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

Review Lecture: The Recovery of Structure from Fragmentary Information

D. G. Kendall

Phil. Trans. R. Soc. Lond. A 1975 **279**, 547-582
doi: 10.1098/rsta.1975.0086

REVIEW LECTURE

THE RECOVERY OF STRUCTURE FROM FRAGMENTARY INFORMATION

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

(*Lecture delivered* 13 *February* 1975 – *MS. received* 6 *March* 1975)

[Plate 19]

CONTENTS

The topic to be considered is the recovery of a detailed structure of known general form from a multitude of minute fragments of information, each one of which is embedded in irrelevant material and may be corrupted by error. The simplest specific example is the problem of 'seriation on the basis of incidence of types', several automated solutions for which are now available and work well. These yield provisional solutions which have to be carefully cleaned up by the archaeologist, but they have the advantage (over earlier 'hand' methods) that the degree to which subjective judgements are employed can be carefully controlled. A more formidable problem is that of reconstructing the geographical layout of a mediaeval manor from information derived from contemporary cartularies and later deeds. This, too, can be solved 'by hand', but recently it has become clear that automated solutions are also possible, as in the seriation problem, and again offer certain advantages.

1. Introduction

The subject of this lecture is the recovery of structure of a known general form from a multitude of minute fragments of information embedded in a mass of irrelevant material and corrupted by error (see note 1, p. 576).

I am not aware of any previous attempt to treat this area of research in a unitary manner.

All my examples will be drawn from what one might call British Academy rather than Royal Society contexts, and I shall mention only one (Benzer 1959, 1961) of the scientific applications (the mapping of genes and viruses), where however Dr Francis Crick tells me that the situation is so nearly 'error free' as to make the present type of method unnecessary: 'direct action' suffices. Of course, one welcomes such a situation for the opportunity it provides for testing the method in exceptional circumstances, and this check has in fact been carried out.

After listing the applications, actual or envisaged, it became clear to me that a lecture attempting to cover all of them would be unintelligible to a large fraction of the audience. I will therefore concentrate on one particular application, on which much of the work is quite recent, and I must refer you to the accompanying *Notes* for brief accounts of and references to cognate work on other problems.

2. Map-making in one dimension: seriation

I like to think of these tasks under the heading: the building up of maps out of scraps; in other words, doing jigsaw puzzles when some of the pieces are missing, when most of them are broken, and when the picture of *The Stag at bay* has been torn off.

Map-making in its simplest form is one dimensional, and here we have the problem of seriation; how to arrange a series of 'objects' in approximately the correct serial order, when we are provided with a mass of rather unreliable information about serial (e.g. chronological) 'proximity'. This, if we are dealing with 'objects' which are archaeological closed finds ('graves'), will often consist of information about the 'presence or absence' in such a 'grave' of various '*types*' of pottery, or jewellery. If the 'objects' are undated charters, then the information might be the presence or absence of a named witness. If the 'objects' are the 'point-mutants' of a virus (which one can think of as identified with locations on the virus), then typically the information will relate to experiments in which, by recombination in the presence of certain 'deletion-mutants', the 'wild type' has been recovered. And so on.

Whatever form the information about a single 'object' takes, we seek to use it to build up an appropriate index of proximity, or rather, of *serial* proximity (see note 2), for *pairs* of such 'objects'. We then have the task of fitting together the 'objects' so that pairs with high 'proximity' lie closer together in the serial order than those with low 'proximity'.

If we have a total of N 'objects' then there will $\frac{1}{2}N(N-1)$ pairs. The information about 'objects' and 'types' will normally be stored in a 'matrix', or 'table of double entry', where the *rows* represent 'objects' and the *columns* represent 'types'. The seriation problem is then one of rearranging rows so as to give the matrix what is considered to be a meaningful form.

In figure 1*a* we see such a 'presence-or-absence' matrix, often called an incidence matrix (see note 3), for 5 'graves' B, A, E, C, and D in relation to 4 types of decorated pottery. (The example is, of course, an artificial one.)

When the *rows* are rearranged as in figure 1*b*, you will see that in each column, the 1's are bunched together in a compact block (see note 4). When an incidence matrix has this form I have called it a Petrie matrix, after Sir W. M. Flinders Petrie, who first tackled such problems on a large scale about 75 years ago (1899, 1901). In this example there is essentially only one way of rearranging the rows so as to get the Petrie form (of course, it is preserved

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

by a complete reversal of the order). Thus in this artificial example we should decide, on the basis of the evidence presented, that ABCDE or EDCBA is the correct chronological order. The decision between these two options would have to be taken on other grounds; perhaps, say, from carbon-dating evidence.

(*a*)

B	1	1	0	0
A	1	0	0	0
E	0	0	0	1
C	0	1	1	0
D	0	0	1	1

(*b*)

A	1	0	0	0
B	1	1	0	0
C	0	1	1	0
D	0	0	1	1
E	0	0	0	1

FIGURE 1. (*a*) An artificial incidence matrix; 5 graves and 4 types. (*b*) The same matrix, with rows (graves) rearranged to recover the Petrie form (1's packed together in each column).

In a simple case like this the problem is quite trivial and can even be solved by mental arithmetic. But it is a different matter if (as in Petrie's famous example) we have an incidence matrix with 900 rows, and 800 columns. Today we should think this an appropriate task for a computer – and not at all an easy one, even then. Petrie himself, however, thought nothing of doing it by hand; an altogether staggering feat (see note 5).

In a moment I am going to tell you a little about what are called 'scaling' methods. These are designed to detect or (what is appropriate for us) to *recover* structure in a set of 'objects' when the latter is endowed with a measure of similarity (or, it may be, dissimilarity) for pairs of 'objects'. What we mean by 'similar' or 'dissimilar' will depend on the context. If we are seeking a serial chronology, then we use a similarity measure which *we hope and believe* is positively related to chronological proximity. This is sometimes set up arbitrarily and subjectively, but I have proved (1969, 1971 *a*) that in a perfectly posed 'Petrie problem' there is a natural way of setting up a similarity measure, here called K_1, for pairs of 'graves' which is such that *if* the incidence matrix can be thrown into Petrie form, then the similarity measure K_1 contains all the information needed to discover the acceptable rearrangement, or rearrangements.

Actually one can extend this result, both to deal with 'abundance' rather than mere 'incidence' matrices, and also to obtain a series of different similarity measures $K_1, K_2, K_3, \ldots$, any one of which suffices for the discovery of the 'petrifying permutation'. This later result (see note 6) was established in 1971, and has proved to be an acceptable basis for seriating large collections of 'graves', when using one of the 'scaling' algorithms. Of course, by an 'abundance matrix' we mean a table which tells us, not only whether some 'type' of pottery or jewellery is present, but also, if so, in what abundance, and the Petrie property for an appropriately rearranged matrix is generalized in a natural way so that it is meaningful also for abundance matrices. I will not, however, complicate matters by going into the mathematical details here.

What must be said is that the seriation resulting from the composite algorithm just mentioned, known as HORSHU, will only be chronologically significant if the underlying typology is 'chronologically oriented'. If it is not, then while one may still recover a seriation, it may be a spatial or a sociological rather than a chronological one. Whether this is so or not, in a practical case, is a matter for archaeological judgement. Computer solutions can remove years of drudgery, can supply an otherwise almost unattainable degree of objectivity, and can also help by making it apparent (through repeated runs from different *random* starting situations) that several possible solutions exist, or that no acceptable solutions exist. But it cannot remove the elements of subjective or professional judgement which are needed at the *outset*, in deciding whether or not to accept a typology, and at the *end*, in deciding whether or not to accept a solution, and how then to modify its finer detail in accordance with other evidence and one's general background knowledge of the particular problem.

	A	B	C	D	E
A	1	1	0	0	0
B	1	2	1	0	0
C	0	1	2	1	0
D	0	0	1	2	1
E	0	0	0	1	1

K_1

A	B	C	D	E
2	2	1	0	0
2	4	2	1	0
1	2	4	2	1
0	1	2	4	2
0	0	1	2	2

K_2

A	B	C	D	E
5	5	4	2	1
5	9	6	4	2
4	6	10	6	4
2	4	6	9	5
1	2	4	5	5

K_3

FIGURE 2. The similarity matrices K_1, K_2 and K_3 corresponding to the incidence matrix in figure 1. The row/column order is the correctly seriated order, to enable the structure to be more clearly seen. Notice that only with K_3 is there a unique most dissimilar pair.

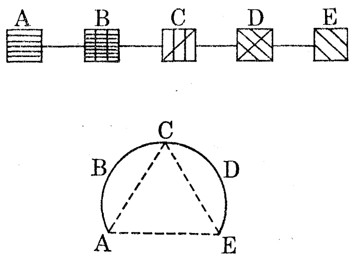


FIGURE 3. Diagrammatic representation of the incidence matrix in figure 1, and a sketch to show the origin of the 'horseshoe' effect.

Here in figure 2 are the first three of the similarity measures for our artificial example, the 'graves' being given their 'correct' order. Notice that as we proceed from left to right, we get better discrimination; in the first similarity matrix no less than 6 'pairs' all score the minimal value of zero, and unless one is careful this can lead to a curious phenomenon. Figure 3 illustrates this. At the top we see the 5 'graves' arranged in correct order and the shading of the squares which here represent them helps us to remember what is being assumed about their contents. From this we see that, superficially at least, there is nothing at all in common between the pairs A and C, C and E, and A and E. This is also true of the description of the situation summarized in the first similarity measure, and so if one uses the scaling algorithm invented by Roger Shepard and perfected by Joseph Kruskal in its strictest form (see note 7), in which all 'pairs of pairs' are compared, and 'tied' pairs are treated as 'ties', then one gets a seriation of characteristic 'horseshoe' form. The essence of this scaling algorithm is that *'objects' are to be represented by points, in this case in 2 dimensions, and the more similar pairs of 'objects' are to be represented by the closer pairs of points.* Thus it comes about that the computer tries to make ACE into an equilateral triangle, and that in essence is the reason for the horseshoe effect (and the name of the algorithm).

The higher-order similarity-measures compare 'graves' via other 'graves', and in this way achieve better discrimination; this helps to break down the horseshoe effect and give a more nearly linear order. A different treatment of 'tied' similarities is also useful in this respect and the two devices can be employed together to 'unbend the horseshoe'; figure 4 (with 'tied ties') shows how the horseshoe unbends, for our artificial example, when one replaces K_1 by K_2 and then by K_3.

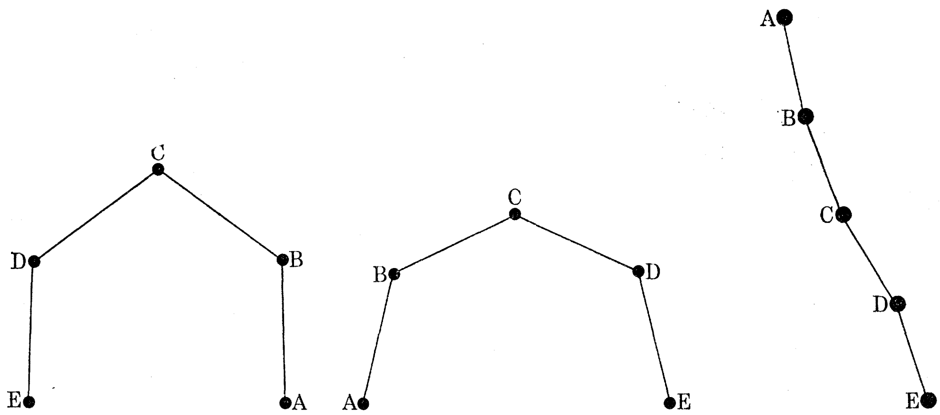


FIGURE 4. Outputs from HORSHU for the artificial example in figure 1. The similarity measure is: first K_1 (left), then K_2 (centre) and then K_3 (right); K_3 'unbends the horseshoe'. (The 'secondary' treatment for ties has been used here.)

Notice, however, that the 'horseshoe-shaped' picture has an advantage of its own, which has not been remarked upon before; if we draw onto the picture line-segments joining all pairs of strongly similar 'graves', then we can judge for ourselves whether or not these segments 'respect' the seriation, as they should do, or merely run criss-cross over the figure. In the latter case we would have no hesitation in rejecting the seriation as an artefact. Once the 'horseshoe' is unbent, we can no longer so easily carry out this test, because the picture is almost one dimensional, and the segments become confused. For this reason I believe that the horseshoe, *with such strong links*, should *always* be drawn, even though the seriation itself is taken from an 'unbent' horseshoe obtained by the 'unbending' procedures. It is, of course, for this reason that we carry out the seriation algorithm in two dimensions and not in one dimension. *We want to give the algorithm every possible chance to fail.* This is one of the great advantages of the HORSHU technique over some of its rivals; that when it ought to fail, it will normally fail dramatically.

I should mention here that the three scaling pictures just shown, and one of those to follow later, were drawn through the kindness of Dr Robin Sibson of Cambridge, who has not only prepared an extremely flexible and powerful version of the scaling algorithm, but has also added many new devices to it, which are of great value in seriation and in other problems. It would not be possible, however, to go into the details of these here. I will only say that he

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

makes it possible to consider not *all* but in a sense *the most relevant* 'pairs of pairs of objects', and that he too has a device, different from mine, for modifying a similarity measure so as to become more discriminating (Sibson 1972; see also his article in Hodson, Kendall & Tautu 1971). For related development see also Wilkinson (1974).

