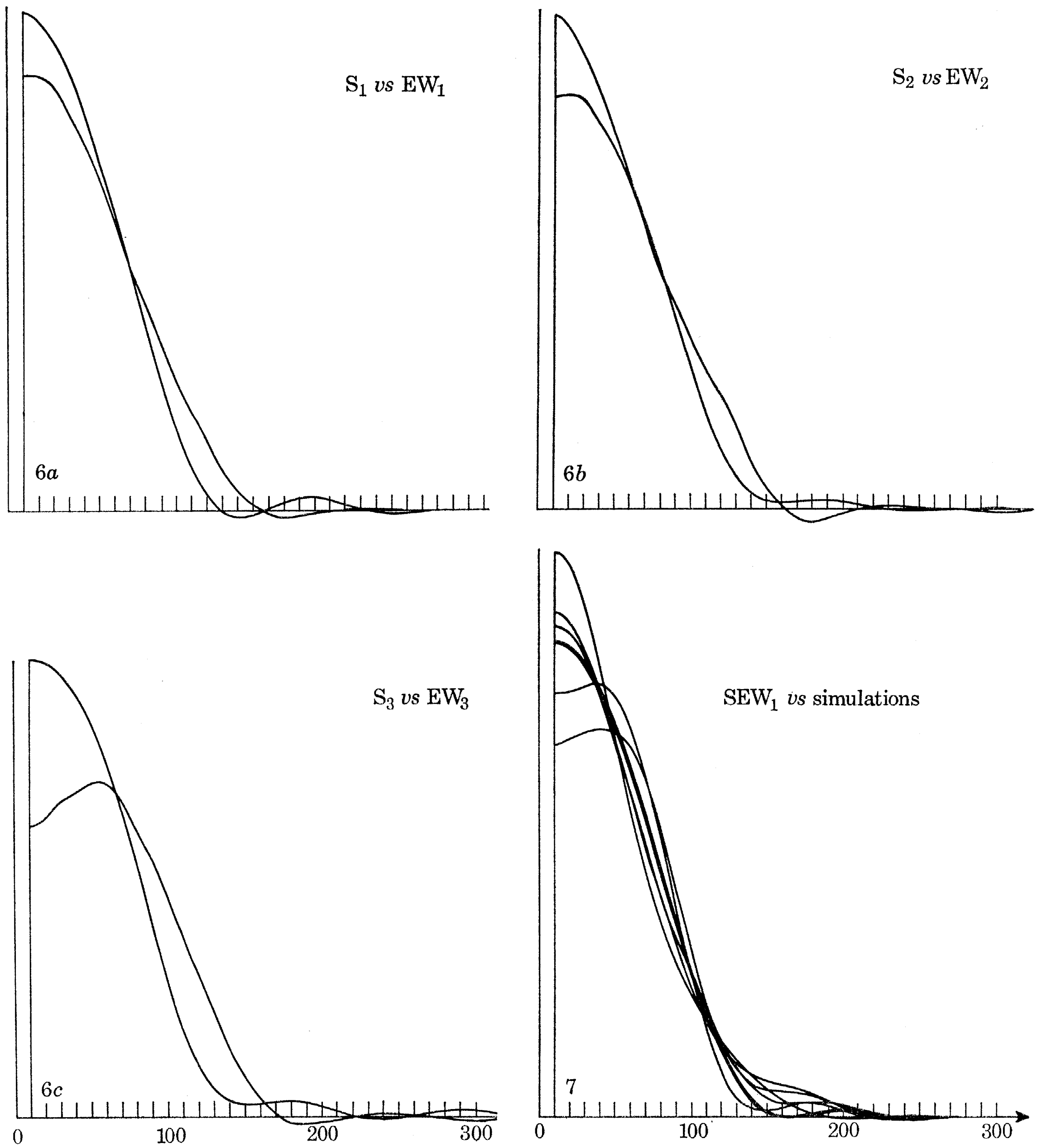


FIGURE 6. Spline transforms for the distribution of circle diameters (in feet): Scotland versus England and Wales, sets 1–3.

FIGURE 7. Spline transforms for the distribution of circle diameters (in feet): SEW_1 versus 5 simulations.

To make this question meaningful we must try to imagine what might happen if we took a 'random' set of data, similar in all statistical respects to the actual data save only in definitely *not* having any underlying quantal effect, and subjected it to the same analysis; i.e. computed its cosine quantogram over the agreed range (0.09, 0.59). Of course we should learn nothing worth knowing by doing this just once; we should have to do it at least, say, 100 times, the successive 'imitation' samples being all independent of one another.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

We now find ourselves thrown back on a deeper-lying question: what do we mean by a sample, artificial and random in character, and definitely not involving a quantal effect, and yet 'similar in all statistical respects to the actual data'? Well, in the first place, it should be of the same size (i.e. have the same value for N), though this is not in fact an essential requirement because we have very largely eliminated the effect of sample-size by including the factor $\sqrt{(2/N)}$ in the definition of $\phi(\tau)$. (Statisticians please note; the value of N has little or no effect on the autocorrelation function nor the approximating stationary normalized Gaussian process.) So we need not be too particular about the N value, and this fact will serve us well at a later stage.

I submit that by 'similar in all statistical respects to the actual data' we should mean that in the imitating samples, there will be roughly the same proportions of small, medium-sized, and large diameters. That is, we must try to arrange that both the actual data and the imitating data will have the same *coarse-grained* distribution of diameter-values. To achieve this we must have a look at the coarse-grained distribution of diameter-values for the actual data-sets; figures 6*a*, *b*, *c* show these for the S and EW data separately, for each of the three progressively enlarged data-sets 1, 2 and 3. It will be seen that there is indeed a sense in which one can speak of the coarse-grained distribution observed for X; it ranges from 10 ft (at which there is a natural cut-off) to several hundreds of feet, and the density when appropriately smoothed falls off in a fairly regular manner. An anomaly of SEW_3 need not trouble us because we shall not in fact make any serious use of that data-set.

It is a lucky accident that we can very well imitate this observed coarse-grained distribution of diameter values by one half of a Gaussian curve, with the peak shifted to the 'cut-off' at 10 ft. This is very fortunate because half-Gaussian random variables are extremely easy to simulate on a computer, so that we can produce as many replicates of the data as we like, and as they will all be derived from a half-Gaussian distribution there will be no question whatsoever of their being quantal with $0.09 < \tau < 0.59$.

I almost forgot to mention that the distributions in figure 6 are displayed not in *histogram* form, as would be more usual, but in *spline transform* form, as is far more convenient and appropriate. For the details of the spline transform, invented 3 years ago by Liliana Boneva and Ivan Stefanov (both of Sofia) and myself, see our paper (1971) which stemmed from a visit to Bulgaria under the Agreement between the Royal Society and the Bulgarian Academy of Sciences. All the reader needs to know is that the spline transform program converts a sample of observations in a unique invertible manner into a smooth curve, in such a way that the ordinate of the curve is large and positive where the observations lie thickly. Where they lie only thinly, or do not lie at all, the curve oscillates through small positive *and negative* values (part of the cost of this particular type of smoothness).

In figure 7 I show (i) the spline transform for the diameter distribution of the 112 circles in SEW_1, together with (ii) the spline transforms for 5 independent 'imitations' of SEW_1 (using the half-Gaussian simulation device). It will be observed that the curve for SEW_1 differs no more from the five imitations than they differ among themselves.

A word about the 'cut-off' at 10 ft, mentioned above: Thom's smallest circle has a diameter of 10.8 ft, and it is clear that there is some (perhaps not determinate) cut-off at the low-end of the diameter scale; one does not find, or even look for, neolithic circles of diameter 2 or 3 ft. I therefore treated 10 as an absolute cut-off for the data at the lower end, both for the observed and artificial samples, and the spline transforms have been computed taking this into account.

The formula generating the simulated circle diameters was

$$X = 10 + 57.176|Z|, \tag{5.3}$$

where Z is an 'artificial' standard Gaussian variable. The numerical coefficient was chosen to make the average simulated X-value equal to the average X-value for SEW_1. It would in retrospect have been better to use the SEW_2-value, but figure 6 shows that this would have made little difference. For SEW_2, the average X value is 58.8, while for SEW_1 it is 55.6. (From the appendix it will appear that it would have been better to 'match' r.m.s. (X) rather than $E(X)$.)

In my first attempt at handling the Monte Carlo aspects of this problem I took what now seems to me a rather stupid path, but I will report on it here because it does, in fact, have something to teach us. I ran 100 simulations of the SEW_2 data, and computed the cosine quantograms for each one. Let us pause for a moment and think out exactly what this means. It is *as if* we said to the computer: go back 4000 years, observe the method and design then used for the construction of these monuments, build yourself 169 of them, with the appropriate proportions of small, medium-sized and large diameters, *but* totally disregard any quantal practices which may at that time have been customary. Then return to the present, simulate Professor Thom, have him survey the monuments, and report the diameters on the print-out device; having carried out these instructions to the full, go back and do the same thing again 99 times, *independently*. Then construct the 100 cosine quantograms for the data thus simulated, and plot these for examination.

What I then did, in this first analysis, was to require the computer to plot out max $\phi(\tau)$ against τ, and min $\phi(\tau)$ against τ, where the 'max' and 'min' operations were carried out over the whole set of 100 simulations of SEW_2. Finally the cosine quantogram of SEW_2 was drawn on top of the picture thus produced, and figure 8 was the result. It will be noticed that the 'max' and 'min' curves are 'almost' horizontal, thus confirming the essentially stationary character of the stochastic process, and that the 'min' curve is 'almost' a reflexion (in the line $\phi = 0$, i.e. the τ-axis) of the 'max' curve. This again is as one would expect, in view of what I know about the statistical structure of this particular stochastic process. (There are no large negative values in the estimate which we have of its autocorrelation function ρ.) We can think of the 'max' curve as a sort of flood-level, showing the greatest height to which the $\phi(\tau)$-waters have risen at each separate τ, and there, proudly showing its summit above them (but only just), is what we might well call Mount Thom, at a τ value corresponding to $q = 5.44$ ft.

