

19. Fourier analysis of field boundaries

John Peterson

School of Information Systems, University of East Anglia, Norwich NR4 7TJ, U.K.

This paper describes how Fourier analysis can be applied to the distribution of field boundaries of cadastral systems. It considers an example of current practice and, by means of a case study, seeks to show how the technique may be enhanced and used in conjunction with other sources of evidence.

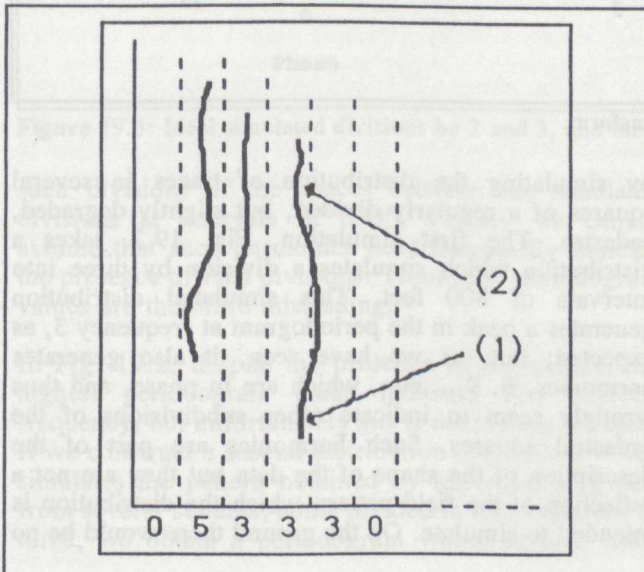
An initial study of the application of Fourier analysis forms part of a recent work by Rita Compatangelo (Compatangelo 1989). Her subject is a Roman cadastre in the Salentine peninsular, at the south-eastern extremity of Italy. This is a typical centuriation, with the major land divisions — which are often access ways — forming a grid of squares whose sides have a length of 2,400 Roman feet (*pedes monetales*). This 'module' is in this case equivalent to 705m.

In certain parts of this cadastre there are also many other existing boundaries at the same orientation as the grid, which may be the remains of regularly planned internal subdivisions of the squares. Compatangelo looks for signs of such regularity because this may give clues to this cadastre's date and function, by allowing comparison with other better known Roman cadastres.

Compatangelo's approach is to generate periodograms — charts which show the relative importance of frequencies contributing to the observed pattern. Frequencies with large amplitude may reveal underlying regularities in the field pattern, even in the presence of 'noise' arising from later modifications.

In the analysis of discrete data, periodograms are obtained by generating Fourier transforms of n values which represent the distribution of the data within some fundamental interval. In this case the numbers are derived from a figure proportional to the length of boundaries falling in a series of equally spaced bands

Figure 19.1. Generation of raw data from possible cadastral traces.



parallel to one orientation of the cadastral grid, see Fig. 19.1. In order to compensate for degradation of the traces, which may have become irregular or been wiped out, data is summed from the corresponding bands of several grid squares. The fundamental interval is the module of the centuriation.

The Fast Fourier transform (FFT) takes this real vector z and creates a vector c whose values are the complex coefficients of the discrete Fourier transform of z . In the case of the MathCAD FFT routine used by the author, these coefficients satisfy

$$C_j = \left(\frac{1}{\sqrt{n}}\right) \cdot \sum_k z_k e^{2\pi i j k / n}$$

FFT routines will produce $(n/2) + 1$ values of c , and the periodogram is the set of $|c_j|$. The periodogram values are proportional to the amplitudes of cosine curves, with phase $\arg(c_j)$, which, when added together, generate the distribution represented by the vector z . A cosine curve with phase zero will have maxima at the ends of the fundamental interval.

This is equivalent to the form of representation (cf. Rayner 1971) in which any wave form is generated by the sum of a series of sinusoids with frequencies 0, 1, 2, 3 etc., which have suitably chosen amplitude and phase. The periodogram values show the relative importance of the frequencies, and the corresponding phases measure the degree to which the cosine wave at that frequency is 'in step' with the fundamental interval.

Fig. 19.2 shows an example of FFT using the MathCAD package. The distribution of features (upper chart) is represented by 64 values, held in a vector z . The periodogram (centre chart) is the distribution of the amplitudes of the 32 cosine curves, of frequencies 1 to 32, which contribute to the spatial distribution (the amplitude of frequency zero is not shown because it has no information to provide about frequencies). The corresponding phases (lower chart) are expressed as values from $-\pi$ to π . Those frequencies which are more nearly in phase are represented by points nearer to the central line. In this example these frequencies are 3, 6, 9, etc.