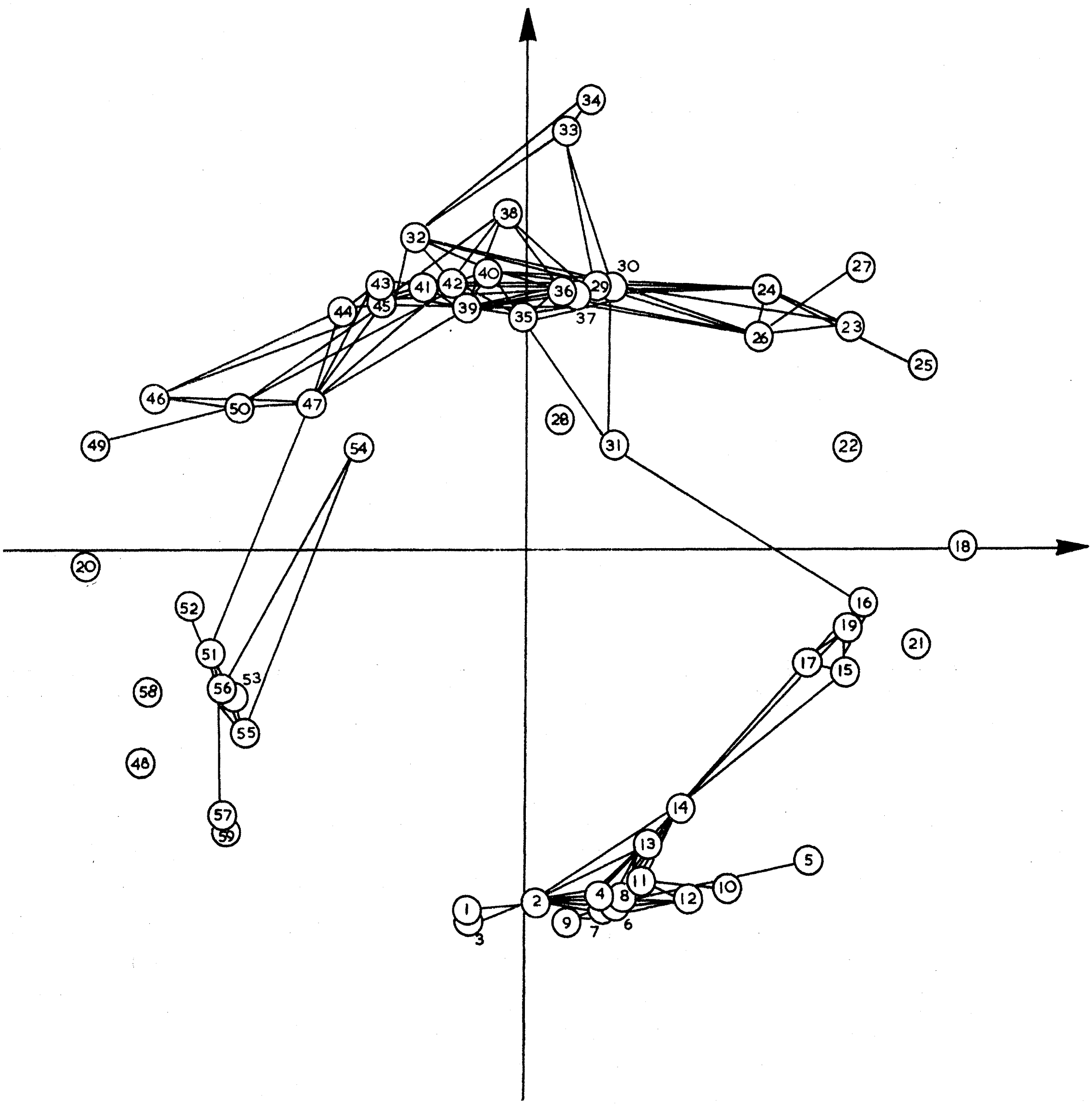


FIGURE 5. Münsingen–Rain. HORSHU output by using K_1. The numbers indicate Hodson's own ordering of the graves. (Only *strong* links are shown.)

I conclude this very brief review of seriation by showing (figures 5–9, from Kendall 1971*d*) my algorithm HORSHU in action on a real problem: Professor F. R. Hodson's collection (1968) of 59 La Tène graves at Münsingen–Rain, containing 70 types of fibula. Here you see a typical 'horseshoe', an 'unbent' horseshoe, a linear projection of the latter, the computer rearranged incidence matrix, and finally a plot of the computed order C_1 (horizontal axis) against Hodson's agreed archaeological order H (vertical axis); 85% of the graves are not more than 5 steps out of place. Further details will be found in Kendall (1970, 1971*d*).

The picture of the rearranged incidence matrix, almost but not quite of Petrie form, provides an opportunity to point out that while the HORSHU algorithm is based on a theory which strictly applies only when the exact Petrie form can be realized, nevertheless it has the necessary robustness to give good results in the typical practical situation when, for various obvious reasons, only an approximately Petrie form is to be expected.

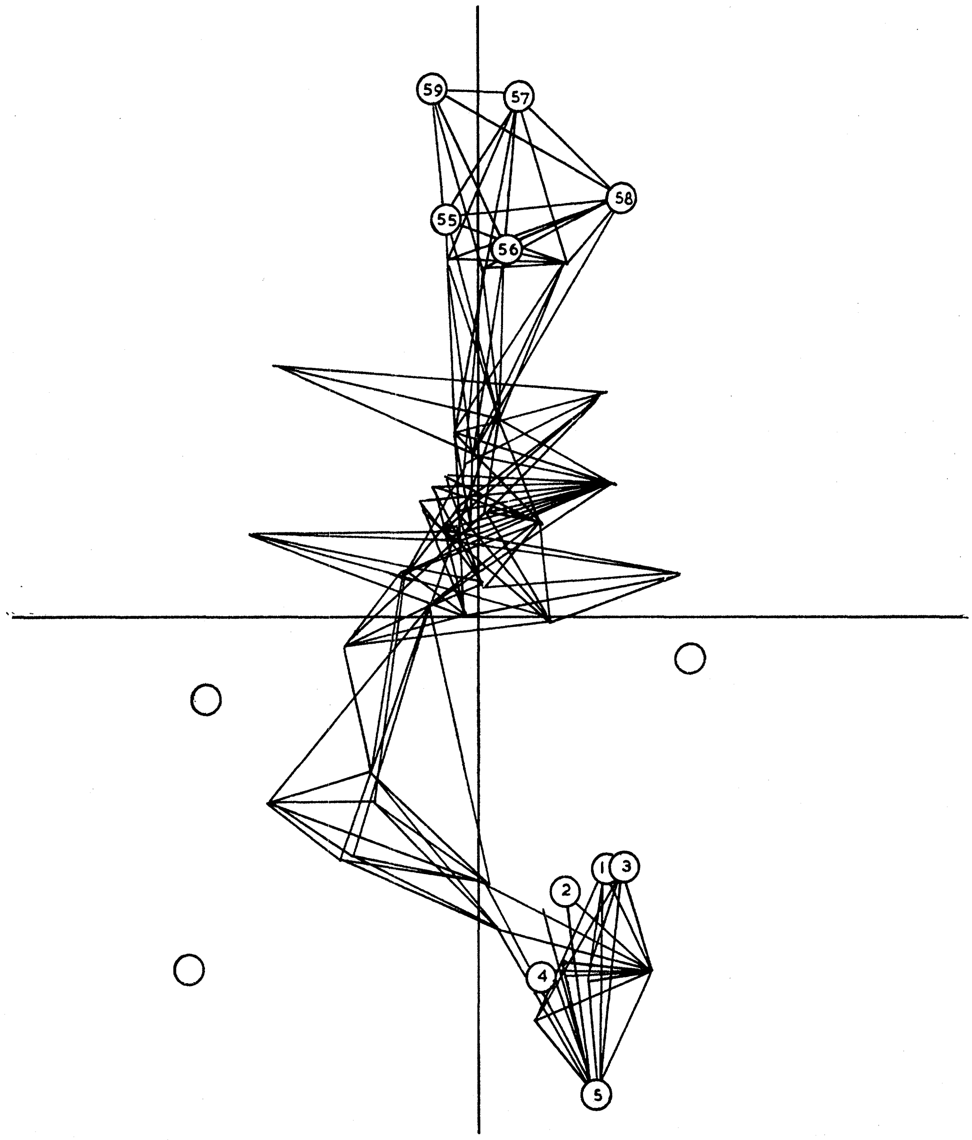


FIGURE 6. Münsingen–Rain. HORSHU output by using K_2 and the 'primary treatment of ties'; the horseshoe has been unbent. The numbers indicate Hodson's own ordering of the graves.

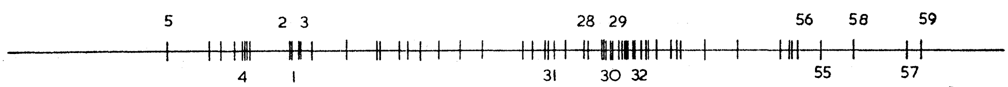


FIGURE 7. Münsingen–Rain. The linear projection of the 'ordering' in figure 6 onto the first 'principal component' (followed by a one dimensional MD-SCAL run from that starting configuration). The numbers indicate Hodson's own ordering of the graves.

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

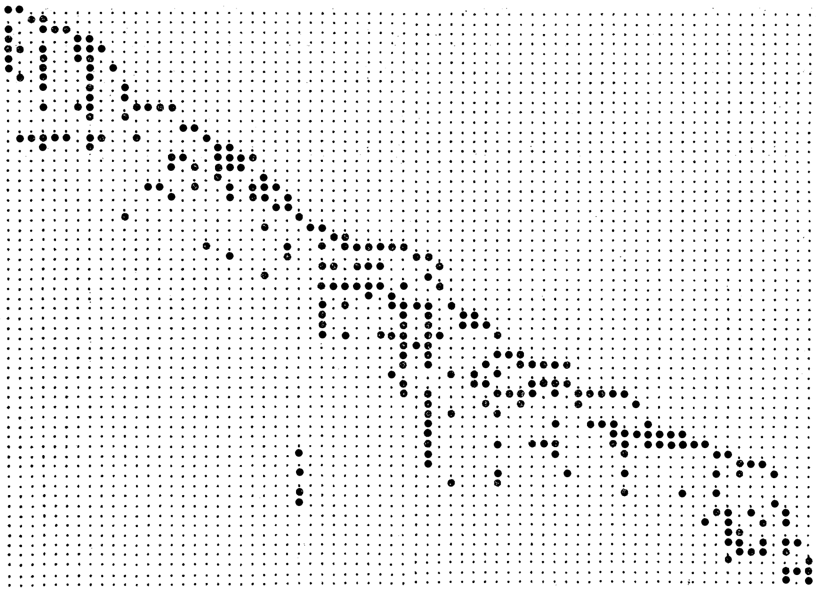


FIGURE 8. Münsingen–Rain. The incidence matrix (•, 1; ., 0) with the rows (= graves) rearranged according to the *computer*-ordering. An example of 'semi-Petrie form'. (The columns have also been rearranged to clarify the structure.)

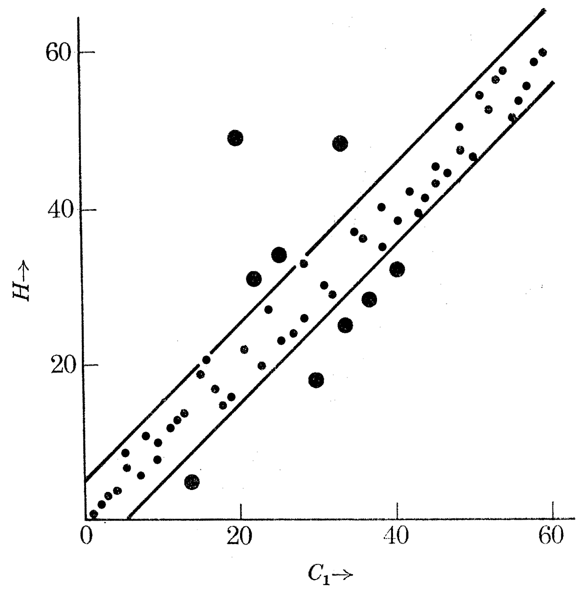


FIGURE 9. Münsingen–Rain. Plot of the HORSHU computer-ordering (horizontal axis, labelled as C_1) against Hodson's ordering (vertical axis, labelled as H). Eighty-five per cent of the graves are out of place by not more than 5 units.

3. Map-making in two dimensions

It is a little easier to grasp the working of the scaling algorithm in the genuinely two dimensional case, and to this we now turn. We shall be using it to make maps, and they will be real geographers' maps. It is true that the scaling algorithm can be used, and often is used, as a means of 'seeing pictures in the fire' (see note 1); this is particularly true of some

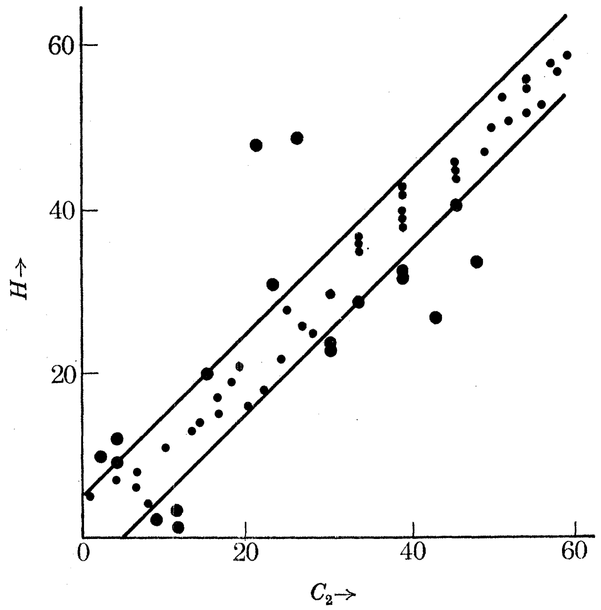


FIGURE 10. Münsingen–Rain. FLATTMAP analysis. This is to be compared with figure 9; the horizontal axis now shows the ordering C_2 derived from the pooling of 3 independent and largely concordant FLATTMAP runs. The vertical axis corresponds as before to Hodson's ordering.

of its applications in psychology for which it was first designed. But I must remind you that we are concerned today with the very different business of trying to recover structure of known general form. Thus the computer has to generate the kind of picture we are looking for, or it has failed; a totally different sort of picture will imply that there is something seriously wrong with our assumptions.

The 'similarity' (now spatial 'proximity') evidence used for map-making can be of three kinds, and we shall be looking at illustrations of each one of these three cases.

I. The 'pairs of objects' are arranged in order of decreasing (or increasing) 'similarity'.

II. The 'pairs' are separated into 2 groups; those pairs whose members are 'similar', and those pairs whose members are not 'similar'.

III. The 'pairs' are separated into 3 groups; those 'pairs' whose members are 'similar'; those 'pairs' whose members are not 'similar'; and those 'pairs' whose members may or may not be 'similar', information here being lacking.

Let us begin by taking a trivial example of the first situation. Figure 12 shows the A.A. recommended-route distances between six cities: Cardiff, Derby, Gloucester, Hull, London and York. Figure 13 shows merely a list of pairs of cities arranged so that the closest pair comes at the top, then the next closest, and so on. Observe that a lot of information has now been thrown away. We can, however, use this as a list of pairs ranked according to 'similarity', and the scaling algorithm (here again I used Sibson's version) gives us the map shown in figure 14. It is not very accurate (such programs are not really intended to do a good job on so few as 5 or 6 'objects'), but it still presents us with a reasonably satisfactory view of this part of the United Kingdom (see note 8). As required, Hull and York form the closest pair of cities on the map. We are not to complain if they are shown as being too close; the information to decide such a point was not given to the computer.