This result was gratifying and persuasive, but one can in fact do better. First, in view of the near-horizontal nature of the 'max' and 'min' curves, which subsequent computations confirmed, it is clear that a secular trend with τ is not to be feared, and therefore that the relevant statistic from each separate simulation is just

$$S = \sup\{\phi(\tau): 0.09 < \tau < 0.59\}.$$

However,

$$I = \inf\{\phi(\tau): 0.09 < \tau < 0.59\}$$

is in fact just as interesting, and indeed just as useful, for as a study of the autocorrelation function of the stochastic process revealed only negligible negative ordinates, we can deduce that deep troughs will not (or hardly at all) be correlated with high peaks. Therefore, to a useful degree of approximation, we can double the amount of information at our disposal by treating each $(-I)$ as if it were the S derived from *a different independent simulation*, and so on collecting

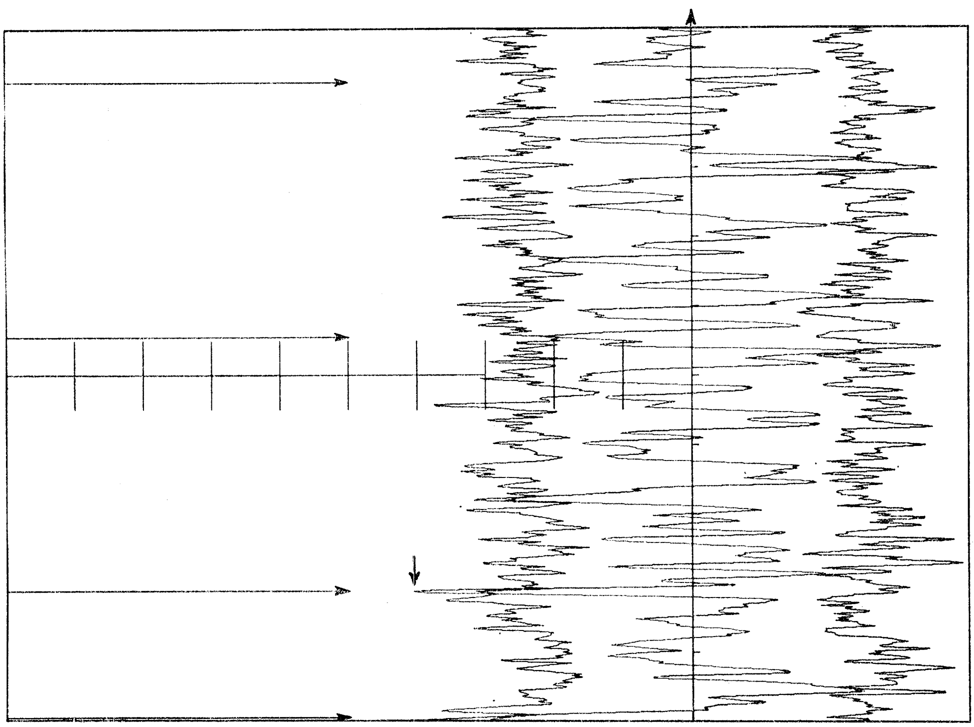


FIGURE 8. Cosine quantogram for SEW_2 ($\tau_0 = 0.09$, $\tau_1 = 0.59$), and the supremum and infimum curves for 100 simulations. The horizontal arrow indicates 'Mount Thom'.

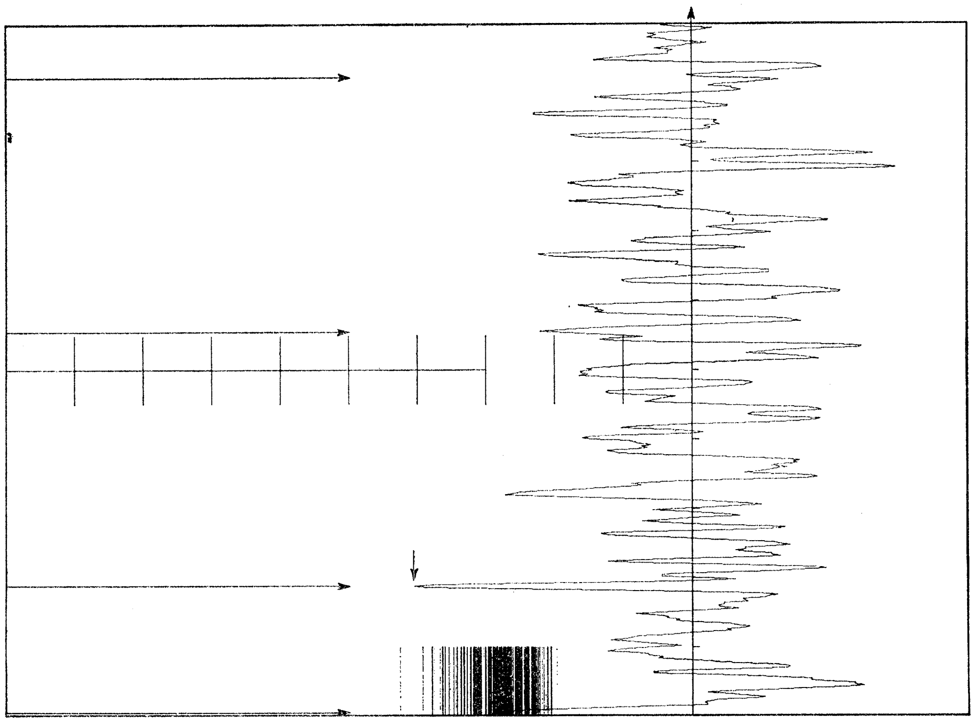
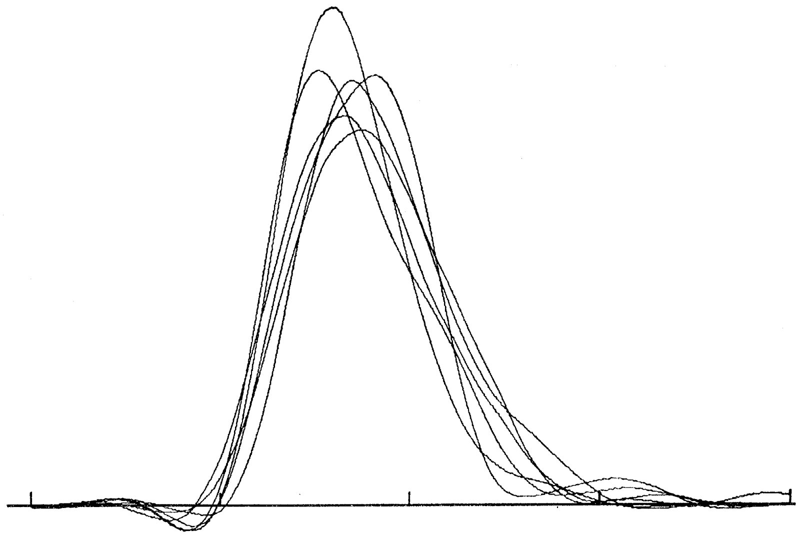


FIGURE 9. Cosine quantogram for SEW_2 ($\tau_0 = 0.09$; $\tau_1 = 0.59$), and the 200 simulated 'flood levels'. (The horizontal arrow indicates 'Mount Thom', and the vertical arrows indicate $q = 5.442$ ft and its multiples and submultiples. The scale running down the centre of the diagram is one of standard deviations.)

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

(10)



(11)

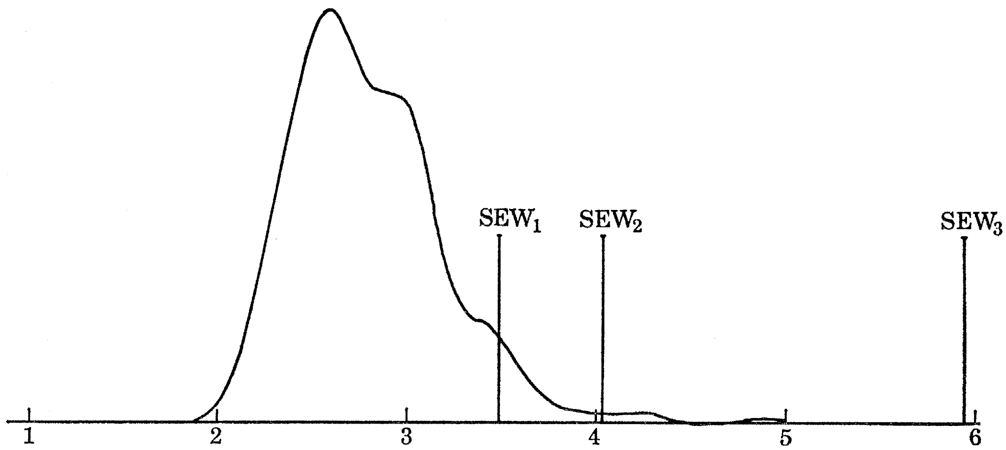


FIGURE 10. Spline transforms of three sets of 100 random S values and three sets of 100 random $(-I)$ values.

FIGURE 11. Spline transform of the whole sample of 600 simulated S values. The vertical markers show the peak heights in the cosine quantograms for $SEW_{1,2,3}$.

the 100 values of S and the 100 values of $(-I)$ we have, in effect, 200 values of S. With computer time costing what it does today, one cannot afford to ignore windfalls like this.

The reader is now invited to look at figure 9, where we again use the 'official' τ range, 0.09 to 0.59. In the top left-hand corner will be seen a thick batch of horizontal straight lines. There are 200 of these, and they are the 200 S-values obtained in the manner described as above, with $0.09 < \tau < 0.59$. That is, they are the 'water-levels' reached in 200 independent 'floods' and it will be seen that Mount Thom, like Mount Ararat, was submerged only once – and even then, only just.

Similar diagrams, not reproduced here, were constructed for SEW_1 and SEW_3, and the 200 S values were freshly and independently obtained each time. When it was realized that sample size had virtually no effect on the S distribution, it was then seen that we had in fact $200 + 200 + 200 = 600$ independent S values, and so a rather substantial amount of information about the S distribution over the τ range (0.09, 0.59) for the stationary normalized Gaussian stochastic process having this particular autocorrelation function.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Figure 10 shows a spline transform display for the distributions of the three batches of S values and the three batches of $-I$ values, in order to check that there was no appreciable difference between them. Figure 11 shows the spline transform for the S distribution using all 600 S values, and also on this picture are indicated, as vertical lines locating positions on the S scale, the height of Mount Thom for SEW_1, SEW_2 and SEW_3, respectively.