In Compatangelo's analysis the raw data take the form of possible traces at one of the cadastral orientations, summed over 20 squares. She divides the fundamental interval of 705m into 44 bands and produces three measures for each band: the sum of the length of traces, the number of occurrences of a trace and the ratio of these two. The periodograms of these discrete distributions have peaks which may represent the sought-for original divisions. For example, Compatangelo's Figure 53 shows highest amplitudes at frequencies of 2, 6, and 20, and possibly at 7 and 11. She appears to consider all of these to be potentially

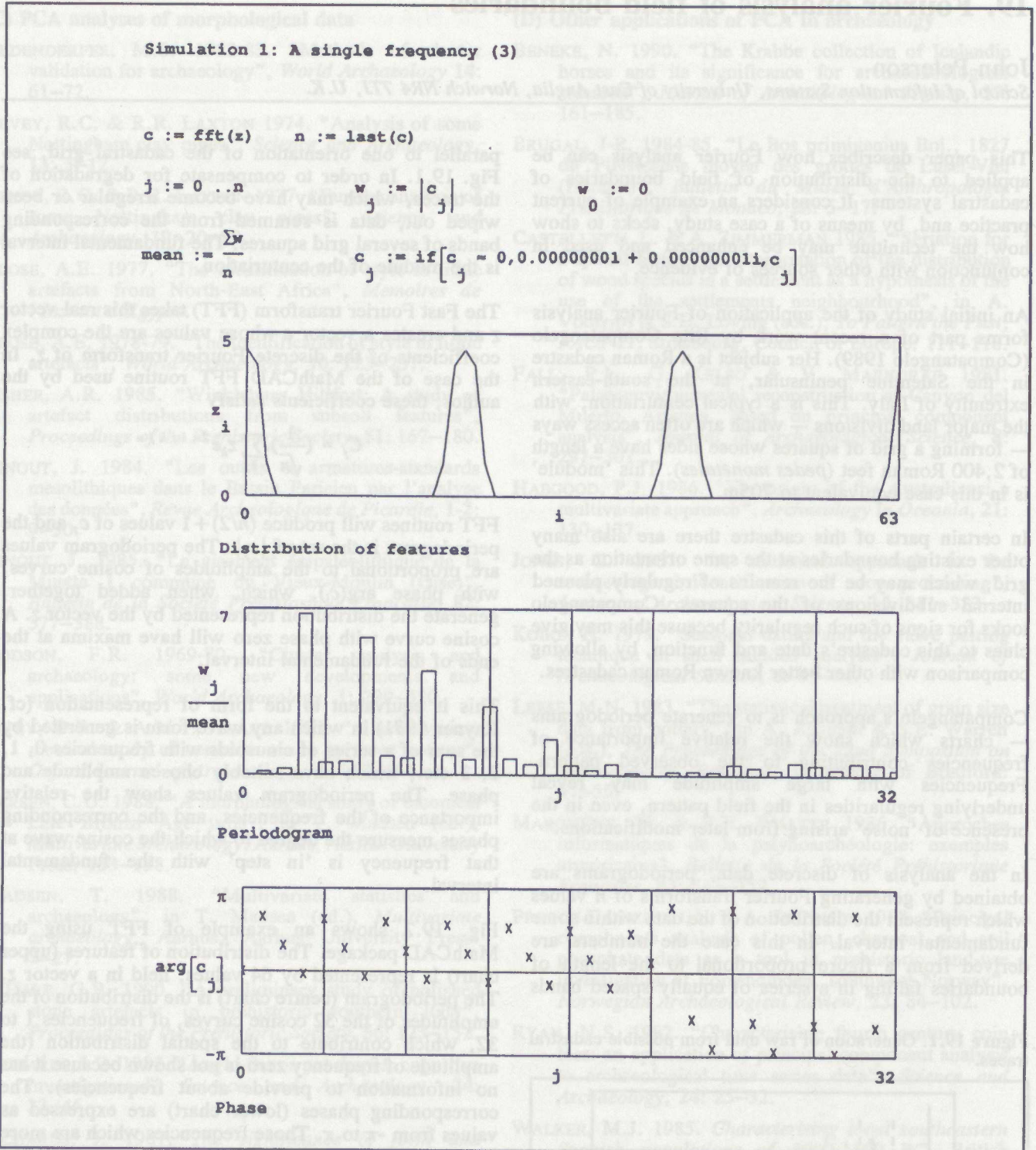


Figure 19.2: An ideal simulated division by 3 and its Fourier transform.

meaningful.

This work is most interesting and original, but it gives rise to some questions. Do we need some form of 'control'? What would we expect to see if we submitted a regularly divided cadastre to this process? Do we need to assess the statistical significance of a particular peak value, and how can we do so? What is the significance of phase?