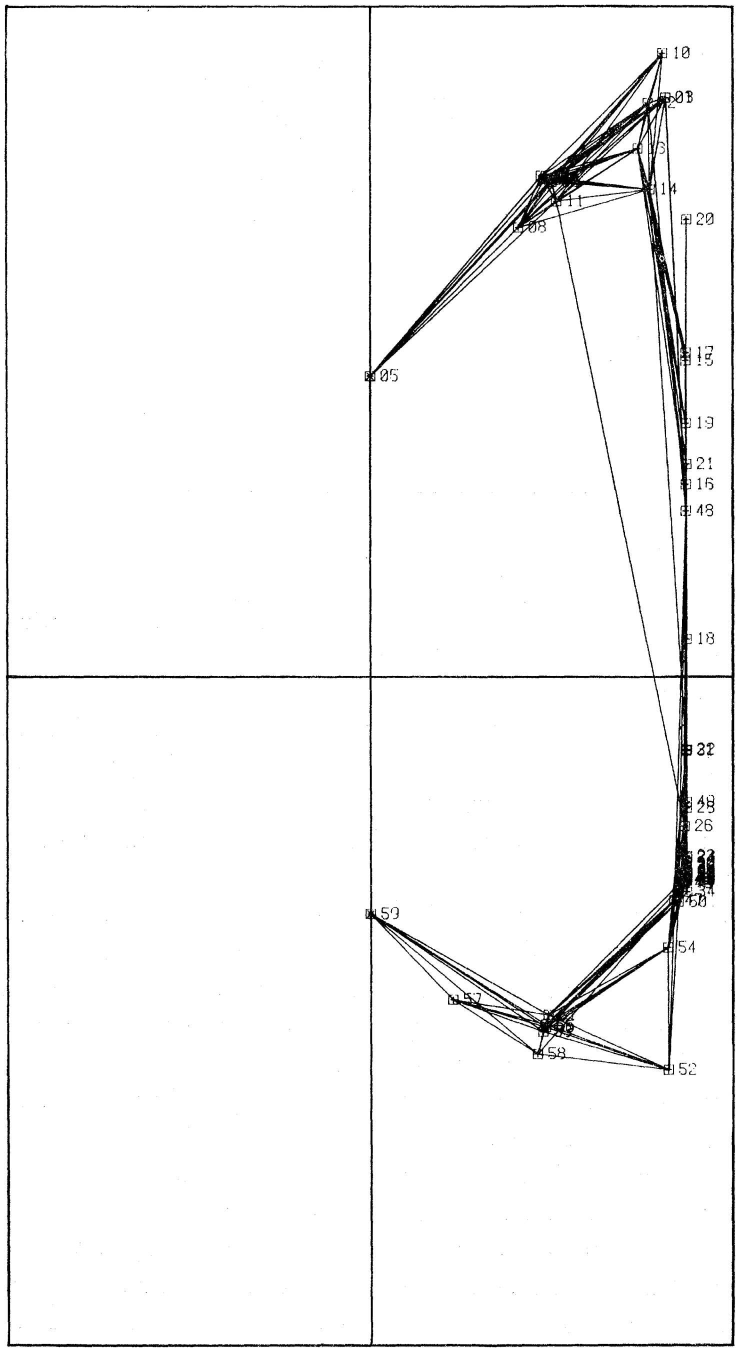


FIGURE 11. Münsingen–Rain. The FLATTMAP seriation has been artificially blown out into a horseshoe form (by fixing the two extremities found for the seriation and then introducing a set of forces designed to extend the configuration to the right). Differential extension, with the greatest displacement in the centre of the seriation, should improve this figure and facilitate the inspection of the 'strong links', here insufficiently separated.

	CRD	DRB	GLC	HLL	LND	YRK
CRD	.	142	56	224	153	229
DRB		.	93	88	123	88
GLC			.	168	103	180
HLL				.	166	38
LND					.	194
YRK						.

FIGURE 12. The road distances (in miles) between six cities (Cardiff, Derby, Gloucester, Hull, London and York).

Ranking of pairs of towns			
1st – 3rd	HLL/YRK	CRD/GLC	DRB/HLL
4th – 6th	DRB/YRK	DRB/GLC	GLC/LND
7th – 9th	DRB/LND	CRD/DRB	CRD/LND
10th–12th	HLL/LND	GLC/HLL	GLC/YRK
13th–15th	LND/YRK	CRD/HLL	CRD/YRK

FIGURE 13. The 15 *pairs* of cities in figure 12 are here arranged in order of increasing city–city distance.

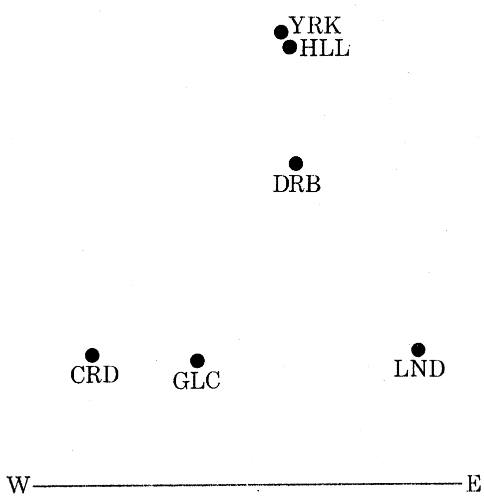


FIGURE 14. A map of the six cities, obtained from the data given in figure 13, by using the MD-SCAL program.

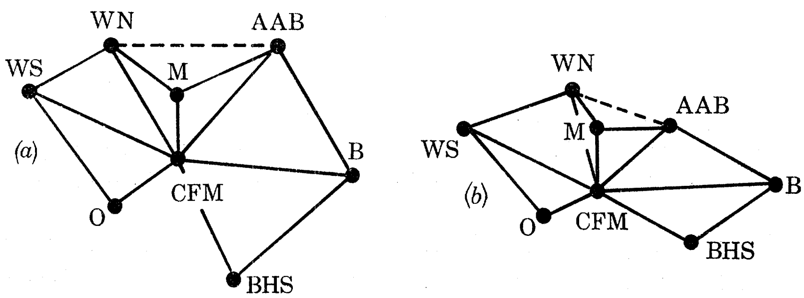


FIGURE 15. The eight Otmoor parishes (BHS: Beckley, Horton and Studley, etc.), (*a*) as they actually lie on the ground; (*b*) as reconstructed by MD-SCAL, using only the rank-order of the standardized intermarriage-rates for the period 1600–1850, approx. The spots represent the population centroids for the actual parishes, in (*a*), and the MD-SCAL-computed 'parish positions', in (*b*). Line segments indicate that the two terminal parishes have some common boundary (this information was *not* employed in the computation, and is thus available as a check).

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

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

Another rather different example of this kind of mapping-situation is illustrated by the work that I did some years ago on intermarriage rates between Otmoor parishes (1971 *b*).

Otmoor, which lies to the northeast of Oxford, was originally a very isolated and marshy group of parishes, and indeed still retains a somewhat independent character. The diagram on the left of figure 15 shows a 'contiguity graph' for the parishes, with each of these represented by a point located at the centroid of its component townships, and these points are linked by a line-segment wherever 2 parishes share a common boundary. (In one case the common stretch of boundary reduces to a point, and then the line-segment is shown 'dotted.')

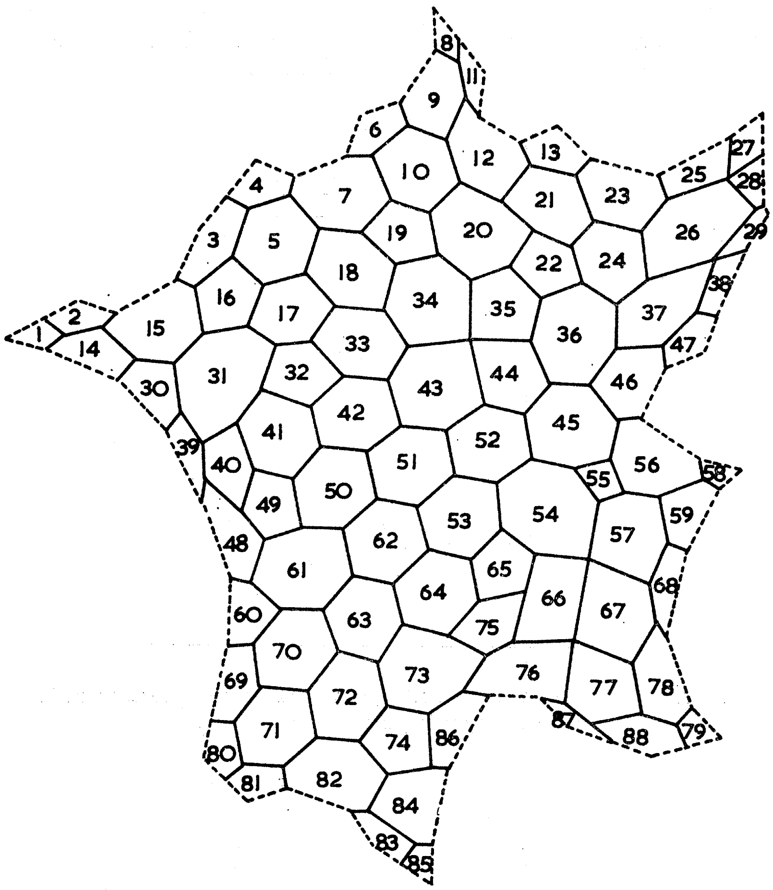


FIGURE 16. A computer map of the 88 Departments of France (Corsica and Seine omitted). This map has been constructed by MD-SCAL. Here the data are the ranked Wilkinson distances between paired Departments, derived merely from a knowledge of which pairs of Departments abut, and which pairs do not.

An interdisciplinary group (Hiorns, Harrison, Boyce & Küchemann 1969) working in Oxford has collected the numbers of marriages between parties living in each pair of (different) Otmoor parishes during the approximate period 1600–1850, and from these I calculated a standardized intermarriage rate, and used it as an index of 'proximity-similarity', thus obtaining from the scaling algorithm the surprisingly good map shown on the right of the same figure. In fact very few marriages were involved in the information used. Most Otmoor marriages were entirely within one parish, or involved one party from 'the outside world'. Nevertheless this mere trace of information about proximity has sufficed to give us a map (see note 9).

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

These last two examples were both of type I; we turn now to an example of type II.

Here (figure 16) is my computer map (1971*c*) of the 88 Departments of France, which illustrates what one can do when one merely knows that some pairs of Departments are 'similar' (share a common boundary) and all other pairs are 'dissimilar' (do not share a common boundary). Detailed comparison of the computer map with a true map (figure 17) shows the really astonishing accuracy of the former. Here the 'abuttal' information (see note 10) was converted into 'similarity' information by ranking the dissimilarity between two Departments *A* and *B* according to the size of the minimum number of intermediate Departments which have to be crossed in order to travel from *A* to *B*. This number, increased by unity, is the Wilkinson 'distance' from *A* to *B*.

Among geographers, W. R. Tobler of the University of Michigan has been a pioneer in the use of scaling methods for map-making. See, for example, his paper with Wineberg (1971), and his earlier work (e.g. with Mielke & Detwyler 1970).

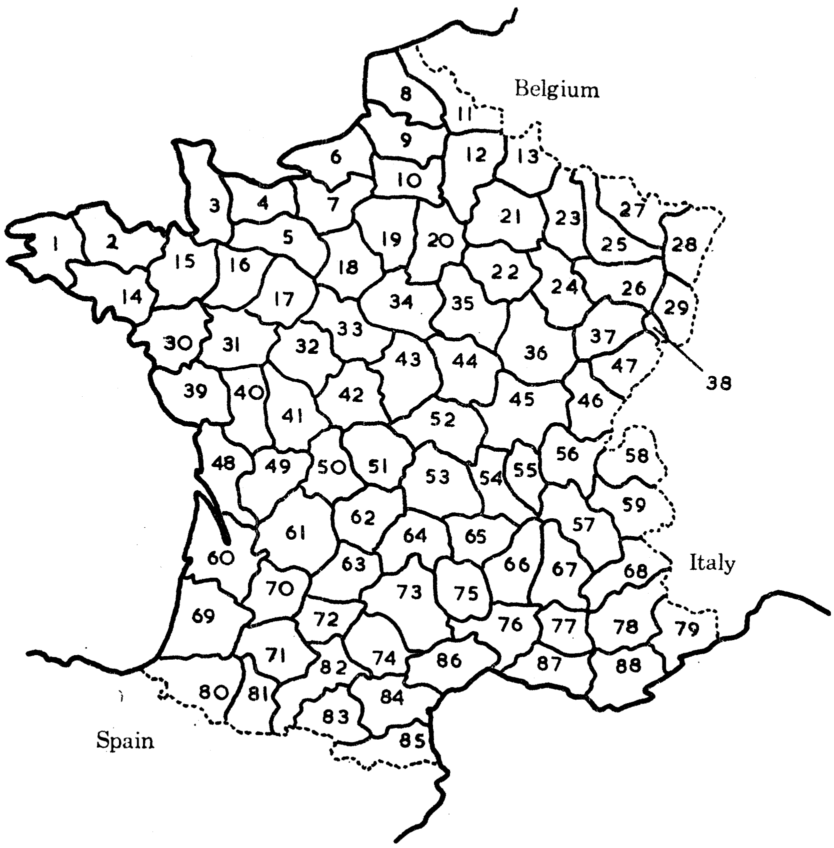


FIGURE 17. A true map of the 88 Departments of France; the system of numbering the Departments is the same as that used in figure 16, so that each integer labels the same Department in the two figures.

4. Manorial reconstitution: the Whixley problem

We have a decidedly worse situation in problems of type III, where the pairs of 'objects' are separated into three groups; those which are 'similar' (do abut), those which are 'dissimilar' (do not abut), and those for which information is lacking. It is this third problem which interests me most at the moment, and I wish to discuss it in some detail. Let me first indicate how it arises.

A slide† which may not be very clearly visible to everyone shows a portion of the Manor of Laxton in the seventeenth century; that is, in pre-enclosure days. You will observe that the arable land occurs in 'field'-shaped compact parcels which I shall call *flatts* (the term used varies from one part of the country to another). These flatts had *names*, and are now the primary 'objects' of study, so in due course we shall be interested in the similarity (in a proximity sense) between two flatts.

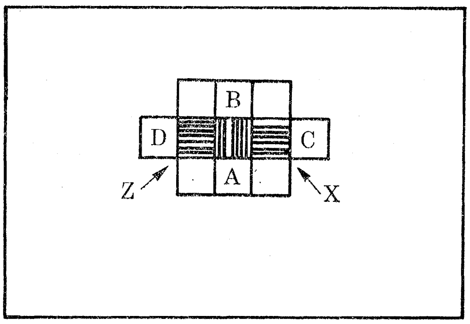


FIGURE 18. A fictitious example showing 11 'flatts' and the 'selions' in 3 of these. The central flatt is named Y; the text explains how a surveyor would describe the abuttals involving flatt Y.

Each flatt in itself consists of a bundle of more or less parallel strips, called *selions*. Normally each strip was cultivated by a particular person either as owner or tenant, so a strip has (1) an owner, (2) an area, (3) an orientation, and (4) a named flatt in which it lies.

In a written survey of such a manor, entries occur which I will illustrate with reference to the imaginary example in figure 18. When describing the flatt Y, in the centre of the diagram, the surveyor will say that its strips run from the south (where they abut on A) towards the north (where they abut on B). So we know that Y has A and B as neighbours. We may learn no more than this about Y, but *if we are lucky* and if the selions run as shown in the diagram, then from the description of X we shall learn that it too is a neighbour of Y, and similarity for flatt Z, so eventually we shall have all of X, A, Z and B as neighbours of Y. If 'neighbourliness' is here equated with 'similarity', then in the first case, in which only A and B are 'declared' neighbours of Y, we shall not know the status of the pair (X, Y); these might be adjacent flatts, or they might be quite remote from one another. This last situation, the occurrence of (many) pairs of flatts concerning whose proximity we have no direct knowledge, is *the norm* in the manorial reconstitution problem, and this is responsible for the assignment of that problem to what I have called type III. However, while this makes for greater difficulty, the number of known abuttals per flatt being reduced to an average of 3 (and on occasion to merely 2), instead of the 5 or 6 typical of the 'map of France' problem, we do have certain other new features which go some of the way towards offsetting this extra difficulty.

In the first place there will normally be some flatt-names which have survived to the present day, either as field-names, or as names on one of the early Ordnance Survey maps, or in maps associated with Enclosure or with Tithe Redemption. These, if the identification

† Not reproduced. For this Laxton map see Orwin & Orwin (1938). See also Chibnall (1965) for similar maps and for a thorough account of the manorial reconstruction problem.

is one in which we feel sufficient confidence, can be treated as 'fixed points' of known position.