The significance of the evidence provided by SEW_3 would be quite overwhelming were it not for the fact that it is this datum which is compromised by the inclusion of 'eggs', etc.

The data set SEW_2, which I regard as the primary one, gives a pretty clear result. Inspection of the figures shows that the height of Mount Thom (4.033) was exceeded or equalled on just 5 occasions (out of 600 possible occasions), these five values being

$$4.887,\ 4.254,\ 4.293,\ 4.152,\ 4.208.$$

The hypothesis of a smooth, non-quantal distribution of circle diameters for SEW_2 is thus rejected at the 1 % level.

The data set SEW_1 (112 accurately measured circles) gave a peak at essentially the same position, and of height 3.427. This was exceeded on the five occasions noted above and also on 36 further occasions, out of the same possible total of 600. Thus the hypothesis of a smooth, non-quantal distribution of circle-diameters for SEW_1 cannot be rejected: the significance level attained is only about 7 %.

Whether one accepts the quantal hypothesis or not thus depends very much on which data-set one decides to use. However (see the appendix) the smaller N value for SEW_1 militates against the chance of detecting a real effect, if present, for (other things being equal) *the altitude of Mount Thom varies as the square-root of* N, and we cannot hope to pick it up, in relation to the general 'noise', if N is too small.

6. Conclusions

If 'a conclusion' is required, it will depend to some extent upon one's point of view. If the audience were to put a pistol to my head and demand an answer to the question: 'a quantum – yes or no?', then I should plead for more observations. A significance level of 1 % is normally regarded as a strong recommendation that the experiment be repeated on a larger scale, this being preceded perhaps by a cautious letter to *Nature*.

But if more observations are not to be had, then we must make the most of those with which we are provided. If the audience were to put the question somewhat differently, and say 'does the analysis of the available data justify the expenditure of public monies on a costly and sophisticated aerial re-survey of the circular sites?', then I think few would disagree with my affirmative reply.

Appendix

The purpose of this appendix is twofold; first, I wish to record one or two details of this new technique of quantal analysis for the benefit of those who may wish to employ it in similar problems (or to study in greater detail its application to the present problem); and secondly, I should like to offer at least a few comments on some of the very interesting points which were raised at the discussion on 8 December 1972.

There were many of these queries, some pressed very insistently; Oliver Twist himself could hardly have asked for more. I shall only attempt to deal with a few of them here, and must begin by saying straight away that the more searching the question, the more complicated the answer; it will no longer be possible, as it was in the main lecture, to avoid mathematical complexities entirely. However, I will start with some queries of a rather general character, and work gradually towards those whose answers involve technicalities.

(i) I have been asked whether, if it be accepted for the moment that 'there must be something in the quantum', does it follow *from this judgement alone* that Thom's claims with regard to astronomical alinements must be accepted in their entirety, without further examination? The answer is, no.

(ii) Professor Huber pointed out that in the data set SEW_2, consisting of 169 observations, there are no fewer than 12 in the range 20.5–22.0 ft inclusive; these are the diameters

20.5, 20.6, 20.9, 21.0, 21.0, 21.3,
21.4, 21.8, 22.0, 21.4, 21.2, 21.0.

I shall call this the data set H, and the remaining 157 observations the data set H^*. Professor Huber's remark is a very pertinent one, because this 'clumping' of the data is an observation *ex post facto*, just as was Thom's original observation of the 'quantal' effect, and so Huber is quite entitled to ask whether the 'clumping' is all that is really peculiar about the data, and whether we need to go to the extravagance of a quantal hypothesis. Mr Hogg made a rather similar point, when he spoke of 'a relatively few circles set out using some standard unit, mixed with a larger number where no standard was used'. It is largely a matter of taste whether we admit the Hogg or Huber 'hypotheses' as 'natural' alternatives to that of Thom, but let us do so, and see what happens.

To begin with, let us make some random (so *not* quantal) simulations of H^*, using the simulator (5.3) *but inhibiting this so that for X the interval of values* (20.5–22.0) *is* '*taboo*'. Ten such simulations were created, and each separate set of 157 was then made up to 169 by adjoining the Huber set H of 12 observations to it. Figure A1 shows the c.q.g of the first five of these (those of the second five are quite similar). A rather high lateral 'magnification' has been used, so that here $\tau_0 = 0.155$ and $\tau_1 = 0.205$; the vertical scale is the same as before. It will be seen that there is no special clustering of peaks near the 'Thom' value $\tau = 0.1838$ (indicated by the arrow pointing down from the top of the diagram); also the peaks which do occur within this range (indicated by short horizontal arrows) do not reach significant heights – the largest of the set of ten has a height of 2.76, which would not look at all impressive when marked in on figure 11.

Next, let us see what happens to the c.q.g. of SEW_2 when the Huber observations H are *deleted* from it. This is shown by figure A2, where two c.q.gs are superimposed; both are viewed under the high lateral magnification ($\tau_0 = 0.155$, $\tau_1 = 0.205$) used in figure A1, and that with the highest peak is in fact the c.q.g. for SEW_2 itself (height 4.03). The other curve, with a peak of height 3.23 in almost exactly the same place ($q = 5.450$ instead of 5.441) is the c.q.g. of H^*; i.e. of the 157 original observations which remain when 'Huber's 12' have been 'cast out'. A height of 3.23 is not particularly impressive either, when judged against figure 11, but it is still large enough to excite interest, and the fact that it lies in almost exactly the same position as the original one says, to me, that 'Huber's 12' made very little contribution to the latter.

We have further to bear in mind that H^* contains only 157 observations and that a *real* quantal effect (as I shall demonstrate below) will in general yield a peak whose height is proportional to the square root of the sample size, N. Now $\sqrt{(169/157)} = 1.0375$, so that the

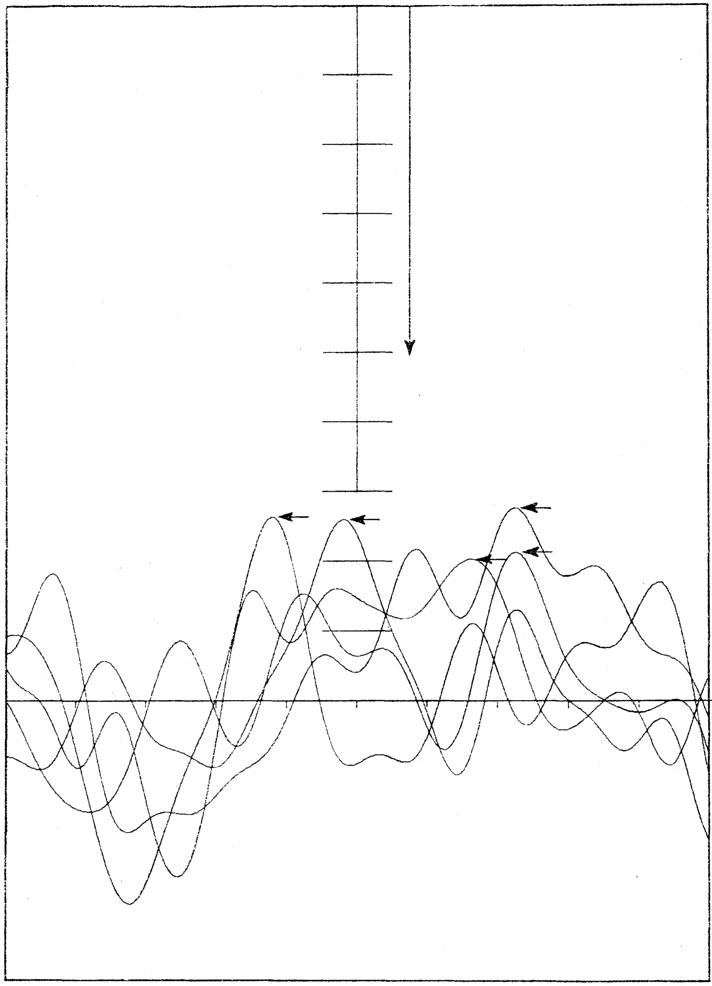


FIGURE A1. Five cosine quantograms ($\tau_0 = 0.155, \tau_1 = 0.205$), each being based on the 12 'Huber' diameters plus independent sets of 157 simulated observations.

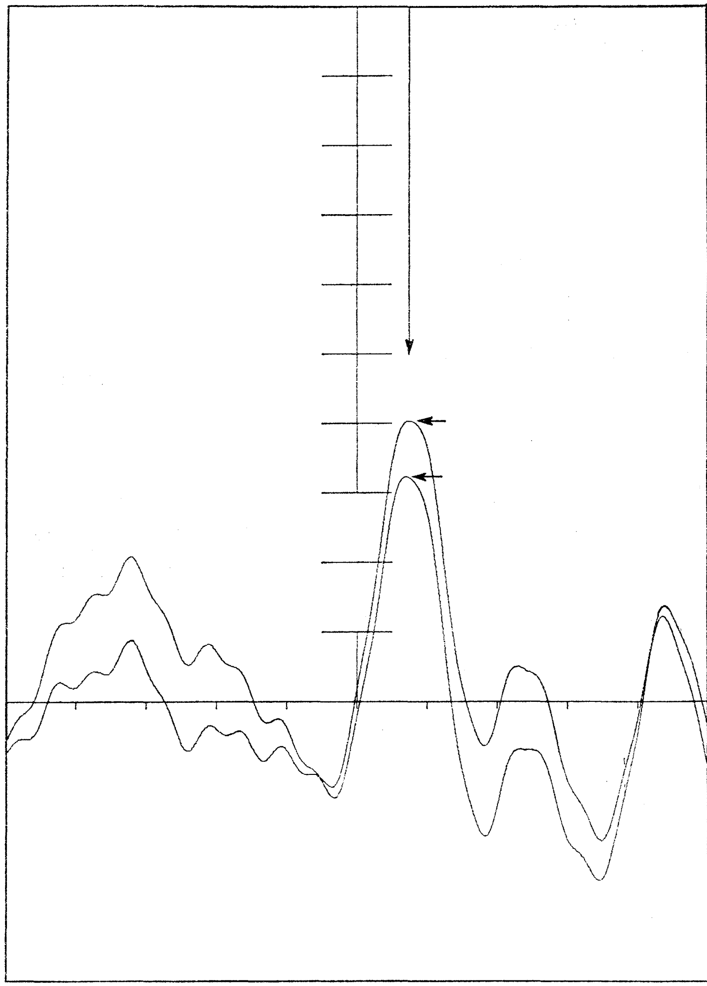


FIGURE A2. The cosine quantograms ($\tau_0 = 0.155, \tau_1 = 0.205$) for (i) SEW_2 and (ii) SEW_2 with the 12 'Huber' observations removed.

comparison with figure 11 would be on a fairer basis if the peak height used were multiplied by this last factor, thus raising it to 3.35, when it begins to stand out rather more in the tail of the S distribution.