These problems were approached with the aid of the MathCAD package running on an IBM PS/2 70. The first question (what do we expect to see?) was tackled

by simulating the distribution of traces in several squares of a regularly divided, but slightly degraded, cadastre. The first simulation, Fig. 19.2, takes a distribution which simulates a division by three into intervals of 800 feet. This simulated distribution generates a peak in the periodogram at frequency 3, as expected; but, as we have seen, it also generates harmonics, 6, 9, ..., etc., which are in phase, and thus wrongly seem to indicate other subdivisions of the cadastral squares. Such harmonics are part of the description of the shape of the data but they are not a reflection of the field pattern which the distribution is intended to simulate. On the ground there would be no

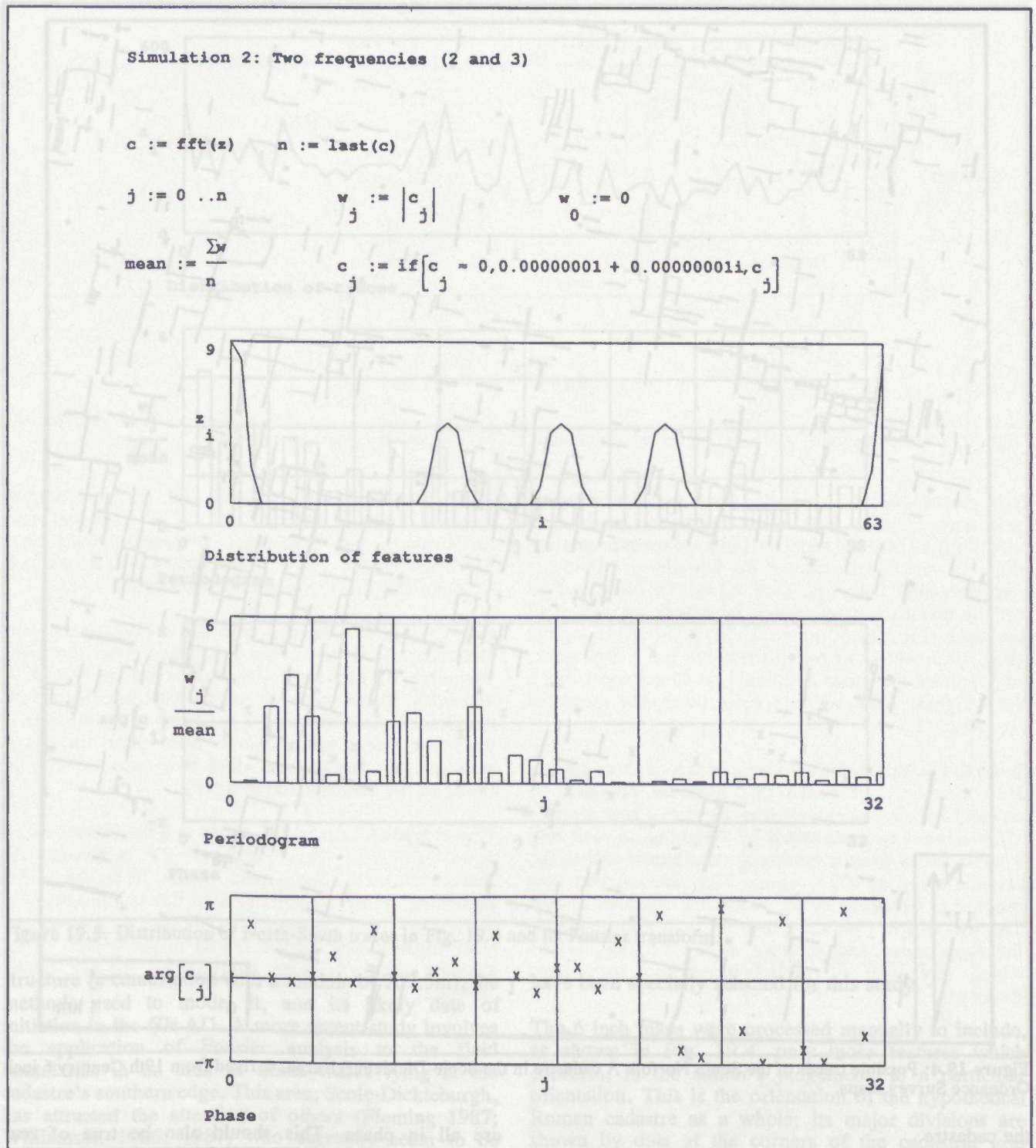


Figure 19.3: Ideal simulated divisions by 2 and 3, and their Fourier transforms.

such divisions of the spaces between the simulated divisions at intervals of 800ft. Clearly we cannot assume that such harmonics need necessarily indicate the presence of such divisions. These high periodogram values are therefore misleading.

In Fig. 19.2, despite the presence of harmonics, the highest periodogram value indicates the 'correct' frequency, but unfortunately this is not always the case. If we construct a second distribution (Fig. 19.3) which simulates the pattern obtained by superimposing data from several squares, some divided in two and some in three, we obtain a periodogram whose highest value

is 6. Again this is an accurate reflection, in terms of Fourier analysis, of the distribution, but it is misleading. A division by 6 was not part of the original simulated distribution of boundaries. This indicates that the appearance of non-prime frequencies which are in phase and which have large amplitude values in the periodogram need not necessarily indicate the presence of corresponding divisions on the ground.