In the second place not all the abuttals given in the manuscript survey will be with other flatts; some will be with lines of communication (roads, bridle-paths, foot-paths, etc.) and with other linear topographical items such as streams. Here again we may be able to identify such items on the modern map, and so treat them, too, as 'fixed'. There is the awkwardness that we will not normally know on which stretch of (nor from what side of) the known stream the 'variable' flatt abuts, and a satisfactory procedure must take account of this. (It is of course assumed that no large scale pre-enclosure map of the manor survives; if one is available, there is virtually no problem.)

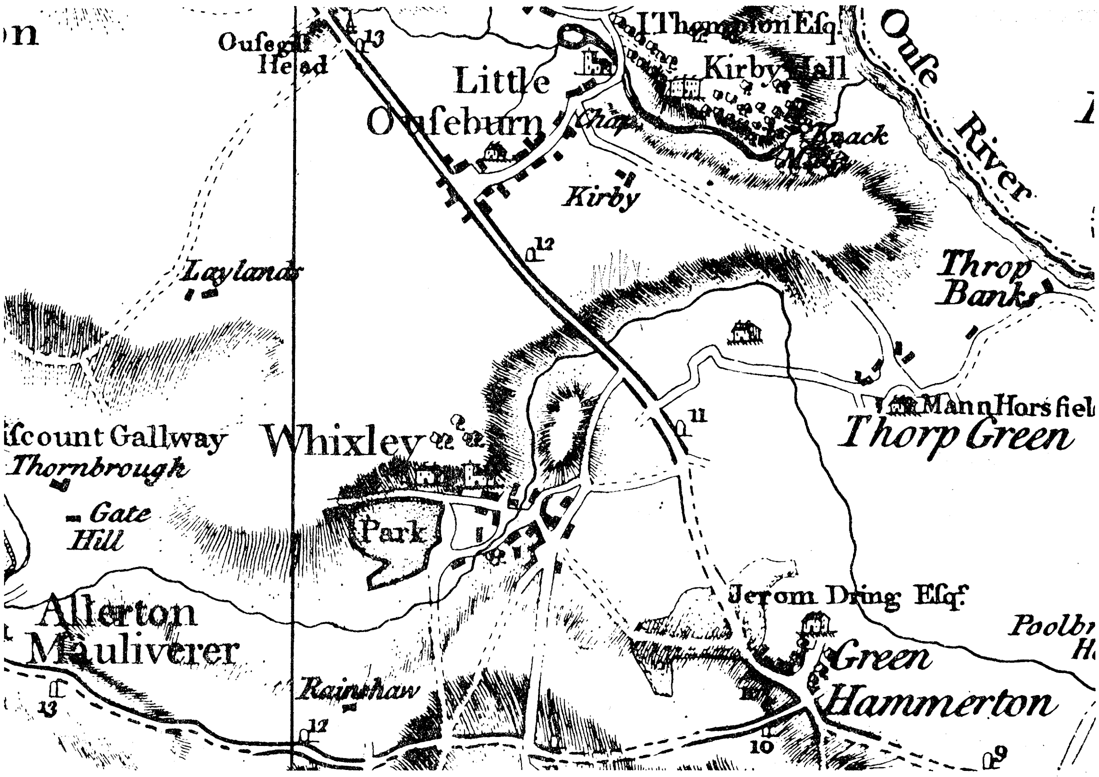
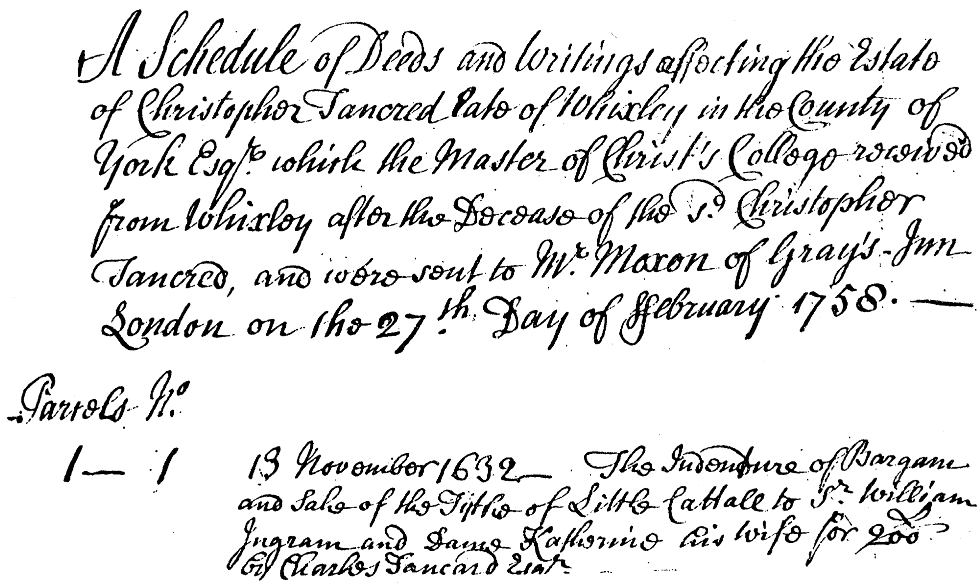


Figure 19. The Manor of Whixley as portrayed by Thomas Jefferys in 1771. Notice the 'quadrilateral' lying immediately south of the village and its crofts, bounded on the east by the roman road *Watlyngstrete*, on the west by the ancient road *Wederbygate*, and on the south by the ancient 'road to St Robert's' (this last is Knaresborough). The aerial photograph in figure 29 covers the top of this quadrilateral area, which is that dealt with in the initial computer experiments using FLATTMAP.

Because of my family connexions with the Manor of Whixley (my Kendall and Styan forebears lived and worked in the parish for about two centuries) it now seems almost inevitable that I should eventually have become involved with the reconstitution of that manor in its mediaeval state, but it was in fact only a chance remark by Mr Paul A. Harris that brought me to look at Whixley in this light (see also Göransson 1961).

Those who are not already familiar with the geography of this area will find it useful to inspect figure 19, which is an enlargement of a portion of the map (about one-inch to the mile on its original scale) engraved by Thomas Jefferys in 1771, thirty years before the enclosure of Whixley by the Act of 41 Geo III (= 1800–1801). (A fine reproduction of the Jefferys map has recently been published by Mr H. Margary.) The arable fields are not shown, and a large-scale pre-enclosure map of Whixley is one of those things which, in our case, we have not got (Reed 1946).†

The Court Rolls (apart from a few leaves) have also vanished, and at first one had very few deeds relating to the manor, although a document found in the Muniments of Christ's College, Cambridge, by Mr Donald Missen (see figure 20) tells us that a rich hoard of title deeds (including Fines from 1 Edw III and Court Rolls from 13 Chas I) passed through the hands of the Master of Christ's on 27 February 1758 on their way to 'Mr Moxon of Gray's Inn' in connexion with the Chancery suit precipitated by the will of Christopher Tancred, who left the Manor to found a charity (still in being) and so excluded his numerous sisters from what they had supposed would be their inheritance.


A Schedule of Deeds and writings affecting the Estate of Christopher Tancred late of Whixley in the County of York Esqr which the Master of Christ's College received from Whixley after the Decease of the sd Christopher Tancred, and were sent to Mr Moxon of Grays-Inn London on the 27th Day of February 1758. —

Parcels No

1 — 1 13 November 1632 — The Indenture of Bargain and Sale of the Tithe of Little Cattall to Sr William Ingram and Dame Katherine his wife for 200 £ Of Charles Tancard Esqr

FIGURE 20. The opening paragraph of the Christ's College manuscript listing the muniments of title of the Manor of Whixley in 1758.

Until very recently it appeared that these documents had vanished beyond recovery. The Governors and Trustees of Tancred's Foundation, through their Clerk Mr S. J. Mosley, have kindly deposited all their Whixley material with the Yorkshire Archaeological Society on permanent loan; this large archive contains an enormous amount of interesting material, but the 1758 'hoard' is not included in it. About 1920 the Foundation sold almost the whole of Whixley to the then West Riding County Council of Yorkshire, and efforts were made to trace the 'hoard' at Wakefield, without success. A few weeks ago the fact of the recent revision of administrative boundaries caused a large collection of muniments of title to be sent from Wakefield to the North Yorkshire County Record Office, and the North Yorkshire Archivist, Mr Michael Ashcroft, now tells me that it contained a large number of sixteenth- and seventeenth-century Whixley deeds, many very decayed. There has not yet been an opportunity for me to examine these, but it seems very probable that they form a remnant of the 'hoard',

† At the time of going to press I am happy to report the discovery of an important working draft of the Enclosure map containing pre-enclosure information; it was identified by Mr Keith Brandwood of Leeds University.

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

the other portions of which may have perished. We may hope that these deeds will in due course supply a number of further 'fixed points', and will also perhaps add to the number of abuttals. I mention these facts in order to bring out the dynamic rather than static character of the present investigation; the information is not constant, but grows while one is working on it (see note 11).

The main source of information about Whixley is however of much earlier date; this is the famous Whixley Cartulary owned by Major Geoffrey Dent, M.C., of Ribston Hall, and carefully preserved by him and made available to scholars. It is a large volume and commences with a great number of copies of fourteenth-century deeds defining the descent of the manor from the Mauleverers through the Quixleys (see note 12) and others to the Banks (who eventually sold the manor to the Tancreds in 1613). The last 50 folios contain a survey of all the flatts and selions, giving full information about the latter (owner, area, direction, name of flatt) as well as the abuttals of the flatts. I am now engaged in the analysis of this information, and in the linking and dating of the preceding charters which are being translated afresh by Mr David Michelmore, the Archivist and Librarian of the Yorkshire Archaeological Society. (A rough translation by the late Dr Francis Collins already exists, but we have found it essential to start again from the original latin manuscript.)

The Cartulary is still in its original binding, stamped 'TB' on front and back covers, and from the survey it seems likely that 'TB' is Thomas Bank who acquired the manor through his marriage with Alice, daughter and coheiress of John le Forester de Quixley (see note 12). From this Alice Bank(e)'s will, found in the Borthwick Institute of Historical Research (see Collins (1889) and figure 21) we see that she expresses a wish to be buried beside her late husband Thomas Bank in the church of the Blessed Mary of Whixley. The will was made on 8 January 1432 (= 1433), so the survey is no later than this. The best information about the Quixley and Bank families is to be found in Glover's Visitation of Yorkshire of 1584/5 (see for example, Foster 1875). This (figure 22 shows a portion of one copy) gives exact particulars of the Bank–Quixley marriage, which took place on the morrow of St Lambert (i.e. 18 September) in 3 Hen IV (= 1402). Thus we can usefully think of the survey as having been made in, say, 1415 (merely a convenient notional date). I have spent a lot of time accumulating further information about these families, but this is not the place to set it out; it suffices to say that a picture emerges of Thomas Bank as a Lancastrian lawyer making a careful assessment of his estate on coming into his own at the opening of a new dynasty.

The Visitation record illustrated in figure 22 tells us that Thomas Bank was a younger son of the family Banke of Bank Newton (in Craven). From an elder brother (Richard) of Thomas Bank (according to Burke & Burke (1844)) there was descended a former President of this Society, Sir Joseph Banks, Bart. (1743–1820), whose magnificent portrait hangs in the Council room. If this is so, then the two branches of the family split apart with the birth of Thomas, say about 1375, and it is interesting to contemplate the curious succession of chances which have brought the two together again today, after a lapse of six centuries.

Fresh from the map of France, my first attack on Whixley was to attempt to map a portion of it by using scaling techniques combined with the Wilkinson metric. While this failed (figure 23) (and indeed it had to fail, for it involved treating not only flatts but also roads as if they were points), yet it led to interesting results. The survey is in two sections, with headings which seem to imply (what I now know to be incorrect) that one section dealt exclusively with North Whixley and the following section exclusively with South Whixley,

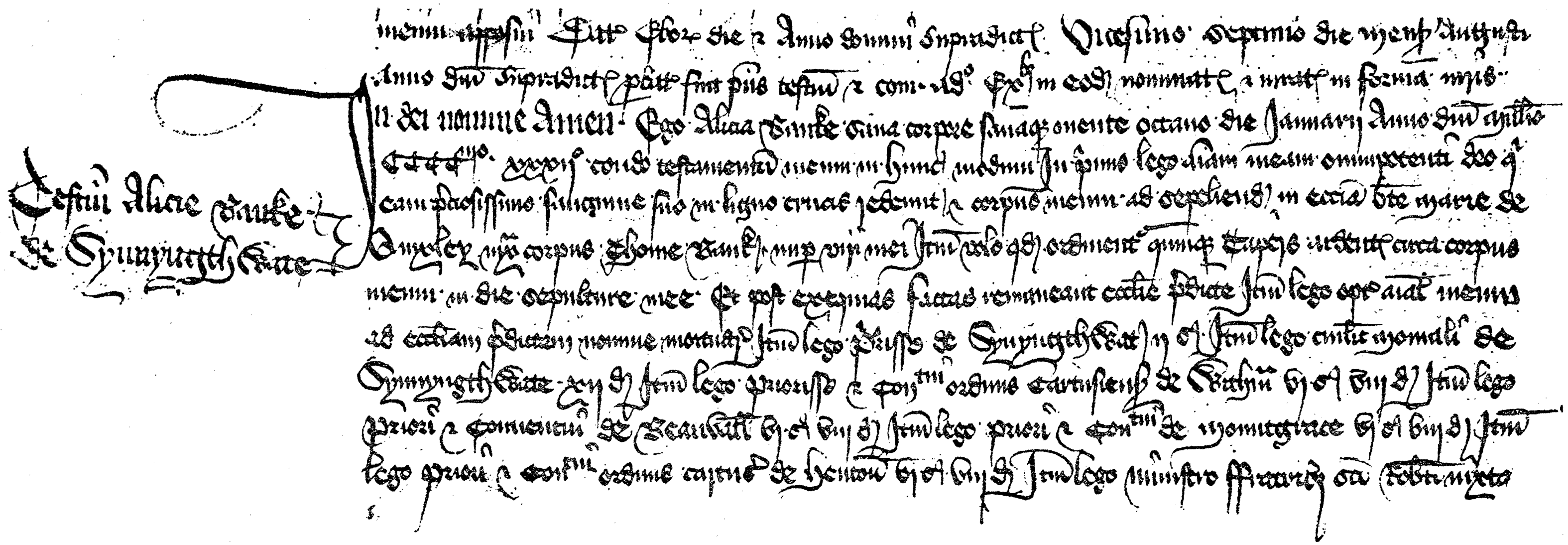


FIGURE 21. The will of Alice Banke, formerly of Whixley, later resident at the Priory of Sinningthwaite. (Borthwick Institute of Historical Research: *York Register*, vol. 3, f. 360.) Will made in January 1432–3 and proved in August 1433. Only the first nine lines of the will are shown, commencing *In dei nomine Amen*. The fact that Thomas Bank has already died is made clear in line 4.