So much for the credit side of the account. On the other hand, we must remember that if the original observations had really been random (up to the first 157) from the 'tabooed' simulator, with the bunch of 12 Huber observations superimposed to make 169 in all, then we should in the first instance have examined the c.q.g. of this material over our official search range from $\tau_0 = 0.09$ to $\tau_1 = 0.59$, so we ought to do this now. Ten such simulations were therefore created, and the table below lists the height (in the customary units) of the highest peak in the given τ range, and the associated q value in feet (the entry labelled 'randomizer' merely serves to identify the run).

TABLE A1. $H+157$ SIMULATIONS

randomizer	height	q value
00293	4.934	11.055
05025	3.558	4.260
11126	3.042	10.499
12968	2.593	5.587
44972	2.720	1.977
53632	2.473	6.757
60561	2.730	3.118
72739	4.052	4.269
94054	3.327	1.950
96561	2.493	3.428

It will be seen that we obtain in this way two peaks which are as high or higher than Mount Thom, so that if this last set of simulated material is taken as the alternative background against which the significance of Mount Thom is to be judged, then its statistical significance is completely eroded, as Huber correctly foresaw that it would be.

We still have to reckon, however, with the fact that the depleted set, 'SEW_2 with H removed', produced a substantial peak near the standard Thom value. Thus *if* independent evidence for $q = 5.44$ ft could be obtained from other monuments, then the supporting evidence from SEW_2 would be impressive. It is of course also true that *if* independent support for $q = 5.44$ ft were available, then we would be freed from the necessity of 'searching throughout a τ-range', and the difficult parts of this investigation could be avoided. Instead of referring to figure 11, we would then assess the significance of a peak height by immediate reference to the standard Gaussian distribution.

These results are clearly rather disturbing, and I am very grateful to Professor Huber for suggesting that such further experiments might be worthwhile. In order that the situation be understood beyond a shadow of a doubt, it seems worth going into a little more detail.

Let us exaggerate the situation by supposing H to consist of some number m of observations all *exactly* equal (with value x_h, say), and let us suppose that the original sample consisted of $N = n+m$ observations, of which m were all of the form $X = x_h$. (In our specific example $m = 12$, $n = 157$, and $N = 169$.) Now suppose that we replace the n 'other' observations by simulated ones, using the simulator (5.3), and to simplify calculations still further let us relax the taboo from the latter, as is only reasonable, because we have already shrunk Huber's interval (20.5 to 22.0) down to a single point, x_h. What sort of c.q.g. would we expect from this material? Well, the m replicates of x_h will of course yield m identical cosine contributions to

(4.1), which will reinforce one another with zero interference, and at any τ value which is an integer multiple of $1/x_h$ they will make a contribution $+m$ to the summation occurring in that formula. It will be clear from our previous discussions that the remaining n terms will contribute to the same summation a term which we might write as

$$\surd(\tfrac{1}{2}n)S,$$

where S is a random variable drawn from the distribution pictured in figure 11. Thus we shall have

$$\phi(\tau) = \surd(1-m/N)S+\surd(2/N)\,m,$$

and therefore we shall obtain a peak of height 4 (equal to that of Mount Thom) if and only if

$$S = \frac{4}{\surd(1-m/N)} - \surd\left(\frac{2}{N-m}\right)m.$$

Let us now insert the values $m = 12$, $N = 169$. The formula then gives $S = 2.795$. Judged by figure 11, which is now appropriate to the problem, this is a typical rather than an extreme value for S, *but in order to obtain the value* 4 *for the peak-height in the c.q.g.* it is necessary for this S value to be attained at one of the multiples of $1/x_h$, and further at a multiple of $1/x_h$ which lies in the interval $0.09 < \tau < 0.59$.

Now in our problem x_h is about 21.25, and its submultiples x_h/k yield τ values in the required range for $k = 2, 3, \ldots, 12$; that is, at 11 sites within the range. To achieve a value $\phi = 4$ it is necessary and sufficient that for the simulated data, S should be about 2.795, *and that* this supremum should occur at or sufficiently near one of these 11 sites. Looked at in this way, our record of two peaks of height four or more out of the ten simulated trials will be seen to be quite reasonable.

(iii) We now turn to some other matters relating to other ways of breaking down the material, noting, as of course we must, that if there is one thing worse than having only limited data, it is a necessity to dissect it. Let us first consider the interesting question, what happens when we look separately at the data-sets

$$SEW_1,\ SEW_{2-1},\quad \text{and}\quad SEW_{3-2};$$

here the first consists of 112 accurate circles (diameters known to within ± 1 ft), the second consists of 57 less accurate circles, and the third consists of 42 'eggs'. It will be interesting to see how far they agree in pointing to 5.44 ft as a quantum, and also how far the statistical size of the 'errors', described as ϵ in (3.1), increase as we proceed from the accurate towards the less accurate or less accurately defined data. We must first explain how the statistical size of ϵ is to be calculated.

If we insert (3.1) into (4.1) with $\tau = 1/q$, we clearly obtain for $\phi(\tau)$ a multiple $\surd(2N)$ of an averaged value for cos $(2\pi\tau\epsilon)$. Let us then adopt the lumped Gauss rather than the von Mises model, and suppose that the ϵ values have a standard deviation equal to σ ft. Then the expectation value for cos $(2\pi\tau\epsilon)$ will be

$$\exp(-2\pi^2\sigma^2/q^2),$$

so that if we write the equation,

$$\text{peak height at quantum} = \surd(2N)\exp(-2\pi^2\sigma^2/q^2),$$

then this provides us with an estimate for σ, and the estimate obtained in this way for SEW_2 is, in fact,

$$\sigma = 1.507\ \text{ft}.$$

It is highly significant that this figure is of the order of the dimensions of a single stone.

Now let us apply the same technique to the 'split' data. We obtain the following results.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

TABLE A2. SPLITTING THE DATA

data	N	height	q value	σ
SEW_1	112	3.482	5.435	1.477
SEW_{2-1}	57	2.242	5.459	1.535
SEW_{3-2}	42	5.359	5.435	0.896

These results are of great interest. They show, first, that the accurate and the less accurate circles indicate the same quantum, *and the same error*; that is, the effect of the greater error in measuring the less accurately defined circles is 'swamped' by the major source of error (which I believe to be stone-size). Secondly, they show that while the 'eggs' indicate the same value for the quantum, they also indicate an appreciably smaller size for the error.

There are clearly two possible interpretations here, and I have no inclination to take sides on this issue. *Either* all layouts are eggs, and the higher errors for the circles arise from trying to squeeze an egg-shaped foot into a circular shoe, *or* the greater freedom of choice available when fitting an egg (e.g. when deciding of how many circular arcs it is to be composed, and so on) has resulted in getting an unnaturally good fit to the data by unconscious selection.

At least it is clear that I was wise to leave the 'eggs' out of my primary analysis. We can present the situation in an alternative, pictorial, fashion because my computer program among other things works out the 'remainders' when the best integer multiple of 5.442 ft is removed from each of the individual observations, and then plots a periodic spline transform (see Boneva, Kendall & Stefanov (1971) for this) showing the distribution of these residuals on the same scale as figures 1 and 2 of the main text. The results for SEW_1, SEW_{2-1} and SEW_{3-2} are shown together in figure A3, with the lumped Gauss curves as a background. It will be seen that the distributions are very similar for SEW_1 and SEW_{2-1}, and of lumped Gauss/von Mises shape, but that the distribution for SEW_{3-2} is *very* much more highly peaked near zero, so much so that its plot goes off the scale altogether. This picture, of course, just tells us the same story in a different way.

(iv) We now dissect SEW_2 along another dimension, this time a geographically meaningful one, into the Scottish data (S_2) and the English and Welsh data (EW_2), the respective sample sizes being 109 and 60. As the sample sizes are so unequal, it seemed a useful idea to split S_2 into two sets of sizes 54 and 55 respectively; these were called S_2A and S_2B. Various kinds of c.q.g. analysis are now possible.

TABLE A3. SCOTLAND VERSUS ENGLAND AND WALES

data	N	height	q value	σ
S_2	109	3.748	5.435	1.433
EW_2	60	2.067/ 1.697	4.53/ 5.46	*/1.677 (two peaks)
$S_2A.EW_2$	114	3.680	5.411	1.447
$S_2B.EW_2$	115	2.917	5.459	1.578

It will be seen that the English and Welsh data does give some support to a quantum at the Thom value, but that it also produces a slightly higher peak at about 4.5 ft. However the sample size is here very small; this suggested 'mixing' it with each of the two halves of S_2 in turn, with the results shown above. There is then consistent evidence pointing to $q = 5.4$, with the usual σ of about 1.5 ft. Notice that S_2 and each of the two 'mixed' data-sets involve only about 113