So, the analysis of simulated regular divisions shows that harmonics are to be expected; we should not too readily assume that a prominent peak in the periodogram necessarily reveals a physical division of

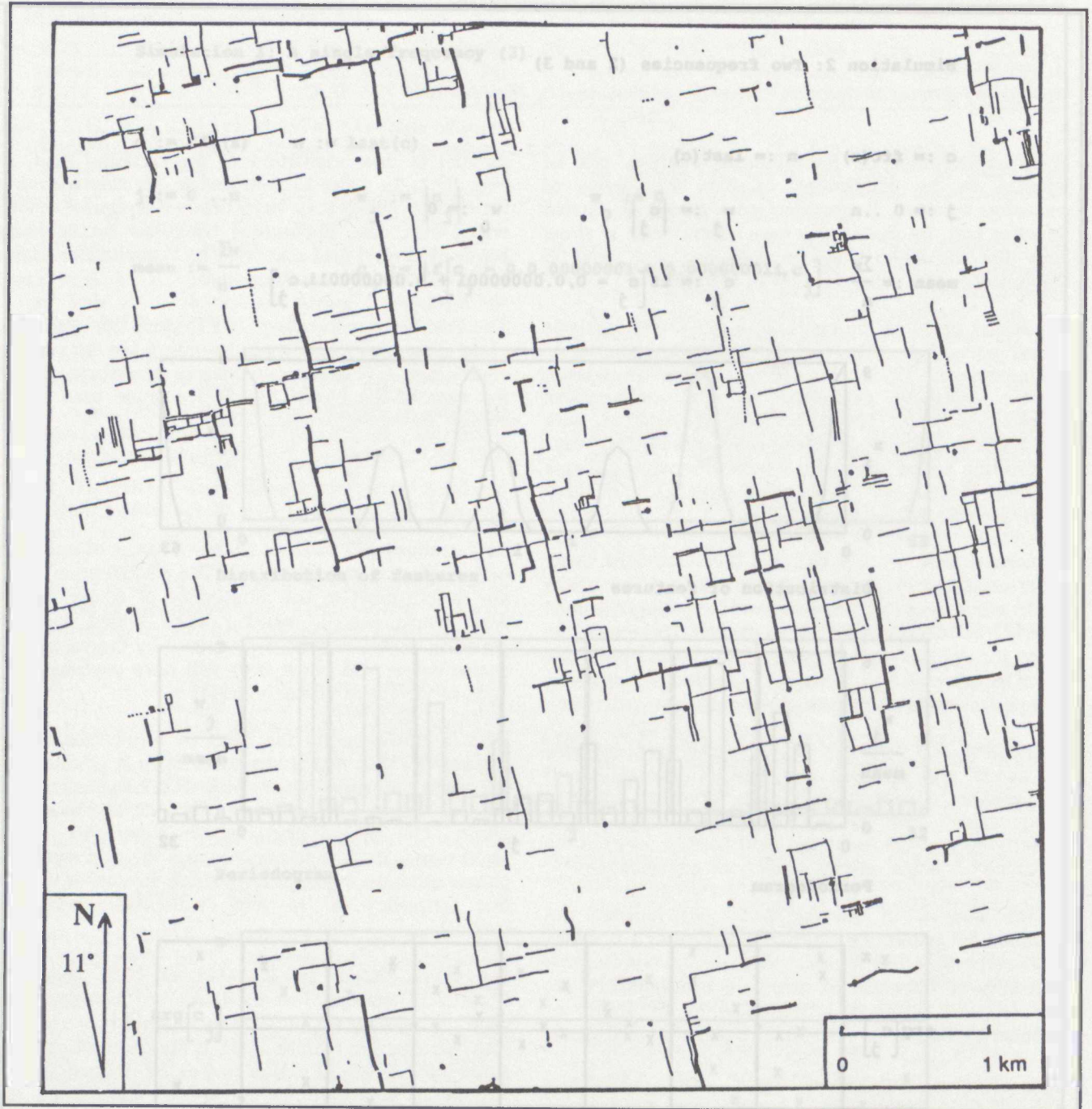


Figure 19.4: Possible traces of the South Norfolk A cadastre in the Scole-Dickleburgh area, derived from 19th Century 6 inch Ordnance Survey maps.

the cadastre.

Nevertheless it could be worth investigating the physical significance of such high peaks, provided that we had some measure of their statistical significance. We need to know how often a peak of a given height would occur in a periodogram derived from purely random data. This problem can be tackled analytically, by the use of Fisher's g statistic (Priestley 1981:406-9), but another approach is possible, using a computer-based Monte-Carlo method which will shortly be described in the context of a case study.

Figs. 19.1 and 19.2 also reveal a third factor in the analysis. This is the phase of the component frequency. Note that the prominent frequencies and harmonics in the periodograms derived from simulated distributions

are all in phase. This should also be true of real distributions with real sub-divisions. A peak in the periodogram may be sufficiently high to be considered significant, but if the phase angle is too far from zero it is then clear that the frequency cannot possibly represent a genuine sub-division of the fundamental interval.