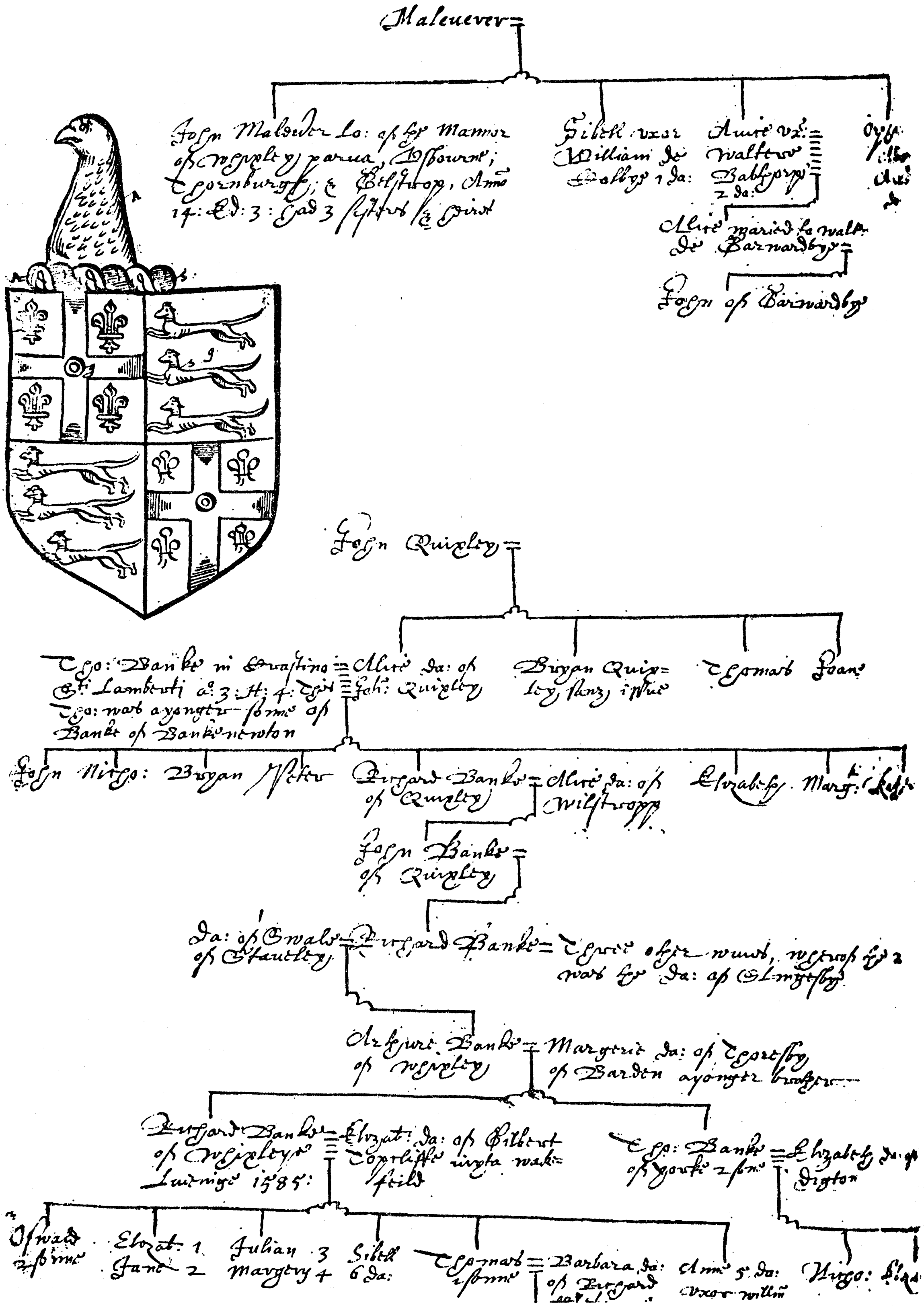


FIGURE 22. A section of a copy of Glover's Visitation of Yorkshire, 1584/5, dealing with Mauleverer, Forster (= Forester de Quixley = Quixley), and Banke of Whixley. Thomas Bank, formerly attorney in the Exchequer for the Duchy of Lancaster, became Chief Baron of the Exchequer for the Duchy. References (supposed to be to our Thomas Bank) occur in the State Papers up to and including the year 1425. He died in or before 1433. He was succeeded by his son and heir Richard Bank. The Cartulary records Richard's birth on 2 February 7 Hen IV (= 1406).

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

this distinction being an important one (associated with differentiated rights of common) and probably derived from an earlier dimidiation of the Manor which is on record.

I therefore arranged for the items mentioned in the two sections to be marked differently on the map, and you will see that while the filled circles ('North Side') fall together, the open circles ('South Side') are split into two groups. Inspection of my typescript version of the manuscript showed that the batch of 13 open-circle flatts on the top right-hand side differ from the remaining open-circle flatts in the following significant respects:

(i) they are dealt with consecutively as a single block in the survey;

(ii) this block occurs right at the end of the survey and concludes it;

(iii) it is in these flatts, but not in the other open-circle flatts, that the Prioress of Ellerton owns land (whence they lie on the North Side);

(iv) in these flatts, but not in the others, it is Richard and not Thomas Bank who holds land.

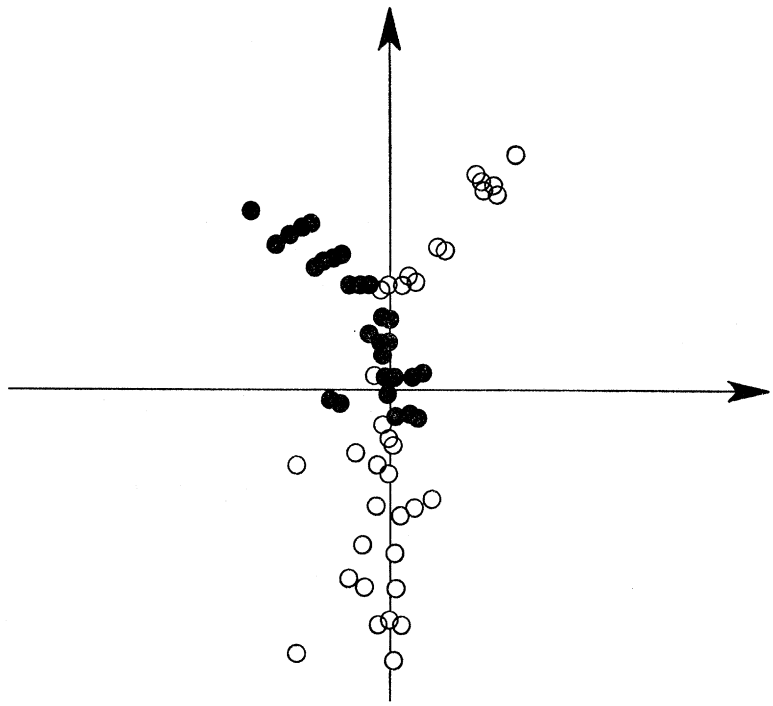


FIGURE 23. A preliminary 'map' of Whixley, made by using MD-SCAL and the Wilkinson metric based on abuttals, all topographical items being treated as 'points' (with no 'fixed' points). Full circles: flatts treated in the first (nominally 'north side') section of the Survey. Open circles: flatts treated in the second (nominally 'south side') section. Only one fully-linked group of flatts was used in this analysis. The 13 'anomalous' flatts are the 13 open circles lying to the top of the right-hand side of the figure. For the explanation of this, see text.

This suggests that the computer did well to segregate the 13 anomalous flatts, and that the entry relating to them is a later addition (thus explaining why they are apparently 'out of place'), for Richard Bank may be the son Richard who succeeded his father Thomas in 1433 or earlier. At this point I remembered that there is a marked change of hand towards the end of the survey, and so wrote to the Yorkshire Archaeological Society to say that, in the computer's opinion, a change of handwriting should occur between folio 161 and its *verso*. Mr Michelmore then wrote back to say that the computer was right. This experience taught me that even a bad map can help to draw attention to significant topographical, genealogical and paleographical facts, and I like to think of it as a modest triumph for the computer concerned.

5. A new approach to the Whixley problem

By now I had had enough of scaling methods and had begun to realize that the Whixley problem requires a custom-built program designed to serve its own peculiar needs, so I proceeded to develop this; I call it FLTTMP (= FLATTMAP).

The idea behind the new program is to represent the flatts of Whixley as stars in a stellar cluster subject to rather peculiar 'gravitational' forces, as follows.

Linked flatts (i.e. those which are *known* to abut) will in general be required to attract one another.

Unlinked flatts (which either do not, or *are not known* to abut) will in general behave neutrally towards one another.

Flatts whose position is known from some external information will be fixed at their true positions.

Identified roads (and bridle paths, streams, etc.) will be represented by line segments approximating to their true positions, and flatts known to be linked to these by abuttal will be attracted to them *laterally*, but without any special preference for one side or the other.

Now such a system of forces alone is inadequate, because apart from the moderating influence of the few 'fixed' elements, there will be a tendency for all the 'variable' flatts to coalesce, yielding perhaps collisions (bad computationally) and a hopeless map. It is clear that two flatts with different names each occupying say 20 acres must be in different positions. Accordingly we introduce *short-range repulsive* forces between *all* pairs of flatts whether linked or not, and between *all* flatt-road pairs, in such a way that two flatts will start to repel one another when their separation is such that, in view of their known size, they would touch, and similarly for flatts and roads.

This, however, produces a consequential difficulty. We are all familiar with the 'cocktail party' effect; the difficulty of forcing one's way through a roomful of strangers towards some person with whom one is, or would like to be, linked. To resolve this, I divided the computation into two phases, and only in the second phase is the system of repulsive forces allowed to operate. During the first phase there is a 'see-through' effect which enables one to glide magically across the room and collide with one's partner. Then Phase 2 intervenes, and the happy pair stand off from one another by a seemly amount.

It turns out that this system works splendidly. Let us look first at some diagrams (figures 24 and 25) illustrating an artificial problem involving 49 'flatts', with first 4 and then 3 of them fixed. In these diagrams 'flatts' are shown as filled circles unless they are 'fixed' (in which case they are shown as open circles). Two circles are linked by a line segment if and only if the corresponding 'flatts' actually abut. Figure 24*a* shows the random starting configuration, while figure 24*b* shows the computed configuration after sixty iterations (thirty in each phase); this is very close indeed to what was actually the 'parent' configuration for the problem – a sort of cut-down 7×7 chessboard. As usual, crossed links indicate a topologically impossible situation; these two figures show very clearly how FLATTMAP resolves such impossibilities.

In this last example there were 4 'fixed' points, these being the four 'corners' of the 'chessboard'. It seemed better to have fewer (say 3) 'fixed' points, and to choose these on the boundary, but not at points so privileged as the corners. What happens then can be seen in figure 25*a*; we get a turned-over corner (the 'elephant's ear' effect). But by re-running this example with increased tension in Phase 1, the crease is unfolded (figure 25*b*).

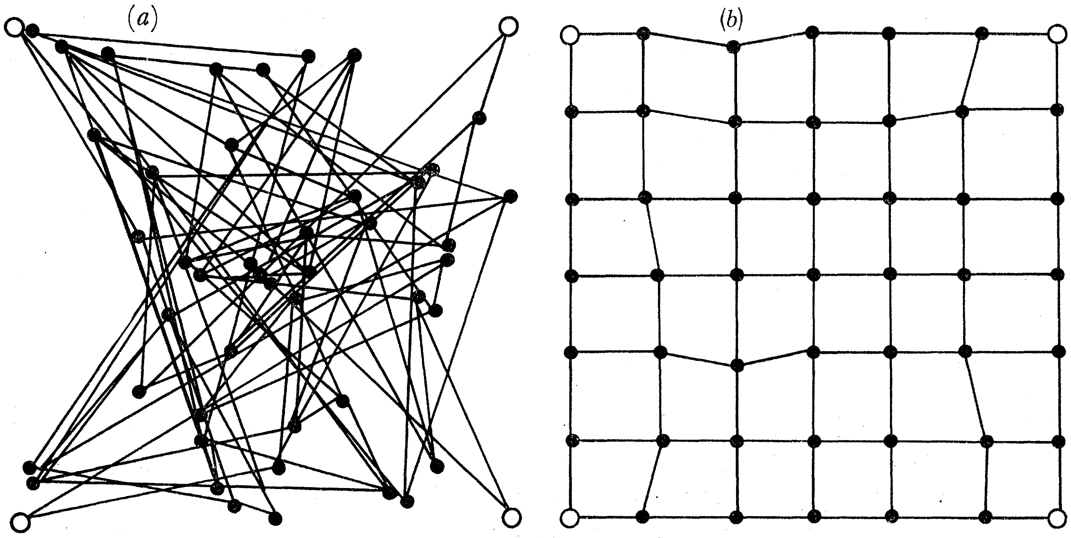


FIGURE 24. Seven-by-seven 'chessboard' with four corner-squares fixed. (*a*) The starting configuration. (*b*) After 2×30 iterations.

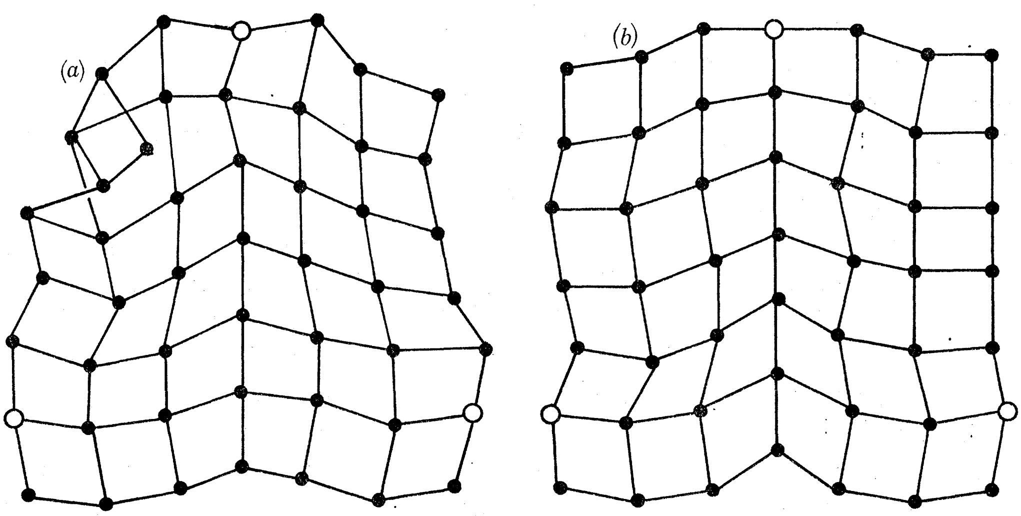


FIGURE 25. Seven-by-seven 'chessboard' with three mid-side squares fixed. (*a*) 'Elephant's ear' effect. (*b*) Corrected configuration achieved by increasing the tension in Phase 1.

In a final experiment in this series the intensity of linkage was deliberately reduced. It will be remembered that for the map of France the typical linkage-figure was 5 or 6 abuttals per flatt, and in the experiment just described this value was reduced to 4. But this is still too high to be representative for the manorial problem, where an average of 3 is appropriate. What makes the manorial problem technically so very interesting is the feeling one has that this value 3 may prove to be *critical*, the margin of the possible. Accordingly by a random

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

rejection process simulating the haphazard run of the selions, the number of links was cut down to an average of 3, and at the same time two 'roads' were thrown across the pattern, links formerly running *over* these being replaced by twice the number of links *to* these.