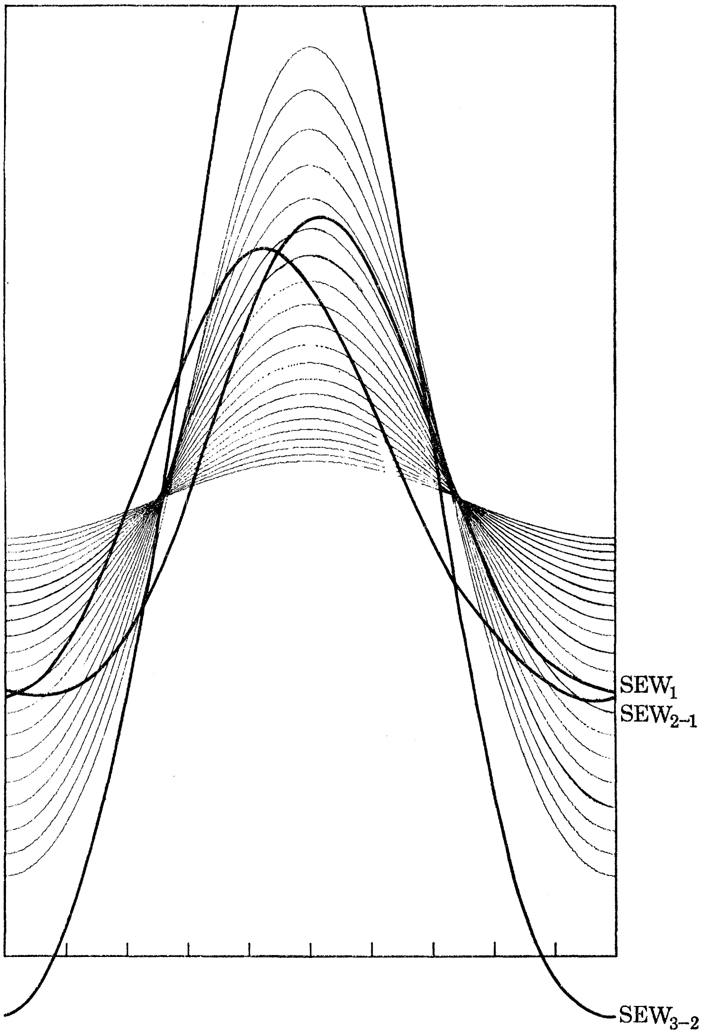


FIGURE A3. Periodic spline transforms for the 'residuals' in the cosine quantogram analysis of SEW_1, SEW_{2-1} and SEW_{3-2}, shown against a background of lumped Gauss distributions.

observations. Thus for comparison with figure 11 there would be a good argument for multiplying up the corresponding peak heights by the factor $\surd(169/113) = 1.223$, thus raising the least of them to 3.57 and the largest of them to 4.58.

As for EW_2 by itself, if we look only at the peak near the Thom quantum, the factor to allow for the small sample size (60) would be $\surd(169/60) = 1.678$, which would raise the peak height to only 2.85. My inclination is to suspect that the evidence for the quantum resides in the Scottish data *only*. If this be true, it is of some archaeological importance.

These results seem to indicate that the evidence for the quantum really resides in the Scottish data set S_2 of 109 observations. It is therefore desirable to repeat the c.q.g. analysis in full for this smaller sample, and to re-estimate the statistical significance of the result. Figure A4 shows a c.q.g. over the 'official' range $0.09 < \tau < 0.59$, from which it will be seen that the peak near 5.44 ft is still the dominant one. To permit a more accurate determination of peak height, the analysis was repeated with the narrower τ-range $0.13 < \tau < 0.23$, and the result is shown in the *upper* curve in figure A5. Here the peak occurs at $q = 5.435$ ft, and is of height 3.748 units; the estimated value of σ comes out to be 1.43 ft. On referring to our sample of 600 S values

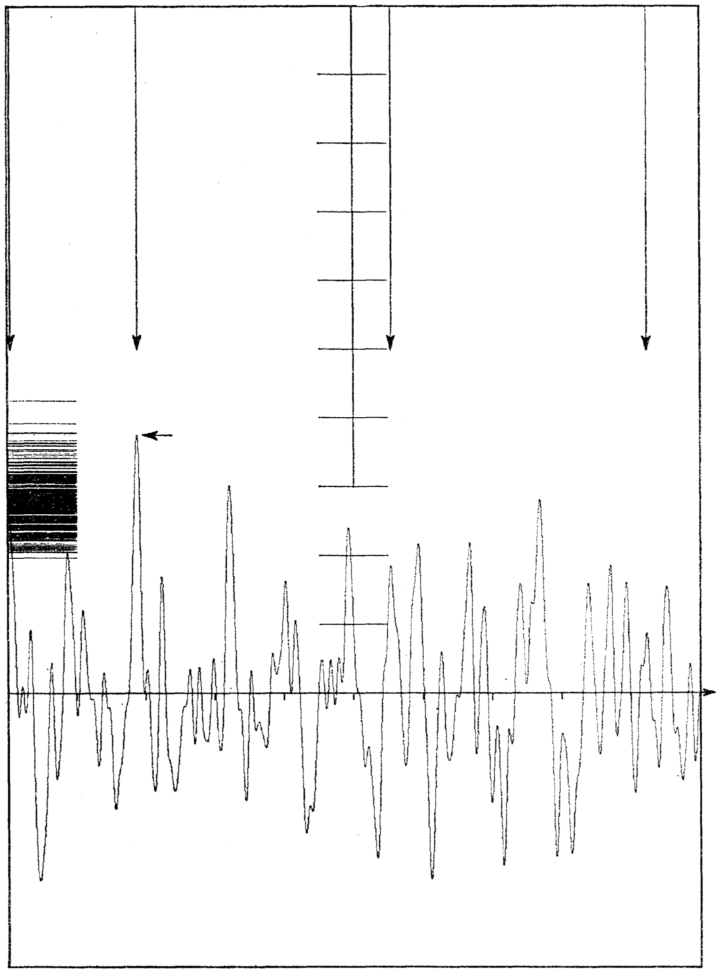


Figure A4. Cosine quantogram ($\tau_0 = 0.09$, $\tau_1 = 0.59$) for the 109 Scottish observations (S_2). The 200 simulated 'flood levels' are also shown.

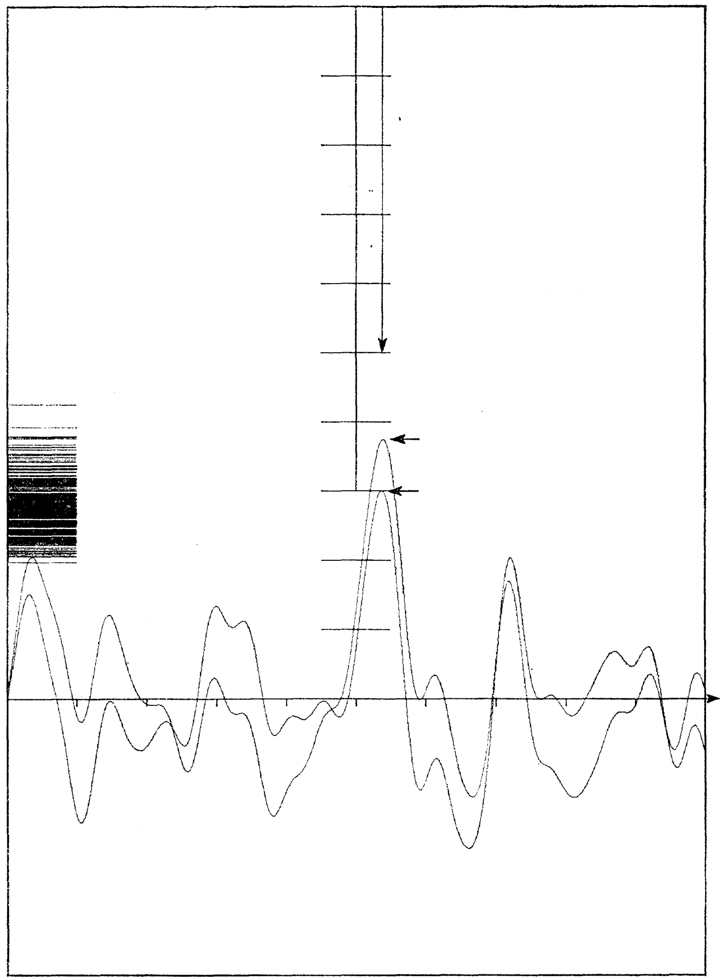


Figure A5. Cosine quantograms ($\tau_0 = 0.13$, $\tau_1 = 0.23$) for (i) S_2, and (ii) S_2 with the 11 'Huber' observations removed. The 200 simulated 'flood levels' are also shown.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

(whose distribution is, we know, very little affected by the value of N), we find that an S value of 3.748 is reached or exceeded just 13 times, so that the significance level is now 13/600 or about 2 %. That is, *even if we throw away the English and Welsh data*, we still find that Mount Thom demands to be taken seriously. Unfortunately 11 of Huber's 'knot' of observations in the interval (20.5, 22.0) belong to S_2 so that as before we have to exercise considerable caution in interpreting these results. We therefore look at what remains of S_2 when the 11 diameters belonging to H are removed from it, leaving us now with only $109-11 = 98$ observations. The corresponding c.q.g. is to be seen in the *lower* curve in figure A5. Here the peak height is 3.004, the location of the peak is at $q = 5.441$ ft, and $\sigma = 1.52$ ft. Once again the familiar values of q and σ are found, but the loss of the 11 observations near $4 \times 5.442 = 21.77$ ft has eroded the peak to a much lower value, 'typical' rather than 'extreme'.

At the same time we must bear in mind that we have 'thrown away 11 of our best observations'; the successive paring away of observations from the initial sample size of 169, even if there *were* a genuine quantal effect, would by the 'square-root law' which we have explained above necessarily cut down the height of Mount Thom to a value comparable to the noise level. This is seen best by noting that

$$4.0 \times \sqrt{(98/169)} = 3.046.$$

Exactly the same point is made in another way by noting the stability of the σ value. Those who consider that the 'quantal effect' is merely an artefact caused by the fortuitous congregation of about a dozen circle diameters near to 21.77 ft have to find some convincing explanation for the invariability of σ.

(v) Professor Huber and several others mentioned the possibility of splitting the data by circle size, and this is obviously worth doing, although we must (to preserve sample numbers) work with the complete data set SEW_2, and avoid trying to split by *both* circle size and geography. To get three nearly equal groups I took three sets of SEW_2 circles:

58 small, diameters from the smallest up to 35.6 ft;
55 medium, diameters from 37.6 ft up to 68.4 ft;
56 large, diameters from 69.1 ft up to the largest.

These gave the following results.

TABLE A4. EFFECT OF CIRCLE SIZE

group	N	height	q value	σ
small	58	1.590	5.371	1.672
medium	55	2.586	5.429	1.446
large	56	2.937	5.447	1.388

The stability of the quantum is quite remarkable, and no less so is (*a*) the stability of the order of magnitude of σ, and (*b*) the fact that σ if anything *decreases* as the circle size increases. As has already been explained in the text, variation in the length of the quantum from one measuring out of a quantum's length to the next, along a radius, would produce a $\sqrt{X}$ dependence in σ, while a variation of the quantum from one site to another would produce an X dependence in σ, yet we find neither. This confirms what I have already suggested; that the 1.5 ft value of σ merely reflects the inherent uncertainty in the location of the starting or end point of the measurements (due to the non-negligible size of the stones).