So we can go further than Compatangelo by considering the three additional factors which have been outlined, i.e. harmonics, statistical significance and phase. A case study may indicate the sort of analysis that the author has in mind.

A suspected Roman cadastre in South Norfolk has been under investigation for several years. Earlier publications (Peterson 1988a, 1988b) have indicated its

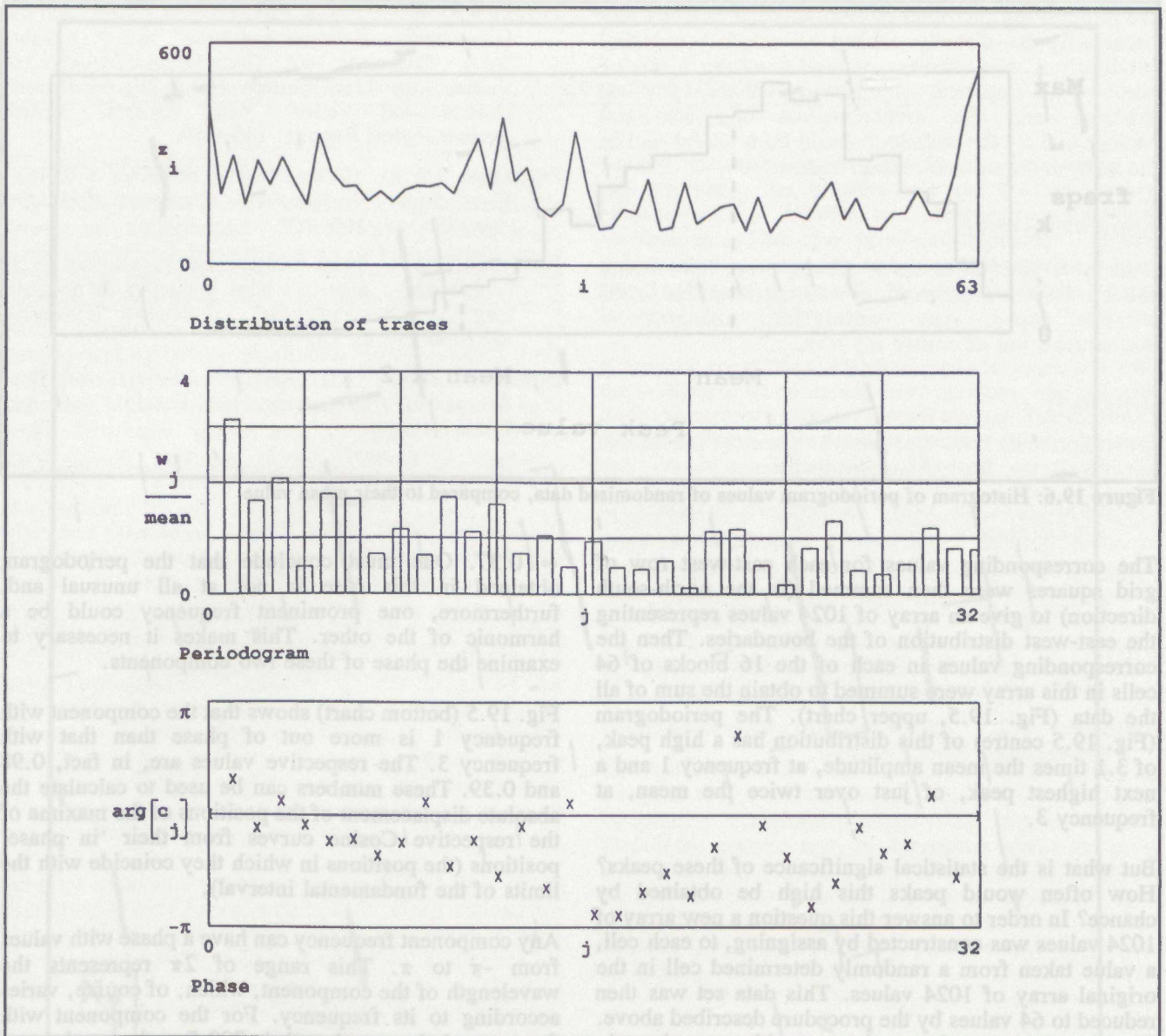


Figure 19.5: Distribution of North-South traces in Fig. 19.4 and its Fourier transform.

structure (a centuriation with a module of 709.5m), the methods used to model it, and its likely date of initiation in the 60s AD. A more recent study involves the application of Fourier analysis to the field boundaries in an area of dense traces lying at the cadastre's southern edge. This area, Scole-Dickleburgh, has attracted the attention of others (Fleming 1987; Williamson 1986, 1987) who currently seem to be convinced that the pattern of boundaries derives from a coaxial field system which was certainly in existence in the Romano-British period, before the construction of a main Roman road which crosses the cadastre obliquely, and which could have had a prehistoric origin.