The 'roads' caused no particular difficulty, but the reduction in the number of links created in a number of cases a situation in which a flatt was connected to the rest of the pattern by one link only ('lamb's tails'). These do indeed 'wag', and what is worse they can

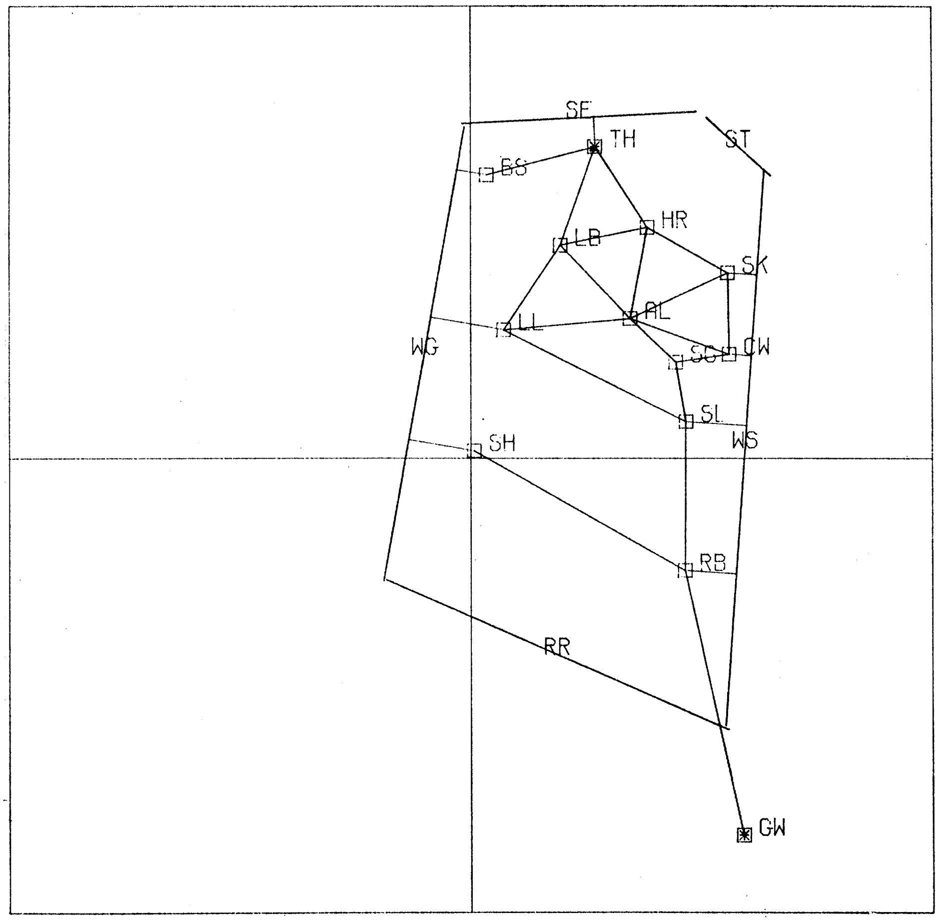


FIGURE 26. Computer-map of mid-south Whixley (TH and GW fixed). RB = *Rowmerbergh* is placed by the computer close to Roman Barfs on the Ordnance Survey map, with which it is subsequently (in figure 28) identified. (Here and in figure 27, the lost way *Sowrelandegate* is treated as if it were a flatt.)

get trapped inside the pattern when they should lie out towards its boundary. To combat this, a strong wind was introduced in Phase 2, blowing outwards from the centre, to extend the 'lamb's tails' (the 'Bo-peep' effect). This worked as intended. Another result of lowered linkage was an increase in 'overlap' anomalies. While these are easily detected by eye, and as easily corrected by hand (followed by a fresh series of iterations to sort the pattern out), it would be more satisfactory to automate this procedure, and I hope to do so. In the interests of space the diagrams illustrating this last experiment have been omitted. It is in any case time for us to start working with real data.

We are now ready to look at the first computer-map of part of Whixley as it may have been in about 1415 (figure 26). Only a sample portion of the whole manor is dealt with here

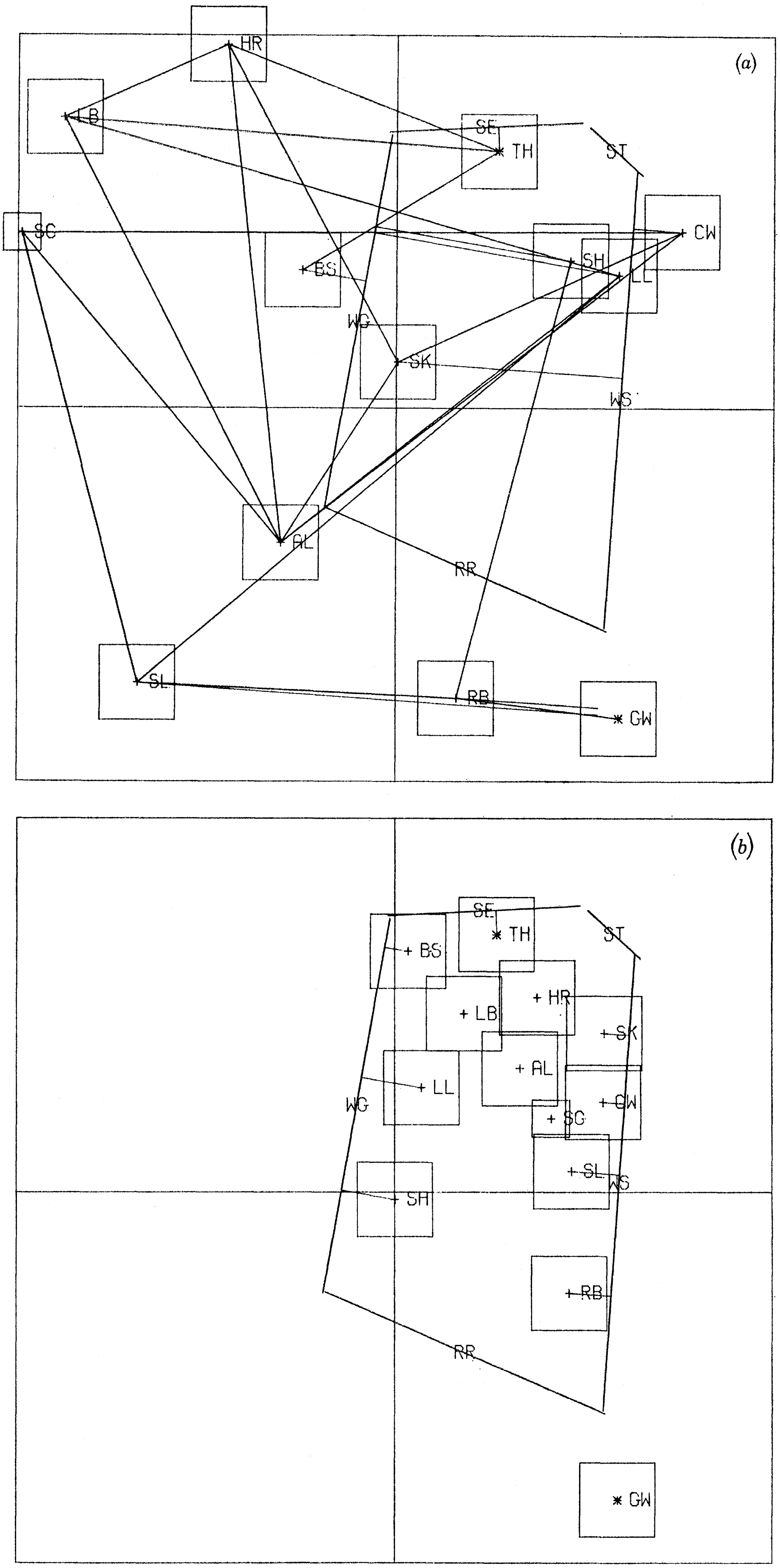


FIGURE 27. Computer-map of mid-south Whixley, with flatts of correct areas. (*a*) The random initial configuration. (*b*) The computed configuration after the completion of both phases of the computation. (The 'area' of GW is notional, and *Sowrelandegate* is shown as if it were a small flatt, SG.)

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

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

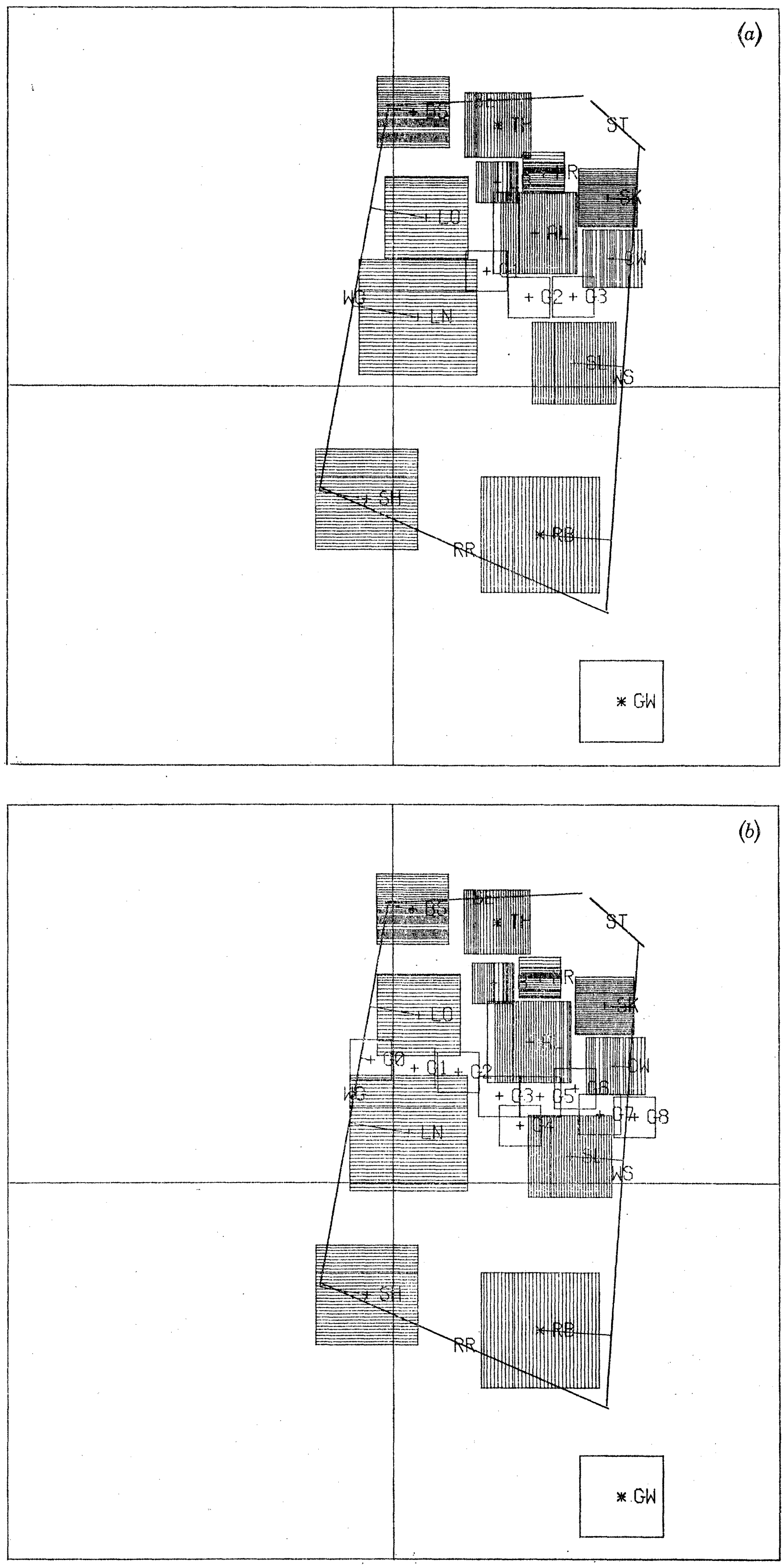


FIGURE 28. Computer-map of mid-south Whixley, with RB fixed, and selions shown. (*a*) *Sowrelandegate* is now a free chain with 3 links. (*b*) *Sowrelandegate* is now a chain of 9 links, tethered to *Watlyngstrete* on the right and to *Wederbygate* on the left.

(cf. figure 19); this is isolated by a quadrilateral of surviving roads, etc., and is small enough to permit a series of experiments with a moderate turnround time. The flatts are located in the plot as small squares, and the linkage of abutting flatts is as usual shown by line segments joining appropriate squares. The flatt names are indicated by a convenient two-letter code. 'Fixed flatts' appear as squares containing the symbol $*$. 'Roads', etc., (also labelled) are shown as line-segments. The plot is produced on transparent paper originally at a scale of six inches to one mile, and the centre of the plot is GR 440570.

Two items are 'fixed'. These are the flatt TH = *Tofillynges* (identified via the Enclosure Award Map and a late eighteenth-century deed with Toft Hill Ing), and GW = *Gildeswath* (identified, after an initial experiment in which it was allowed to 'float' freely, with Gilsthwaite Bridge). The only topologically impossible situation is the crossing of RR ('the road to St Robert's') by the RB–GW link. A location adjustment is obviously called for here. It may be that the road (RR) has been placed too far north, or GW too far south. Later work with other portions of the manor may offer some control over this. Ideally one would like to have the Turnpike Trust papers for the Knaresborough–Green Hammerton Turnpike (= RR), but these have not yet been found, and may not survive.

Figure 27*a* shows a typical 'random' starting configuration, and now the squares representing the flatts have been expanded so as to have nearly the right areas. Figure 27*b* shows the map after the completion of Phases 1 and 2. The square flatts overlap a little, but not so much as might have been feared when it is recalled that they were almost certainly *not* square, and also that we do not know for certain how to relate the local fifteenth-century acre to the modern acre (except via exercises such as the present one.)

After this the computer was asked to draw in all the selions with their proper widths (it was assumed that selion areas differ only via a variation in width; this may not be correct). Previously the lost way called '*le Sowrelandegate* which leads to Hunsingore Mill' was treated as if it were a flatt SG of small area, but now (figure 28*a*) it is treated as a freely floating chain with 3 links (marked G1–G3). There is a suggestion in the Cartulary that this way ran across the area treated, from *Watlyngstrete* on the east to *Wederbygate* on the west, and so in figure 28*b* it has been replaced by a chain of 9 links (labelled G0–G8) and G0 has been linked to WG (= *Wederbygate*) while G8 has been linked to WS (= *Watlyngstrete*). This is my best map to date. Obviously it now needs considerable hand adjustment.

On the whole the gaps worry me more than the overlaps. Gaps are of course to be expected because we must leave room for common and meadow. Concerning the meadow closes, we have a list of these (and at least one of them can be identified by name), and we know their relative areas expressed as 'numbers of days to mow', so it should be possible to estimate their actual areas, and thus perhaps to fit them into the holes in the map. The direct east–west line for *Sowrelandegate* may be an artefact, but on the other hand it is hard to accommodate the detailed abuttals with a southwest to northeast line as suggested by Jefferys' map (figure 19). It is of course quite possible that 'the road to Hunsingore Mill' may have straightened itself out between 1415 and 1771. The new cache of deeds may throw light on such points as these.

The reader is asked to note the position of the flatt BS = *Buretrestubbes* and to compare its selions with those visible in the aerial photograph (figure 29), most kindly taken for me by Professor J. K. S. St Joseph. Here again, we cannot be certain that the 'spectrum' of selion widths has remained constant, but it will at least be worth while trying to match the plot

FIGURE 29. Aerial photograph of Whixley township (top), Whixley Park (with trees, and ancient selions, top left), *Wederbygate* running southwards (downwards as you look) from the Park, and *Buretrestubbes* flatt south of the crofts and to the right of *Wederbygate*. (Photograph by Professor J. K. S. St Joseph: RC8-AF274.)