(vi) This brings us to the vexed question of 'pacing'. Here I have received a great deal of help from Dr H. J. Case, who has consulted his contacts in the Brigade of Guards and reports as follows.

'My drill-sergeant friend tells me that trained Guardsmen should be able to maintain a pace of 30 inches $\pm\frac{1}{2}$ an inch on parade, and even when off duty in the vicinity of the parade ground. For Infantry of the Line and Royal Artillery about $\pm 1\frac{1}{2}$ to ± 2 inches would be acceptable, depending on the degree of training. These figures are heel to heel whereas ours would be toe to toe presumably, but I do not see how that would make a difference. However these military exercises are on asphalt, whereas ours would be on tussocky ground at best and on some degree of slope; but rough ground might not greatly affect the performance of fit and trained men. Guardsmen may be trained by lines painted on the parade ground 30 inches apart and their pace is checked against the drill-stick, but a competent drill-sergeant will detect consistent deviations of ± 1 inch instantly by eye. . . . I would be inclined to think that ± 2 inches was about right for trained men on rough grass.'

This fascinating information left one immediate impression upon me; if guardsmen 'pace' so accurately, but are trained by a drill-stick, then in their case there is very little difference between 'pacing' and 'quantum'; the drill-stick *is* their quantum, even if not physically present but only as a threat, as it were. However, it seemed that it would be interesting, having simulated Professor Thom, to proceed to simulate the Brigade of Guards, and accordingly I modified my simulation program so that it would act as follows.

First it constructs an X-value using the simulator (5.3), and then it computes M, the nearest whole number of Thom-quanta of 5.442 ft thereto. It then adds together M versions of the Thom quantum, *each one of these*, however, being perturbed by an error of mean value zero and standard deviation 0.2357 ft. This last figure was chosen because it is $\sqrt{2}$ multiplied by 2 in. The result obtained from a c.q.g. based on a sample of 169 such diameters was:

$$\text{height} = 13.589, \quad q \text{ value} = 5.441 \text{ ft}, \quad \sigma = 0.674 \text{ ft}.$$

Thus 'pacing' is *far too accurate*, or rather, 'pacing' may have been used, just as a whale-bone rod may have been used, but we shall never know, because the residual errors associated with the sizes of the stones completely swamp errors of the kind associated with good pacing.

Incidentally the reader may be curious to know why one obtains an estimated σ of 0.674 ft; the answer is that $0.674/0.2357 = 2.86$, and the square of 2.86 is about 8. Now 8 quanta make up some 43.5 ft, which is of the same order as the average circle-diameter. Thus there is consistency among the figures, after all.

(vii) If we assume that the errors ϵ are *not* associated in their average magnitude with circle size (and to this conclusion we were driven by the discussion at (v) above), then it is of course of some interest to simulate a set of 169 circles with the ϵ values deliberately controlled in this way. To be specific, we use the randomizer (5.3) to obtain 169 circle diameters X, we compute the nearest integer $\leqslant X/q$, and if this integer is M we replace X by

$$X' = Mq + \sigma\epsilon,$$

where σ is at our choice and ϵ is a random standardized Gaussian random variable. Here $q = 5.442$ and σ is fixed throughout the calculation, but the ϵ values vary randomly from one circle diameter to the next and represent the errors 'associated with the sizes of the stones'. Figure A6 shows the resulting c.q.g over the range $0.13 < \tau < 0.23$ (this narrow range was

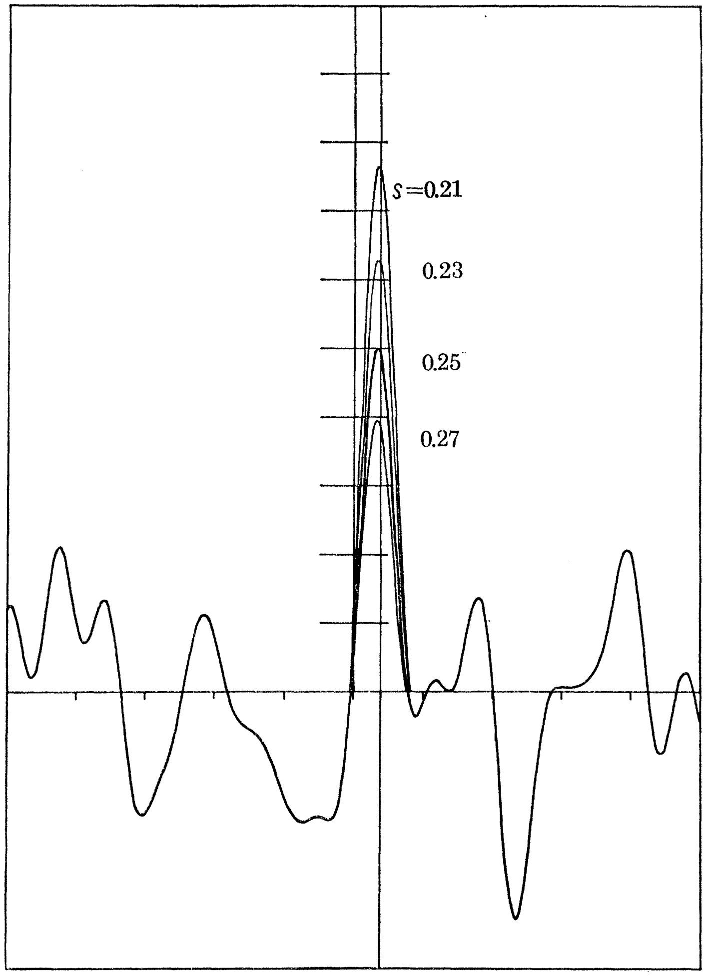


FIGURE A6. Cosine quantograms ($\tau_0 = 0.13, \tau_1 = 0.23$) for simulated *quantal* data with various values for $s = \sigma/q$.

chosen to show up the actual profile of the resulting peak); the four peaks shown have values of σ such that $s = \sigma/q = 0.21$, 0.23, 0.25 and 0.27. With $s = 0.27$ we will have $\sigma = 1.47$ ft, which is the 'observed' value, and it will be noticed that the simulated data then generates a peak of height 4 units, as of course it ought to do if our calculations and simulations are consistent with one another. It will be clear from figure A6 that the peak height is a very sensitive indicator of the true value of σ, so that we can have some considerable confidence in the value $1\frac{1}{2}$ ft which we have found for the latter. Incidentally it should be remarked that while the four peaks have been drawn separately, the c.q.g. throughout the rest of the range has here (for the sake of clarity) been drawn for one s value only. The differences elsewhere were very slight.

(viii) We now turn to a question which has excited considerable controversy. If we accept that there *is* a quantum, of value about $q = 5.44$ ft, how *accurate* is this estimate? The matter is of great importance when we compare estimates of the quantum from different geographical regions (e.g. Scotland, and Brittany). We are not concerned with that question here (apart from our discovery that the evidence for any quantum at all in England and Wales is pretty shaky), but the standard error to be attached to the figure 5.44 remains an important matter to which we must give some attention. The analysis which follows was carried out before the primary importance of the Scottish data was recognized, and so uses an N value of 169 instead of 109. The effect of this is not likely to be very great, and anyone who wishes to examine it could readily do so, either by an asymptotic error analysis, or by repeating the type of calculation used here, which seemed to me more suitable because intuitively more readily comprehensible by non-mathematicians.

The following procedure was carried out 25 times, each computation being quite independent of the others.

First, a sample of X values of size 169 was obtained using the simulator (5.3). For each X value, the nearest integer M to X/q (with $q = 5.442$) was computed, and X was then replaced by

$$X' = Mq + \sigma\epsilon,$$

where $\sigma = 1.507$ ft and where ϵ was a standardized Gaussian random variable, drawn 'fresh from the well' at each occurrence, i.e. for each diameter. Thus each X in the original sample was replaced by a 'target length' Mq which (it is supposed) it was desired to lay out, and this was then perturbed by an error $\sigma\epsilon$ of the appropriate size (arising from the finite size of the stones, etc). On computing (as was then done) a c.q.g. for this sample of X'-values, 169 in number, we would *not* of course expect to get a peak lying exactly at 5.442, because of the perturbing effect of the ϵ's. The fluctuations among the observed positions of the 25 peaks so obtained provides one with a direct indication of the inherent uncertainty in peak location by this method, in the circumstances of the actual SEW_2 analysis. If we let q^* denote the computed location of the peak, using the c.q.g., then we can summarize the experiment by saying that the 25 values for q^* lay in the range (5.41, 5.50). Their arithmetic mean value was 5.446 ft, and their standard deviation was 0.0181 ft. This last I take to be *the best available estimate of the uncertainty in our knowledge of the value of the quantum*. It is *appreciably* larger, by a factor of about 3, than estimates offered previously. I make no attempt here to account for the discrepancy, because I do not claim to understand the method by which the smaller estimate (0.006 ft) was obtained, although it seems clear that it was based on a data-set larger than the one we have allowed ourselves. I hope, however, that everyone will feel happier with a range of uncertainty of about 1 in. The reader should note carefully that this is the degree of uncertainty *in our knowledge of the average value of the quantum*; it has *nothing to do* with how far one quantum may have varied from the next, as they were successively laid out along a radius or diameter, and it has *nothing to do* with how far one quantum-standard in the northwest of Scotland may have differed from another in, say, Carnac. I make these rather obvious remarks because the discussion at the meeting concerning 'the standard deviation of the quantum' was bedevilled by some lack of comprehension of these important (and really quite elementary) distinctions.

(ix) I conclude with a few very brief remarks intended for the mathematical reader. They are deliberately brief because this paper is already much longer than it strictly should have been, and also because my theory and algorithms are both so elementary and naïve that any mathematician could readily fill in the gaps and repeat these computations for himself, on the same or different data, if he were so minded and could spare the time.