The area covered by the supposed coaxial field system was used for the study described here. There were two reasons for this. Firstly, its boundaries were determined independently, by Williamson; secondly, the source data, the Nineteenth Century 6 inch Ordnance Survey maps, are the starting point for both Williamson's analysis and that described here. It would thus be difficult to claim that the survey area or the basic data

have been specially selected for this study.

The 6 inch maps were processed manually to include, as shown in Fig. 19.4, only those features which conform, in the author's judgement, to a particular orientation. This is the orientation of the hypothetical Roman cadastre as a whole; its major divisions are shown by dots at the corners of the cadastral grid squares, forming lines at a constant angle slightly greater than 11° to the west of OS grid north (and at right angles).

In this area the possible cadastral traces running in this direction (the 'north-south' direction) are more numerous and regular, so they were used as input to the Fourier analysis. Since the MathCAD Fast Fourier Transform routine expects the number of input values to be a power of two, the north-south traces in each hypothetical grid square were enlarged to fit a transparent overlay divided into 64 bands. The length of trace lying in each band was measured, by hand, to give 64 values for each square, giving a total of 808 non-zero values over approximately 100 squares.

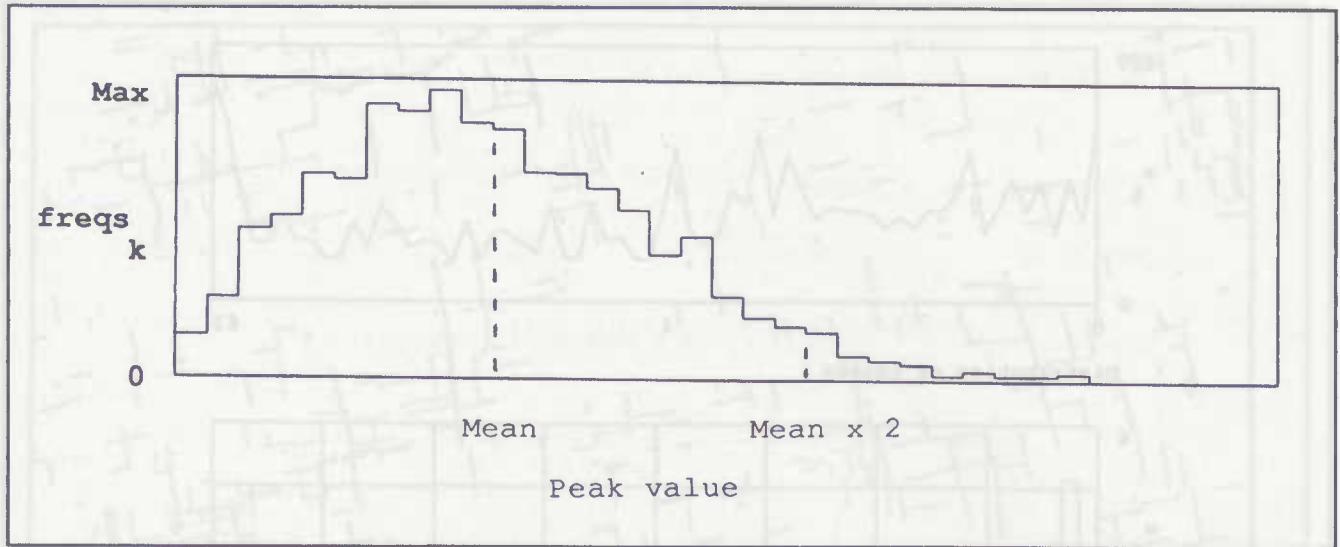


Figure 19.6: Histogram of periodogram values of randomised data, compared to their mean value.

The corresponding values for each east-west row of grid squares were then summed (in the north-south direction) to give an array of 1024 values representing the east-west distribution of the boundaries. Then the corresponding values in each of the 16 blocks of 64 cells in this array were summed to obtain the sum of all the data (Fig. 19.5, upper chart). The periodogram (Fig. 19.5 centre) of this distribution has a high peak, of 3.1 times the mean amplitude, at frequency 1 and a next highest peak, of just over twice the mean, at frequency 3.

But what is the statistical significance of these peaks? How often would peaks this high be obtained by chance? In order to answer this question a new array of 1024 values was constructed by assigning, to each cell, a value taken from a randomly determined cell in the original array of 1024 values. This data set was then reduced to 64 values by the procedure described above. The aim was to produce a data set with approximately the same mean and variance as the original, but in which any sign of periodicity would have been destroyed.

This gives a new distribution and a different periodogram of 32 values (that for frequency zero not included). This randomisation process was repeated 100 times and the 3,200 periodogram values were plotted as a histogram (Fig. 19.6) and tabulated as a cumulative percentage distribution. This indicates that, in this case and for this sort of data, none of the randomly generated values exceeded three times the mean value, but 4% are higher than twice the mean value. At first glance 4% looks quite a low chance; but, given 32 values in each periodogram, it is likely that at least one will exceed this level. If we use the value of 4%, then the probability is $1 - 0.96^{32}$ i.e. 0.73.