(*Facing p.* 572)

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

with the photograph, particularly as the Cartulary tells us in detail how and where the windmill was to be built on three of the selions in *Buretrestubbes* in the fourteenth century.

The small portion of the Manor of Whixley treated here is of course, quite easy to deal with by hand. It is being used here for illustrative purposes, and in the investigation itself as a test-piece on which to check the consistency and successful functioning of the various technical devices as they are developed. A full-scale analysis of the Manor will follow at a later date, and it will incorporate such further information as may by then have become available.

6. Applications to seriation

I am a great believer in the advisability of paying one's debts. For example, the Vicar of Whixley has shown me endless kindnesses, and so it was a great pleasure to be able to tell him, from Alice Banke's will, what was the lost dedication of his church.

But I think that one should also pay one's debts to one's self; archeological seriation led to the Whixley problem, by induction on the number of dimensions. It is now interesting to ask whether the program FLATTMAP can be used instead of my earlier seriation algorithm HORSHU to carry out archaeological seriations. The answer is affirmative.

Here it is interesting to quote Flinders Petrie, whose remarkable studies of pre-historic Egyptian chronology have been the inspiration of all subsequent work on seriation, as saying in 1901, with reference to the slips of cardboard on which in coded form the list of contents of each Egyptian grave was recorded, that he envisaged

> an elastic thread for each type [of pottery], attached to all the slips containing that type... the resultant position of all the slips under the tension of all the threads will be the probable truth; the weakness of each thread being in proportion to the true extent of diffusion of its type.

The idea of a FLATTMAP type of seriation procedure is clearly present in this quotation, although it is obvious that Petrie is speaking only analogically. An actual mechanical model of the kind he describes would, if let off in a random configuration, result in a horrible tangle of threads. But within the 'disk' of a computer we can have notional 'threads' with the delightful property that they do *not* become entangled, because they can pass through one another, and so it is that some 75 years later Petrie's experiment can be and now has been actually carried out.

We begin with a random configuration in one dimension. We equate each 'grave' with a 'flatt'. There will be no 'fixed points', and no 'roads'. The 'see-through' effect in Phase 1 will totally eliminate the collision problem which otherwise arises when one attempts to seriate directly in one dimension. But we have to introduce a new element in order to overcome the possibility of getting an equilibrium situation in which all the threads are slack, and all the 'points' (now 'graves') lie close together in a loose huddle under the mutual attractions of the 'linked points', these attractive forces now being taken to be proportional to the similarity-measure K_1 described earlier in this lecture.

This difficulty is overcome by introducing yet another mock astronomical 'effect': the expansion of the Universe. *All* 'points' will now be repelled from the 'centre' in such a way that the root-mean-square dispersion of the 'points' in one dimension is held at a pre-assigned fixed value, say λ times the sum of the lengths of the linear intervals which represent the 'graves', so that we can think of λ as a sort of 'cosmological constant'. The algorithm is then run with a Lagrange multiplier subroutine which holds the linear dispersion of the 'graves' at the set value.

Three closely concordant runs of FLATTMAP with this modification have been computed, by using the Münsingen–Rain data of Roy Hodson (1968) (the 59 La Tène graves containing in all 70 varieties of fibula). On pooling the three FLATTMAP seriations we obtain the result shown in figure 10. It is not quite so good as the HORSHU seriation which we saw earlier in figure 9, but is none the less very striking. More seriations than this would normally have been carried out, and it has not been possible yet to study the effect of varying the 'cosmological constant' λ, which was set and held at a value merely guessed to be appropriate. Subsequent work with FLATTMAP should thus yield better results than this. A question of great interest is whether FLATTMAP will be more appropriate than HORSHU when working with very large assemblages, where time and storage limitations can be serious.

One argument against seriating directly in one dimension is the difficulty of dealing with collisions; points must be allowed to 'pass through' one another, or one may in extreme cases find one's self left with the random starting order. I have explained already how that difficulty is dealt with by part of the in-built structure of FLATTMAP. Another more serious argument against working in one dimension is that a one-dimensional seriation procedure cannot fail to yield *some* seriation, even if the material does not justify such a treatment. You will remember that this is why HORSHU is always run in the first instance in two dimensions: so that it can 'fail' if it wants to. However, one can adapt FLATTMAP to help to meet this criticism by using Phase 1 to generate and record the seriation, and Phase 2 to provide an 'evaluation' of it, by taking the following steps.

(i) Move with the one dimensional seriation into two dimensions.

(ii) Locate the two representative points which are at either end of the generated seriation, and *fix* them, leaving the remaining N-2 points *free*.

(iii) Transfer the origin of coordinates to the centre of the segment joining the 2 fixed points.

(iv) Magnify the strengths of all links between variable points and the 2 fixed points.

(v) Scale up the configuration along the dimension representing the seriation, so that the 'threads' in the interval between the 2 fixed points are under high tension.

(vi) Now introduce a *one dimensional* gale blowing outwards from the origin *along the dimension of the seriation* (to extend the 'lamb's tails').

(vii) Introduce also a uniform cross-wind blowing *at right angles* to the seriation, so as to distend the segment between the 2 fixed points into an arc-shaped form.

(viii) Finally, ask the plotter to draw linear segments between all pairs of representative 'points' which are strongly linked (corresponding to the 'strong links' of HORSHU).

We now look to see *whether the 'strong links' respect the generated seriation*. Notice that we are thinking now about real life problems in which the answer is not known in advance. We have a seriation from phase 1 – we cannot fail to get one – and we want to know if it is truly concordant with the linkage structure on which it was based. If we have a bad seriation we shall see that this is so straight away, because the 'strong links' will run criss-cross over the figure, just as they do at the outset of a flattmapping operation. But in fact (figure 11) this is not so; they very largely respect the generated seriation, and, of course, in the Münsingen case we know that they *must* do so, because, as we have seen, Hodson's order, the HORSHU order, and the FLATTMAP order are all essentially the same. I am, however, not entirely satisfied with these 'flat-chested' horseshoes, and intend to make the cross-wind non-uniform, in order to give them more 'bosom'.

For the same reason there is no need to carry out what in practice would be the other obvious and very necessary test of the generated seriation, which is to re-arrange the *rows* (representing the '*graves*') of the 'incidence' (or 'abundance') matrix in accordance with the generated seriation in order to see whether it then follows, as it ought to do, that in each *column* (representing a '*type*' of pottery or jewellery) the entries different from zero are blocked together (in the 'incidence' case) and are unimodal (in the 'abundance' case). The FLATTMAP ordering, being so close to the Hodson ordering, which in its turn is closer still to the HORSHU ordering, cannot fail to pass this test too, but there has been no time (and there is no need) to carry it out.

7. Closing remarks

I wish to close with a very important comment addressed primarily to the mathematicians in the audience. If you are successful in this kind of operation, you will at the most have been able to provide the archaeologist or historian with a *very bad* seriation, or map, which he will then immediately proceed to refine with the aid both of his own professional judgement and of external information. What he gains from your activities is the knowledge that his starting point is a long way better than a purely random one, and that it is based on identifiable aspects of the data and not on presumption, or whim. You will also be able to supply a series of independent seriations or maps, derived from different random starting configurations, so that if the problem admits several widely different solutions then you will discover several of these, and your colleague will have been warned of the extent to which the solution of his own problem may be indeterminate without additional information. *You must not expect to achieve more than this*, and you must actively warn your colleague against the facile acceptance of your solution. The really hard work will only now be starting, and he must be left in no doubt of that. For the rest, you will find that the investigation is its own reward. To have gained even a little insight into another discipline, to have tramped the fields of your manor and learned to read what is written in the landscape all around you, and to have learned also to read, however stumblingly, the quaint latin of a scribe dead five centuries ago, and checked his scrupulously accurate arithmetic, and realized that here indeed you have found a numerate colleague: these are rewards you will long savour. And this is to say nothing of the best of all: the friendships which spring naturally out of such interdisciplinary frolics.

But if you feel tempted to go and do likewise, keep well away from Whixley. One mathematician, one manor: otherwise it is more than elastic threads that will get crossed.

I wish to acknowledge the kindness of the following persons and corporate bodies in allowing me to reproduce copyright material: the Committee for Aerial Photography, Cambridge University (figure 29); the Director of the Borthwick Institute of Historical Research (figure 21); the Director of Leisure Services, Humberside County (figure 22); the Edinburgh University Press (figures 5–9 and 15); the Editor of *Nature* (figures 16 and 17); H. Margary, Esq. (figure 19); and the Master and Fellows of Christ's College, Cambridge (figure 20). I further wish to thank the Director of the Cambridge University Computing Service for computing facilities, and my daughter, Bridget Kendall, for kindly verifying some heraldic details.

I also owe a very special debt of gratitude to Major Dent, without whom we would have no Whixley Cartulary, to Mrs Dorothy Owen and Mr David Michelmore, without whom I could not have learned (a little) to read it, and to the good people of Whixley for making me so free of their lovely village. It is impossible in this small space to acknowledge the many kindnesses of others, from those shown to me by Sir Anthony Wagner, Garter Principal King of Arms, to those of my newly found Styan cousins, but I hope that they all realize how much I owe to them.

Notes

1. *Preliminaries*. The phrase 'of a known general form' requires emphasis. The present class of problems must be clearly distinguished from others, perhaps superficially similar, where the motivation of the analysis of a data-set is 'to see whether any interesting structural features turn up'. This last is an entirely respectable activity and has generated much distinguished work, but in the present group of problems we know what kind of structure we are looking for, and if a totally different kind of structure 'turns up', then we have failed. These remarks are the more necessary because one of the techniques to be described ('scaling methods') was originally designed for and is most often used for the analysis of data without any presumptions about the structure in it. As will be seen, however, scaling methods can be successfully employed in each of the two roles referred to here.

2. *Serial proximity*. In 'seriation' (or 'sequence-dating' as it was called by Flinders Petrie) the aim is merely to place the items studied in the correct or nearly correct time *order*; it is not intended that individual 'steps' in this sequence should represent equal intervals of time, nor is it usually expected that the method employed will distinguish between one acceptable order, and its reverse. 'Which way round to read the order' (i.e. 'the sense of time's arrow') is to be determined by external evidence, which in an archaeological context might be, for example, ^{14}C dating or the results of some similar technique.

3. '*Incidence*' *and* '*abundance*' *matrices*. In an incidence matrix we are merely informed whether some 'type' is (1) or is not (0) present in some 'grave'. All the statements made about incidence matrices can be and have been generalized to abundance matrices, where the extent of the 'presence' is quantified by stating the actual or relative number present. For this, see Kendall (1971 *a*, *d*).

4. *Petrie form*. The analogue of this which is used in the treatment of abundance matrices requires that in each column, the sequence of entries is 'unimodal'. That is, it is decreasing, or it is increasing, or it increases to a maximum and then decreases. The idea behind the definitions is of course that of an innovation coming into and then going out of fashion. No one really believes, in the archaeological case, that the development and decay of fashion is as simple as this; nor, if it were so, would the data necessarily faithfully reflect it (there are several obvious reasons for that). But the idea forms a convenient starting point from which to approach the seriation problem, and one seeks a technique which will work perfectly for matrices which *can* be given the exact Petrie form, and which will still work reasonably well when the matrices can only approximately be given the Petrie form. Figure 8, p. 554 will give some idea of the extent to which deviations from the Petrie form will be tolerated by my HORSHU and other algorithms to be described in this paper. (In this figure, the 1's are 'spots', and the 0's are 'dots'.)

An important question is, how can we know when we are dealing with a 'pre-Petrie' matrix (one which can exactly or approximately be thrown into the Petrie form by row re-arrangement)? An important early study of this question was made by Fulkerson & Gross (1965); further contributions will be found in Wilkinson's monograph (1974) and important new contributions have recently been made by R. R. Laxton (1975). One solution to the problem is to suppose that the answer is affirmative, to use one of the algorithms, and to see if the output is acceptable. If it is not, the conclusion is then that we were mistaken. But this can be expensive.

5. *Petrie's seriation.* Sir Alan Hodgkin has pointed out that Petrie's feat tells us something about the superiority of the performance of the human brain in carrying out such tasks when compared with that same brain's ability to design 'artificial intelligences'. We are apparently better at taking short cuts, than at formulating the principles of short-cutting.

6. *Similarity measures for seriation.* Let N_{ih} be the abundance measure (1 or 0 in the incidence case) for the ith 'grave' in relation to the hth 'type'. Then the first similarity measure for two 'graves' labelled i and j is

$$K_1(i,j) = \sum_h w_h \min (N_{ih}, N_{jh});$$

here the w's are arbitrary positive weights. Once K_1 has been set up in this way, the higher-order similarity measures are constructed recurrently by using the formula

$$K_{r+1}(i,j) = \sum_k w_k \min (K_r(i,k), K_r(j,k)).$$

Notice that in setting up K_1, we compare two 'graves' i and j via the different 'types' labelled h, while in setting up the higher-order similarity measures, we compare them via *other 'graves'* labelled k. Notice also that is is not necessary for the same set of weights w to be used throughout; this can change from one application of the formula to the next, even for the same data. The theorem relating to these similarity measures holds whenever the abundance (or incidence) matrix can be thrown into the Petrie form, by some rearrangement not yet identified. *In practice*, it turns out that the predictions of the theorem are realized even when the data-matrix only has this property in some approximate sense, and in this 'robustness' lies one of the main virtues of the HORSHU method.

7. MD-SCAL. This is the name of the Shepard–Kruskal algorithm; at greater length it is 'non-metric multi-dimensional scaling', and it is designed to work in an arbitrary number of dimensions (here usually 1 or 2). It is important to grasp that the algorithm works basically with '*pairs of pairs* of objects'. Each 'object' is represented by a point in a map (initially a random configuration), and the algorithm progressively modifies the map in such a way as to minimize 'mismatches between pairs of pairs'. Let (A, B) and (U, V) be two pairs of 'objects', and let (a, b) and (u, v) be the corresponding pairs of points in the map. The 'objects' have been endowed with a similarity measure S; suppose that $S(A, B)$ is greater than $S(U, V)$. Then in its 'global' version the algorithm strives to eliminate the mismatch which arises if it happens that, in the current map, distance (a, b) is greater than distance (u, v). Of course, in a small number of dimensions there is insufficient room for manoeuvre to permit the resolution of all mismatches, and because of this a compromise is built into the algorithm. It can be run in various modes. One variation, referred to here as 'global scaling', examines *all* 'pairs of pairs', and another ('local scaling') examines, say, only pairs of pairs of the form $((A, B), (A, V))$ which have one component in common. Another variation is

in the treatment accorded to 'tied' similarities. For these technicalities reference should be made to Kruskal's articles (1964*a*, *b*), to various papers in the Mamaia volume (Hodson *et al.* 1971), and to Sibson's article (1972). The simplified account in the text refers to 'global' scaling with the 'secondary' treatment of ties.