The first point to make is that we can *either* accept (4.1) as a natural functional statistic to work with (and this, on the whole, is the point of view I personally favour), *or*, if we feel we must have a 'reason' for what we are doing, one can be found within the framework of likelihood theory (Edwards 1972). If X is an observed circle diameter, we write $X = (M+Y)q$ (where q is the quantum); here M is a random integer and Y is a random variable in the interval $(-\frac{1}{2} < Y < \frac{1}{2})$, and we take $2\pi Y$ to have the von Mises distribution with parameter k. As I have already explained, this is effectively equivalent to assuming a lumped Gaussian distribution for $\theta = 2\pi Y$, and that perhaps is how we should think of the matter in the first instance, but the switch to the von Mises distribution is analytically permissible (because of the famous empirical 'parrot' effect) and for the purposes of maximum likelihood it is a change for the better. We can have a β effect (in the sense of Thom) if we write $X = c+(M+Y)q$, where $0 \leqslant c < q$, and this will merely induce a phase shift of amount $\beta = 2\pi c/q$ radians into the von Mises distribution. For the sake of generality I will leave that effect in, for the moment.

The likelihood of the observations can now be written in the form

$$\prod_{j=1}^{N} p_{[X_j/q]}\, K(k)^N \exp\{kN(A\cos\beta + B\sin\beta)\},$$

where $[x]$ denotes 'the integer closest to x', $K(k)$ is the function in (3.2) (simply expressible in terms of modified Bessel functions), and

$$A = (1/N)\,\Sigma\cos\theta_j, \quad B = (1/N)\,\Sigma\sin\theta_j$$

(here θ_j is $2\pi Y_j \pmod{2\pi}$). The probabilities p_m which occur within the product represent the probability that a length mq will be 'required' by the 'architect'.

The maximum likelihood estimate of β can be obtained from

$$A = C\cos\hat{\beta}, \quad B = C\sin\hat{\beta},$$

where C is the non-negative square root of A^2+B^2. If we carry out this preliminary maximization, we have next to maximize with respect to $k > 0$; it turns out that there *is* a unique maximum, given in fact by

$$I_1(\hat{k})/I_0(\hat{k}) = C,$$

or by

$$\hat{k} = 2C(1+\tfrac{1}{2}C^2)$$

when, as is normal, C is small. Substitution of this value for k then gives the conditionally maximized likelihood

$$\prod_{j=1}^{N} p_{[X_j/q]} \exp\{NC^2\}.$$

We now have to compare this with the likelihood on the non-quantal hypothesis,

$$f(X_1)\,f(X_2)\,\ldots f(X_N), \text{ say.}$$

But for q not too small, we shall have, approximately,

$$\prod_{j=1}^{N} p_{[X_j/q]} = q^N . f(X_1)\,f(X_2)\,\ldots f(X_N),$$

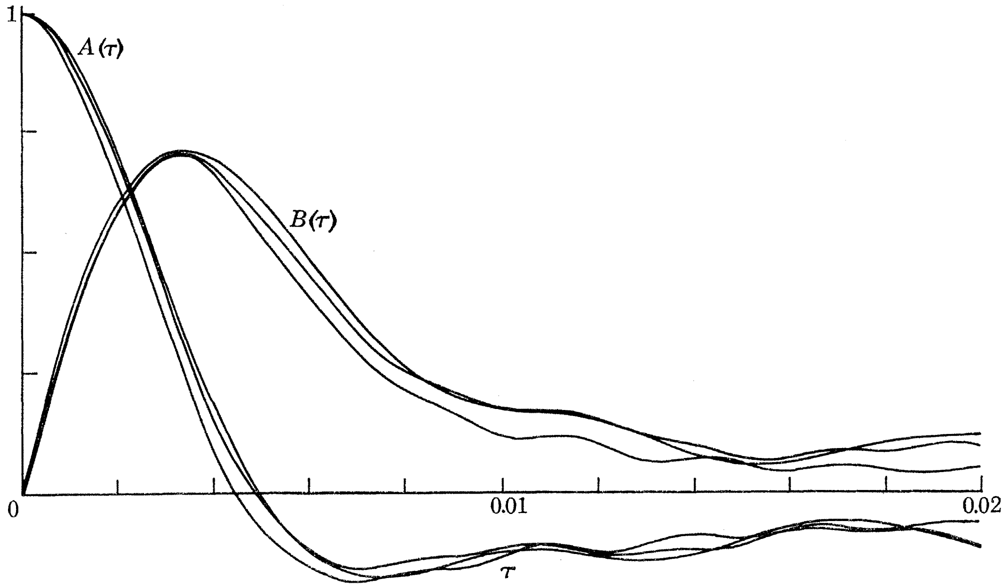
and so, constants apart, *the relative support for a value* $\tau = 1/q$ will be

$$\mathcal{S} = \tfrac{1}{2}\{\phi(\tau)^2 + \psi(\tau)^2\} - N\ln\tau.$$

We recall that we shall only be interested in the behaviour of this function over the range $\tau_0 < \tau < \tau_1$ agreed in advance.

We now observe that, when we are far enough away from $\tau = 0$, the stochastic processes $\{\phi(\tau)\colon \tau_0 < \tau < \tau_1\}$ and $\{\psi(\tau)\colon \tau_0 < \tau < \tau_1)\}$ are effectively stationary Gaussian stochastic processes with zero mean and unit variance, and what is more, their auto-correlation functions and cross-covariance functions can all be estimated (at least for small u) in terms of $A(u)$ and $B(u)$ (where here we recognize the dependence of A and B on $u = 1/q$). Figure A7 shows the

(A 7)



(A 8)

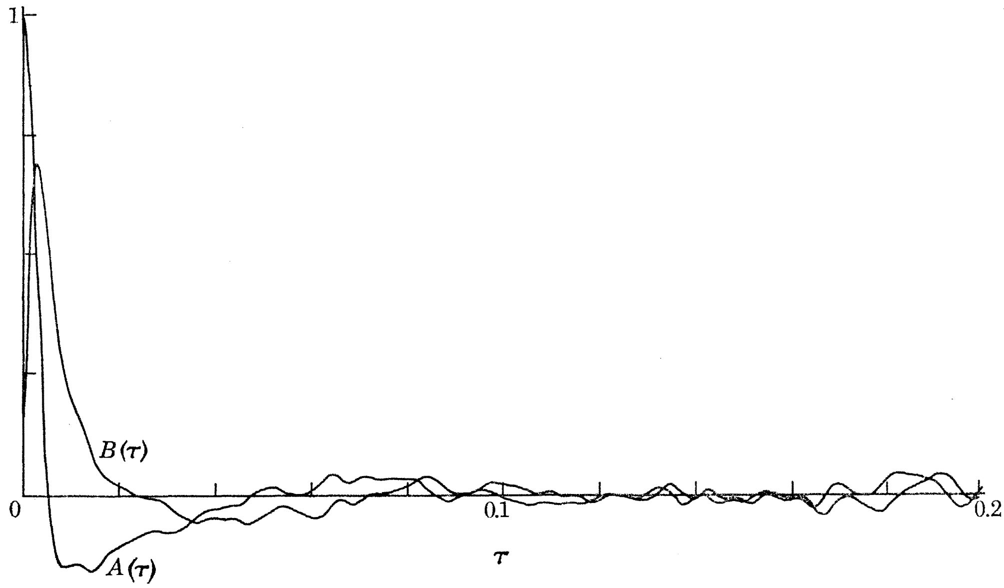


FIGURE A7. Estimates (based on $SEW_{1,2,3}$) of the autocorrelation function $A(\tau)$ for each of the ϕ and ψ stochastic processes, and of their cross-covariance function $B(\tau)$.

FIGURE A8. This is similar to figure A7, but the computations are now based on simulation-sets of size 2000, and extend over a wider range (over which estimates based on the small real data-sets would be quite unreliable).

A and B functions for $0 < \tau < 0.02$ for each of $SEW_{1,2,3}$, while figure A8 shows the A and B functions for $0 < \tau < 0.2$ for very large samples generated by the simulator (5.3). It will be seen that both functions fall rapidly to zero; the peak of the B function near $\tau = 0.004$ corresponds to a systematic lag between the stochastic processes ϕ and ψ, which are otherwise essentially independent of one another.

From these considerations we learn that

$$U(\tau) = \tfrac{1}{2}\{\phi(\tau)^2 + \psi(\tau)^2\}$$

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

is a stationary stochastic process each of whose individual ordinates has approximately a negative-exponential distribution with unit parameter. An initial analysis, by what I call the modular quantogram (m.q.g.), consists in making a computer plot of $U(\tau)$ against τ over the appropriate τ range, and looking for peaks which are large in relation to the negative-exponential noise. These must then be assessed more carefully, using a Monte Carlo study of the supremum-functional for the U process (I have not done this), and if there are competing significant peaks they can then be compared on the basis of the $\$$ values. The last step is facilitated in my program by the drawing of constant $\$$ loci (isopithons) on the computer output; all this is done automatically, of course. In my version the isopithons are computed from the approximate formula given above, but it would be a simple matter to compute them exactly if a Bessel function sub-routine is available.

If a significantly high peak on the m.q.g. is found, the β-effect must next be tested for. This is done by testing the value of $\psi(\tau)$ at the peak τ as a standard Gaussian variable; if it lies within the range ± 2 we reject the β-effect as an unnecessary complication, and this is what happened with the Thom data.

We then go right back to the beginning and set $\beta = 0$ (no β maximization is now called for), and then proceed as before. This time we find that

$$\begin{aligned} \$ &= \tfrac{1}{2}\phi(\tau)^2 - N \ln \tau, \quad \text{when} \quad \phi(\tau) > 0, \\ &= -N \ln \tau, \quad \text{when} \quad \phi(\tau) \leqslant 0. \end{aligned}$$

Accordingly we now have to examine the plot of $\phi(\tau)$ against τ over the agreed τ range, that is, we have to look at the c.q.g., and the formula for the isopithons becomes

$$\phi = \sqrt{\{2(\$ + N \ln \tau)\}}$$

for so long as the expression within the square-root is non-negative. Detailed study of the argument shows that when the isopithon hits the τ axis, then it is to be continued downwards in a vertical straight line. (Of course this last detail is of no importance, because we shall never consult the isopithons unless we have to discriminate between significantly high positive peaks of ϕ-value 3.5 or more.)