Nevertheless, one might think that the chance that *two* values exceed twice the mean value would be considerably lower. But, again using the 4% figure, we can calculate that the chance of seeing exactly one value exceeding twice the mean value is ${}^{32}C_1 \times 0.04 \times 0.96^{31} = 0.36$. Thus the probability of seeing a periodogram with at least 2 such values is $0.73 - 0.36$

$= 0.37$. One must conclude that the periodogram obtained in this case is not at all unusual and, furthermore, one prominent frequency could be a harmonic of the other. This makes it necessary to examine the phase of these two components.

Fig. 19.5 (bottom chart) shows that the component with frequency 1 is more out of phase than that with frequency 3. The respective values are, in fact, 0.98 and 0.39. These numbers can be used to calculate the absolute displacement of the positions of the maxima of the respective Cosine curves from their 'in phase' positions (the positions in which they coincide with the limits of the fundamental interval).

Any component frequency can have a phase with values from $-\pi$ to π . This range of 2π represents the wavelength of the component, which, of course, varies according to its frequency. For the component with frequency 1 the wavelength is 709.5m; thus a phase of 0.98 radians represents a spatial displacement, of the maxima of the Cosine curve from the 'in phase' position, of

$$\frac{0.98}{2\pi} \times 709.5 = 111m$$

This displacement is large and it is perhaps explained by the way in which the traces tend to be most dense on the left hand side of the distribution (see Fig. 19.5).

The wavelength of the component with frequency 3 is $709.5/3 = 236.5m$ and its phase is 0.39. Thus its spatial displacement is

$$\frac{0.39}{2\pi} \times 236.5 = 15m$$

This component is much more nearly in phase, which suggests that it is unlikely to be solely a harmonic of the component with frequency 1, because there is 96m between their maxima. Despite the lack of statistical significance of its amplitude, it may indicate a genuine original subdivision of the cadastral squares into three.