8. *The map of part of the United Kingdom.* For the sake of a better understanding of what comes later, the reader should make sure that he knows what information was and what information was not made available to the computer which 'drew' this map. It was supplied with the contents of the table in figure 13; it did *not* 'know' the distances listed in figure 12. After production the map was oriented by hand, with Cardiff to the west, York to the north, and London to the east; it is clear that information of that sort could not possibly have been conveyed by the contents of figure 13. It is normal for such computer maps to be drawn on transparent paper, and this is a great convenience because it makes it easy to correct for the 'reflexion error' which can occur (and indeed will occur with a probability equal to one half). I am indebted to Dr J. H. Andrews of the Department of Geography, Trinity College, Dublin, for the information (Andrews 1974) that the maps of the escheated counties of Ulster (1609–10) occasionally contain such a 'reflexion error'; he says:

> in two of them [the maps] the whole of the outline, together with the positions (though not of course the actual writing) of the names, appears as if in mirror image with east and west reversed... As a plotting error, when plotting is done from purely verbal data, this mirror effect is easily understood; given verbal abutments but no compass points, a plotter is in fact just as likely to produce a reversed outline as a correct one.

In the reference cited above the reader will find an account of the way in which these seventeenth-century Irish maps were made, and also many further remarks of great interest in the present context.

9. *The map of Otmoor.* It is important to notice a special convention in figure 15, where the line-segments serve to remind us of the *true* spatial relations in the situation; they were *not* available to the computer, and were *not* used in the construction of the computer-map, which was based solely on the rank-ordering of the standardized intermarriage rates. Elsewhere in this paper, however, line-segments are frequently used to indicate the presence of a highly-ranked similarity in the data used by the computer.

10. *The map of France.* This problem is of type II, because the pairs of Departments are simply sorted into two groups; those that 'touch', and those that 'don't touch'. The Wilkinson metric (it was E. M. Wilkinson (1974), who kindly suggested its use to me) plays in this context a role rather like that of the higher-order similarity matrices employed in the seriation problem. Its computation can be lengthy and bulky; Wilkinson has written a very quick and compact algorithm whose use is recommended in problems of this kind. In a practical problem we would not normally know *a priori* whether the 'touch–don't touch' data was sufficient to link all the objects into a single connected graph. If the graph happens to be disconnected, then of course some of the Wilkinson distances will not be properly defined, and the method must then be used separately on each connected component. For example, this would happen with France if the Department of Corsica were included. In fact both Corsica and Seine (= Paris) were excluded; this is why the number of Departments is 88 and not 90. The numbering used is not the official coding used for French Departments,

but an artificial one designed to make it easier to compare the two maps. It runs from left to right, and from top to bottom, in 'typewriting order'.

The computer supplied 'point' positions for each Department. These were linked by the line-segments indicating the contiguities which were employed in the computation, and the polygonal cells intended to represent the Departments themselves were then constructed by hand, by joining the centroids of adjacent triangles (or quadrilaterals) composing each 'star' corresponding to a Department-point. This method cannot be used if a Department has a maritime or alien boundary, and then a variety of artifices was used to obtain 'reasonable' external boundary lines for these 'cells'. If this experiment were ever to be repeated, an automatic and more systematic passage from the 'point' graph to the 'cell' graph would be desirable, but the map shown here at least suffices to indicate how a close approximation to the original geometry can be recovered from abuttal data.

This map, although based only on abuttals, owes its accuracy both to the large number of 'objects' which are to be placed in position, and to the richness of the linkage between them. If this number of 'objects' is N (here equal to 88), then given a possible cellular map the labels could have been placed on it in $N!$ different ways, so that the amount of information (in a technical sense, and measured here in 'naperian' bits) involved in a successful selection of the correct labelling is

$$N \ln N - N + o(N)$$

(the corrections due to the fact that a correct orientation is not asked for are swallowed up in the error-term). Now suppose that a typical 'object' is linked on average to k others; then there will be a total of $\frac{1}{2}kN$ such links, in a problem of type II, and an equal number of *declared* links, in a problem of type III. The total number of all possible links is $\frac{1}{2}N(N-1)$. Hence the data consists precisely in this, that a certain subset of size $\frac{1}{2}kN$ is specially distinguished within the set of all $\frac{1}{2}N(N-1)$ links. To have been presented with such a subset is to be presented with

$$\tfrac{1}{2}kN \ln N + \left(\tfrac{1}{2}k \ln \frac{\mathrm{e}}{k}\right) N + o(N)$$

naperian bits.

Suppose first that k is less than 2 (of course k need not be an integer). Then the above argument shows conclusively that for large N our problem is insoluble; the amount of information available is insufficient to enable us to recognize the true configuration from among all the logically possible ones. Next suppose that $k = 2$. The major terms in the asymptotic expansions now coincide, but at the level of the second terms we have an advantage amounting to $1 + \ln(\frac{1}{2}\mathrm{e})$ naperian bits per 'object', so that if the number of 'objects' is large then we may hope to be successful. When k is greater than 2 we are very much better off, because now the leading terms of the expansion are discrepant, and for large N we shall have more information than we need in a multiplicative rather than in a merely additive sense. For example, in the map of France where (at least in the interior) k is about 5 or 6, we have something like 2 or 3 times as much information as is required. There are, however, snags in this argument. We may have 'enough' information, but it may be weakened by concealed redundancies, it may be corrupted by error, and what is relevant and error-free in it may be difficult to extract. While the caveat is often left unstated, information-theoretic arguments of this kind can only give reliable *negative* answers; at the positive level they are at most suggestive of the true state of affairs. The reader must also be reminded that we have supposed

N to be large enough for the approximations to work; quite moderate values of N are normally sufficient for this, however. Perhaps a more serious consequence of 'small N' is that then the relative effect of the boundary becomes greater (maritime Departments have few neighbours), and would be more marked still if we were trying to map a highly contorted region which was 'almost all boundary'.

We have suggested that formally $k = 2$ is the threshold across which behaviour changes, but in practice I suspect that we will feel its effect even when $k = 3$. This is one reason why the manorial reconstitution problem, next to be studied, is technically so interesting. For there, on the simplest model, we should expect $k = 3$, and so we shall be trying to work in the vulnerable region.

When considering the chances of success with a given k-value, the terms of reference of the investigation must be considered; we are looking for structure 'of a known general form'. Thus, if we are looking for a (two dimensional) map, then we will not normally use $k = 2$, because that could only give us a seriation, or perhaps a disjoint set of seriations. It would be an interesting academic exercise to try to reconstruct the map of France with $k = 2$. Ideally one should get a 'ribbon' capable of being laid down in the plane in such a way as to yield a true map.

11. *The 'lost hoard' of Whixley deeds.* Possible locations for this are the institutions associated with the *ex officio* Governors and Trustees of Christopher Tancred's Charity: the Master of Christ's College, Cambridge; the Master of Gonville and Caius College, Cambridge; the President of the Royal College of Physicians; the Treasurer of Lincoln's Inn; the Master of the Charterhouse; the Governor of the Royal Hospital, Chelsea; and the Governor of the Royal Hospital, Greenwich. Other possible locations are the offices and institutions associated with the Clerks and Receivers of the Charity, and the Stewards of the Manor. All these possibilities have been or are about to be explored. A search of the Masters' Exhibits in the Public Record Office yielded a few interesting Whixley deeds, but none of those listed in the Christ's College calendar. Further searches will be made in the Fine Rolls, and in the Yorkshire Registry of Deeds.

12. *The Quixley family.* The members of this family are known sometimes as Quixley, sometimes as Forster, and occasionally as 'le Forester de Quixley'. The late Dr Francis Collins suggested that perhaps Forster was an occupational (and Quixley a locational?) surname, and so we may here be concerned with a family possessing a hereditary title of Forester, perhaps in relation to the group of manors anciently possessed by the Mauleverers of Quixley. In this connexion it is interesting to note that the Quixley/Banke entry in Glover's Visitation (see for example figure 22) is preceded by a reference to John Mauleverer as Lord of Whixley, Little Ouseburn, Thornborough, and Gelsthorpe in 14 Edw III (= 1340–1341). (The apparent reference to 'Whixley parva' is a copyist's error. There should be no comma after 'parva', which qualifies 'Usbourne'.) From the Whixley Cartulary we know that on his death the Manor of Whixley passed to his 3 sisters and co-heiresses named in the Visitation (Sibilla, Avice, and Orfania), and we know how it passed through them and their families and eventually was reunited (via John de Gerwardby and others) in the hands of John Forester de Quixley (the father of Alice Bank, who was the wife of Thomas Bank). These complexities are still being tidied up. However, it seems worth noting that the Johnston manuscripts in Bodley's Library, Oxford, contain a lengthy Mauleverer pedigree, which assigns to John Whixley of Whixley Esq. (fl. 51 Edw III = 1377) (the father of Alice Bank(es))

a variant of the traditional coat of the Mauleverers of Quixley (gules, three greyhounds courant in pale argent, collared sable(?)bezanty) differenced† by assigning to the hounds (i) a collar compony or and sable and (ii) on the shoulder an escallop (?) sable (?).

It will be noticed (figure 22) that the Bank family quartered (in 2nd and 3rd) a coat also clearly derived from that of the Mauleverer family, although the details of the collar appear to be different in this copy of the Visitation, and the escallop (or whatever it was) has disappeared. The particulars of the hounds and of their collars vary from one copy of the Visitation to another, but they are printed by Foster (1875) as 'argent' and 'gobony d'or et sable', respectively, and this appears to be correct. The implication seems to be that John le Forester de Quixley was in blood a Mauleverer, and hence also those members of the Bank family descended from Alice Bank, although the possibility cannot be wholly excluded that here, as occasionally may have happened, the coat descended in right of lands rather than in right of blood. None of these questions are very material to the present investigation save in so far as they direct attention to possible earlier Whixley deeds. In the same spirit it has been found necessary to pay especial attention to the three religious houses holding land in Whixley, as both the original deeds of gift and the later dissolution records contain valuable clues.

† Thus Burke. But Henry Johnston (Bodley MS. Top. Yorks. c. 13) charges the hounds with what looks like a huntsman's hat, labelled as g(ules) or perhaps s(able); the former tincture seems implausible as it would cause the charge to merge with the field.

References

Andrews, J. H. 1974 The maps of the escheated counties of Ulster, 1609–10. *Proc. R. Irish Academy* **74**(C), 133–170 (with folding map and 3 plates).

Benzer, S. 1959 On the topology of the genetic fine structure. *Proc. natn. Acad. Sci. U.S.A.* **45**, 1607–1620.

Benzer, S. 1961 On the topography of the genetic fine structure. *Proc. natn. Acad. Sci. U.S.A.* **47**, 403–415.

Burke, J. & Burke, J. B. 1844 *Extinct and dormant baronetcies*. London: John Russell Smith.

Chibnall, A. C. 1965 *Sherington: fiefs and fields of a Buckinghamshire village*. Cambridge University Press.

Collins, F. (ed.) 1889 *Index of Wills in the York Registry* 1389–1514. *Yorkshire Archaeol. Soc.* (*Record Series*) VI. (This contains the reference for the will of Alice Banke, part of which is reproduced here as figure 21 through the kindness of the late Norah K. M. Gurney, M.A., until recently Director of the Borthwick Institute of Historical Research, University of York.)

Foster, J. (ed.) 1875 *The Visitation of Yorkshire, made in the years 1584–5, by R. Glover, &c.* London, privately printed. The extract from this Visitation shown in figure 22 is part of f. 152 in an unofficial manuscript copy or part-copy now in the Hull Local History Library of the Humberside Libraries (shelf-mark L929.2(4)). It should be noted that copies of Visitation records, whether printed or in MS., are never wholly to be relied upon, and this is no exception. I am grateful to Sir Anthony Wagner K.C.V.O., D.Litt., Garter Principal King of Arms, for kindly allowing me to collate this slightly misleading copy with Glover's own autograph MS. in the College of Arms. I have followed the latter MS. in preparing the text of this paper.

Fulkerson, D. R. & Gross, O. A. 1965 Incidence matrices and interval graphs. *Pacific J. Math.* **15**, 835–855.

Göransson, S. 1961 Regular open-field pattern in England and Scandinavian *Solskifte*. *Geografiska Annaler* **43**, 80–104.

Hiorns, R. W., Harrison, G. A., Boyce, A. J. & Küchemann, C. F. 1969 A mathematical analysis of the effects of movement on the relatedness between populations. *Ann. Hum. Genet.* **32**, 237–250.

Hodson, F. R. 1968 *The La Tène Cemetery at Münsingen-Rain*. Berne: Stämpfli.

Hodson, F. R., Kendall, D. G. & Tautu, P. (eds.) 1971 *Mathematics in the archaeological and historical sciences*. Edinburgh University Press. (*M.A.H.S.* elsewhere in this list.)

Jefferys, Thomas 1771–5 *A Survey of the County of Yorkshire* (Reprinted by H. Margary, Lympne Castle, Kent in 1973, with an Introduction by J. B. Harley and J. C. Harvey.)

Kendall, D. G. 1963 A statistical approach to Flinders Petrie's sequence dating. *Bull. int. statist. Inst.* **40**, 657–680.

Kendall, D. G. 1969 Incidence matrices, interval graphs, and seriation in archaeology, *Pacific J. Math.* **28**, 565–570.

Kendall, D. G. 1970 A mathematical approach to seriation. *Phil. Trans. R. Soc. Lond.* A **269**, 125–135.

Kendall, D. G. 1971*a* Abundance matrices and seriation in archaeology. *Zeitsch. f. Wahrscheinlichkeitstheorie* **17**, 104–112.
Kendall, D. G. 1971*b* Maps from marriages. *M.A.H.S.* pp. 303–318.
Kendall, D. G. 1971*c* Construction of maps from 'odd bits of information'. *Nature, Lond.* **231**, 158–159.
Kendall, D. G. 1971*d* Seriation from abundance matrices. *M.A.H.S.* pp. 215–252.
Kruskal, J. B. 1964*a*, *b* Multi-dimensional scaling, I and II. *Psychometrika* **29**, 1–27, 28–42.
Laxton, R. R. 1975 A measure of chronological significance in archaeological data. (To appear.)
Orwin, C. S. & Orwin, C. S. 1938 *The open fields*. Oxford: Clarendon Press.
Petrie, W. M. F. 1899 Sequences in prehistoric remains. *J. anthropol. Inst.* **29**, 295–301.
Petrie, W. M. F. 1901 *Diospolis Parva*. London: Egypt Exploration Fund.
Reed, H. 1946 *A map of Verona*. London: Jonathan Cape.
Sibson, R. 1972 Order-invariant methods for data analysis. *J. R. statist. Soc.* (B) **34**, 311–349.
Tobler, W. R., Mielke, H. W. & Detwyler, T. R. 1970 Geobotanical distance between New Zealand and neighbouring islands. *Bioscience* **20**, 537–542.
Tobler, W. R. & Wineberg, S. 1971 A Cappadocian speculation. *Nature, Lond.* **231**, 39–41.
Wilkinson, E. M. 1974 Techniques of data-analysis: Seriation theory. In *Technische und Naturwiss. Beiträge zur Feldarchäologie* (ed. I. Scollar), pp. 1–134. Bonn: Rheinisches Landesmuseum.

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

FIGURE 29. Aerial photograph of Whixley township (top), Whixley Park (with trees, and ancient selions, top left), *Wederbygate* running southwards (downwards as you look) from the Park, and *Buretrestubbes* flatt south of the crofts and to the right of *Wederbygate*. (Photograph by Professor J. K. S. St Joseph: RC8-AF274.)