It will not escape the reader that a peak at $\tau = \tau^*$ must imply a peak of sorts at $\tau = 2\tau^*$, $3\tau^*$, and so on for higher harmonics. I wasted a lot of time (before the likelihood argument was evolved) in working with a linear geometrically discounted combination of peaks and their harmonics, like

$$(1-\rho) \sum_{m \geqslant 0} \rho^m \phi(m\tau).$$

This is *possible*, because the series can be summed in finite terms and so computer-plotted, but I found it quite futile, and I am now satisfied that there is no good theoretical reason for doing anything of the sort. Professor Whittle pointed out to me that if one used a lumped Gaussian distribution for the ϵ *and neglected all but the central* (‘*unlumped*’) *term* (without which approximation the likelihood argument gets stuck right at the start), one winds up with Broadbent's 1955–6 functional statistic (see formulae at the foot of p. 50 in Broadbent (1955)). This can also be regarded as a linear discounted combination of the c.q.g. and its harmonics, but very curiously the ‘weights’ in Broadbent's linear combination alternate in sign, which seems not natural to me; also they tend to zero very slowly indeed. This did not matter to him because his series, too, can be summed, and it was in that form that he used it. Those who wish to compare and contrast Broadbent's work with mine should be warned that on his typical diagram it is the *minima*, and not the maxima, which have to be looked for. He used, as I do, a Monte Carlo basis for assessing significance, and proposed a partial ordering in the plane to assist in the comparison of rival significant peaks. This last device I do not need, because the isopithons essentially play that role, if it be needed at all.

The last mathematical topic on which I should like to touch very briefly concerns the asymptotic formula of Cramér (Cramér & Leadbetter 1967) for the limiting distribution of the supremum of a section of a 'well-behaved' Gaussian stationary standardized stochastic process. This formula is very beautiful, but so far as I am aware no one in the world has the slightest idea how accurate it is in finite, that is, non-limiting, circumstances. Our Monte Carlo calculations give us an opportunity to throw a first gleam of light on this matter, and it is obviously an opportunity not to be missed.

If we write $P(s)$ for the probability that the supremum random variable S will have a value *s or less*, then Cramér's limit theorem asserts that

$$\lim_{s\to\infty} P(s) = \exp(-e^{-z}),$$

where his z is defined in such a way that, in our notation,

$$e^{-z} = (\tau_1-\tau_0)\ \text{r.m.s.}\ (X)\ e^{-\frac{1}{2}s^2},$$

and it is supposed that $(\tau_1-\tau_0)$ is being allowed to increase, as s increases, at such a rate that z is held fixed. This is not, of course, a very helpful statement when we come to apply the formula! A word of explanation is required about r.m.s. (X); this means, as the notation indicates, the root-mean-square value of the circle diameters for the sample (or more strictly, for the coarse-grained distribution from which they are drawn). In Cramér's formula what appears is in fact the absolute value of the second derivative of the autocorrelation function at the origin, which is equal to the expectation of $(2\pi X)^2$. It is interesting to note that the sample-size N nowhere appears in the Cramér formula, and that the only trace of the identity of the data-set used is in the value of r.m.s. (X). For this I have used the value 65.433 ft which was computed from (5.3); that for S_2 would have been 65.4 ft, while that the SEW_2 would have been 72.2 ft. With these explanations the reader will be able to adapt the arguments and charts which follow to suit other data-sets and other values of τ_0 and τ_1.

Now while we do not expect Cramér's formula to be valid in our particular circumstances, because we are not (*and never will be*) in the limit situation, we may none the less reasonably expect that its general analytic form may indicate a helpful way of graduating the information gained empirically from our Monte Carlo experiments using the simulator (5.3), for which r.m.s. (X) = 65.433 ft. To be specific, we can usefully look at the experimental data we have accumulated in this way and see if there is any sign of a linear relation between the transformed variables

$$x = 2000\,(2\pi)^{-\frac{1}{2}}\,e^{-\frac{1}{2}s^2},$$

and

$$y = 20\ln(1/P(s)),$$

where we estimate $P(s)$ by

$$(\text{number of empirical } S \text{ values} \leqslant s)/600.$$

If Cramér's formula were to hold exactly, then we should have the relation

$$y = \{\sqrt{(2\pi)}\times(65.433/100)\times(\tau_1-\tau_0)\}\,x,$$

or in numerical terms,

$$y = 0.820\,x.$$

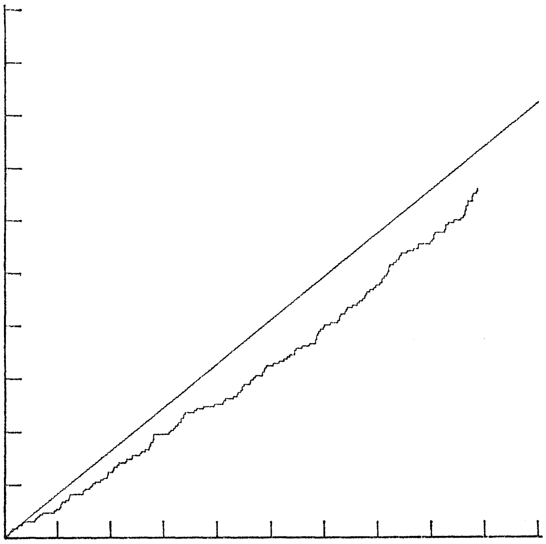


FIGURE A9. The (y, x) plot corresponding to the simulated sample of 600 S values, compared with the straight line which represents the Cramér limit. On the horizontal scale, x = 2000 times the 'standardized Gaussian density' at location s, while $y = 20 \ln (1/P(s))$, where $P(s) = \text{pr}(S \leqslant s)$, and natural logarithms are used. The diagram relates to the situation $\tau_1 - \tau_0 = 0.5$, and r.m.s.$(X) = 65.433$. In other situations the ordinates should be increased in proportion to $(\tau_1 - \tau_0)$ r.m.s. (X).

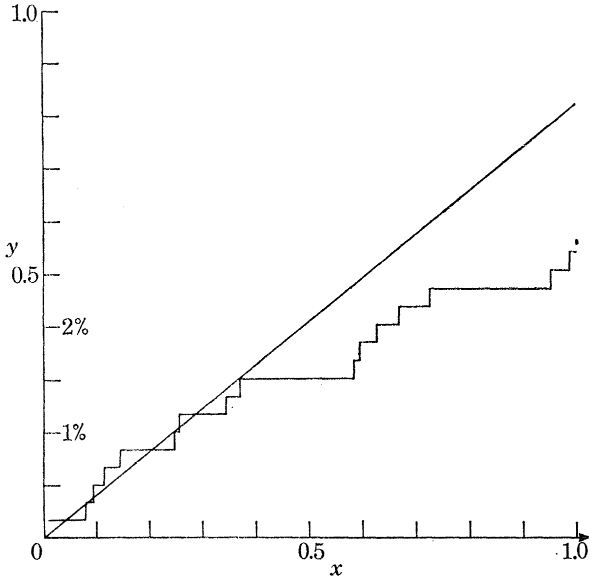


FIGURE A10. This is a tenfold enlargement of the bottom left-hand corner of figure A9. The scales are as before, but significance levels $100(1-P)$ are now indicated.

Figure A9 shows that the actual plot of y against x falls below this line, but approaches it very closely near $x = y = 0$ (which is the 'limit situation'). The diagram covers a range 0.00–8.86 for x, and a range 0.00 to 6.57 for y; thus at the top right-hand end of the stepped curve we have $P(s) = \exp(-0.3285) = 0.72$, so that the diagram in fact represents the upper 28 % of the S distribution.

Figure A10 shows the situation near $x = y = 0$ enlarged tenfold. In this diagram the top right-hand end of the stepped curve corresponds to $P(s) = \exp(-0.02703) = 0.9733$, so that the enlarged diagram represents the upper 2.7 % of the S distribution. It is this last chart which is most likely to be useful in practice. If the stepped curve is used to estimate significance levels, care must be taken to expand its *ordinates* in proportion to r.m.s. (X) $(\tau_1 - \tau_0)$ (the plotted ordinates correspond to 65.433 (0.59–0.09) = 32.7165 for this parameter). Thus if we wanted to work with the doubled range, $\tau_0 = 0.09$, $\tau_1 = 1.09$, we should *double* the ordinates of the stepped curve, thus effectively *squaring* the estimated value of $P(s)$ (as is obviously reasonable).

Finally, note carefully that the 'significance level' is not $P(s)$, but $1 - P(s)$. The following table will be found useful in connexion with figure A10.

TABLE A5. SIGNIFICANCE LEVELS AND THE (x, y) DIAGRAM

$1 - P(s)$ (%)	0.1	0.5	1.0	1.5	2.0
y	0.020	0.100	0.201	0.302	0.404

It will be seen that with our parameter-values, the Cramér limit formula is (perhaps by chance) astonishingly accurate for significance levels of 1 % or better.

In conclusion I wish to thank Miss Mary Brooks for assistance with the computations and the preparation of the diagrams. Some of the latter (e.g. figures 4 and 5) have been prepared directly from computer-plotter output, by Messrs E. L. Smith and R. S. Hammans of the Department of Physical Chemistry, University of Cambridge, who will be pleased to deal with any enquiries relating to their technique. I am very grateful to them for their help, and also to the Director of the Computer Laboratory, University of Cambridge, for a most generous allocation of 'space' and 'time' to this investigation.

References (Kendall)

Boneva, L. I., Kendall, D. G. & Stefanov, I. 1971 Spline transformations: three new diagnostic aids for the statistical data-analyst. *J. R. statist. Soc.* B **33**, 1–70.

Broadbent, S. R. 1955 Quantum hypotheses. *Biometrika* **42**, 45–57.

Broadbent, S. R. 1956 Examination of a quantum hypothesis based on a single set of data. *Biometrika* **43**, 32–44.

Cramér, H. & Leadbetter, M. R. 1967 *Stationary and related stochastic processes.* New York: John Wiley.

Edwards, A. W. F. 1972 *Likelihood.* Cambridge University Press.

von Mises, R. 1918 Über die 'Ganzzahligkeit' der Atomgewichte und verwandte Fragen. *Phys Z.* **19**, 490–500.

Stephens, M. A. 1963 Random walk on a circle. *Biometrika* **50**, 385–390.

Thom, A. 1955 A statistical examination of the megalithic sites in Britain. *J. R. statist. Soc.* A **118**, 275–295.

Thom, A. 1967 *Megalithic sites in Britain.* Oxford: Clarendon Press.

MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES
THE ROYAL SOCIETY A
PHILOSOPHICAL TRANSACTIONS OF