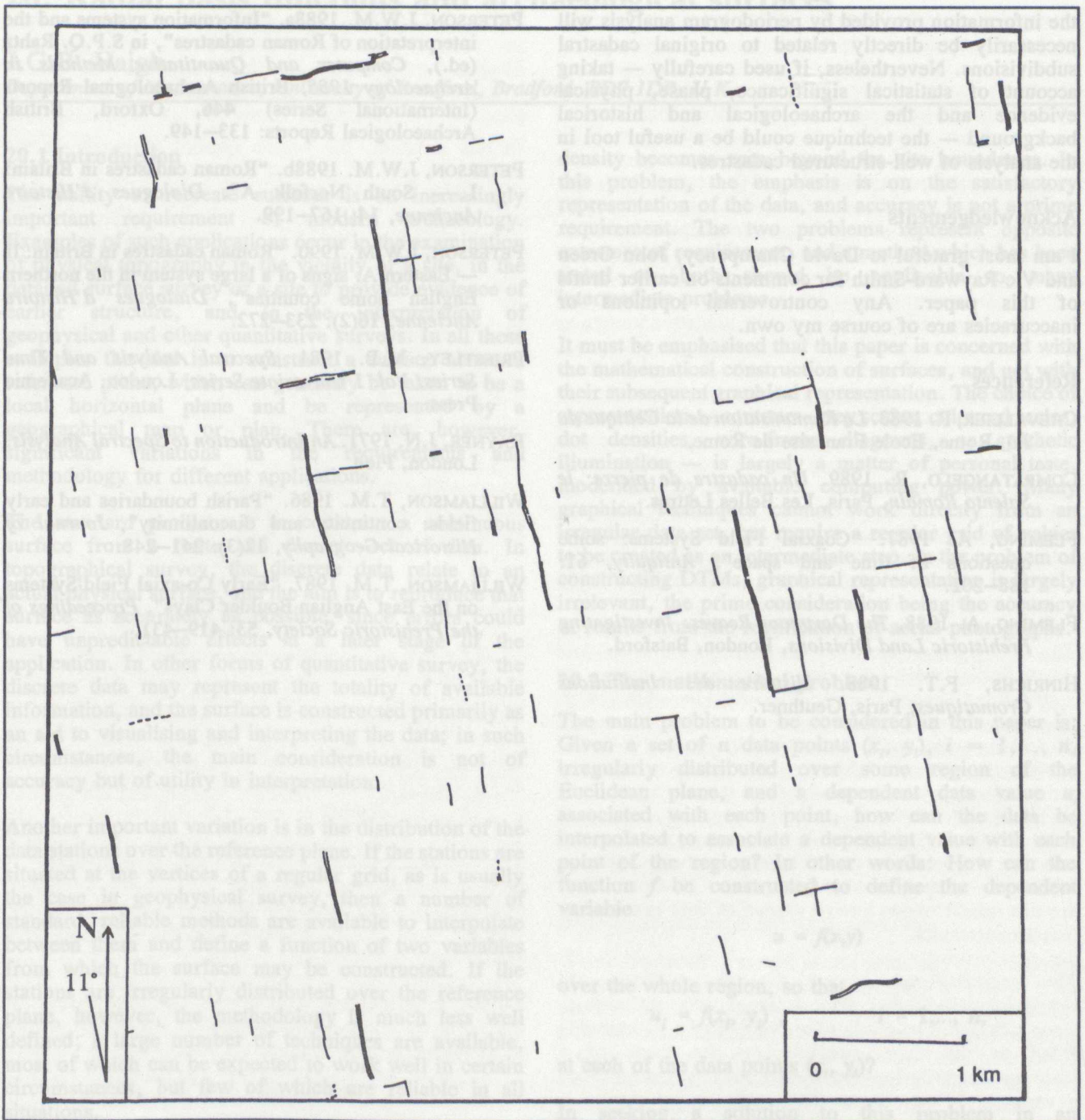


Figure 19.7: Traces from Fig. 19.4 which may represent a division of the cadastral grid squares into three.

Using this clue we can examine the supposed traces for physical evidence of this form of division (Fig. 19.7) — and there does seem to be clear evidence for it. It is also noticeable that these divisions are more uniformly distributed than the possible cadastral traces, considered as a whole. Possibly they reflect an original sub-structure of the cadastre, and the variable density of the other traces reflects different patterns of Roman and later land use.

This form of subdivision was not initially expected. The author's study of a possible Roman cadastre elsewhere in Britain (Peterson 1990) seemed to have revealed a division by four, which is not evident in this part of this cadastre. Nevertheless, two things argue for the idea that this pattern is the work of a Roman land surveyor. Firstly, since 3 is a factor of 2,400, one can

accept that, in a Roman context, squares with sides of 2,400ft can easily be divided in this way. On the other hand, although subdivision by 2 occurs (rarely) in prehistoric cadastres (Fleming 1988:65), other sorts of subdivision are not apparent, and not expected. Secondly the writings of the Roman agrimensores tell us that they did indeed divide land in this way. According to Hyginus, one square could be allocated to three men, and it has been suggested that in his time, around the end of the First Century AD (Hinrichs 1988:81), this area of land was considered to be the normal allocation for one man and his family (Chevallier 1983:38 note 55).

This case study suggests that, in this field, the technique of Fourier analysis needs to be grounded in the reality which the field boundaries represent. Not all

the information provided by periodogram analysis will necessarily be directly related to original cadastral subdivisions. Nevertheless, if used carefully — taking account of statistical significance, phase, physical evidence and the archaeological and historical background — the technique could be a useful tool in the analysis of well-structured cadastres.

Acknowledgements

I am most grateful to David Champeney, John Green and Vic Rayward-Smith for comments on earlier drafts of this paper. Any controversial opinions or inaccuracies are of course my own.

References

CHEVALLIER, R. 1983. *La Romanisation de la Celtique du Pô*, Rome, Ecole Francaise de Rome.

COMPATANGELO, R. 1989. *Un cadastre de pierre: le Salento Romain*, Paris, Les Belles Lettres.

FLEMING, A. 1987. "Coaxial Field Systems: some questions of time and space", *Antiquity*, 61: 188–202.

FLEMING, A. 1988. *The Dartmoor Reaves: Investigating Prehistoric Land Divisions*, London, Batsford.

HINRICHS, F.T. 1988. *Histoire des Institutions Gromatiques*, Paris, Geuthner.

PETERSON, J.W.M. 1988a. "Information systems and the interpretation of Roman cadastres", in S.P.Q. Rahtz (ed.), *Computer and Quantitative Methods in Archaeology 1988*, British Archaeological Reports (International Series) 446, Oxford, British Archaeological Reports: 133–149.

PETERSON, J.W.M. 1988b. "Roman cadastres in Britain: I — South Norfolk A", *Dialogues d'Histoire Ancienne*, 14: 167–199.

PETERSON, J.W.M. 1990. "Roman cadastres in Britain: II — Eastern A: signs of a large system in the northern English home counties", *Dialogues d'Histoire Ancienne*, 16(2): 233–272.

PRIESTLEY, M.B. 1981. *Spectral Analysis and Time Series: Vol. 1 Univariate Series*, London, Academic Press.

RAYNER, J.N. 1971. *An Introduction to Spectral Analysis*, London, Pion.

WILLIAMSON, T.M. 1986. "Parish boundaries and early fields: continuity and discontinuity", *Journal of Historical Geography*, 12(3): 241–248.

WILLIAMSON, T.M. 1987. "Early Co-axial Field Systems on the East Anglian Boulder Clays", *Proceedings of the Prehistoric Society*, 53: 419–431.